Implementation by vote-buying mechanisms

By JON X. EGUIA AND DIMITRIOS XEFTERIS*

Vote-buying mechanisms allow agents to express any level of support for their preferred alternative at an increasing cost. Focusing on large societies with wealth inequality, we prove that the family of binary social choice rules implemented by well-behaved vote-buying mechanisms is indexed by a single parameter, which determines the importance assigned to the agents' willingness to pay to affect outcomes and to the number of supporters for each alternative. This parameter depends solely on the elasticity of the cost function near its origin: as this elasticity decreases, the intensities of support matter relatively more for outcomes than the supporters' count.
JEL: D72, D71, D61.
Keywords: implementation; social choice; vote-buying mechanisms; contests; quadratic voting; inequality.

Many collective binary choices —to reelect the incumbent or to elect the main challenger, to remain or to leave the European Union, to legalize or to ban same-sex marriage— generate fierce political competition between opposing interest groups. Supporters on each side take costly actions to promote their preferred alternative, and the likelihood that each alternative is collectively chosen increases in the total backing it receives. Such political contests often exhibit the following features: the ordinal preferences over outcomes and the willingness to incur costs to bolster a preferred alternative are the agents' private information; costly actions such as voting, campaigning, or donating to a campaign are voluntary; and agents have heterogeneous endowments of the relevant resources.[1]

How does the collective choice depend on the profile of preferences, the cost of supportive actions, and the inequality in resources? Under what conditions will collective decisions align with majoritarian principles or with the utilitarian welfare optimum? Can we describe the class of social choice rules that this tug-of-war political process ends up implementing?

We study a class of *vote-buying mechanisms* —a formal representation of the above contest— over binary alternatives, and we identify the choice rule that each of these

* Eguia: Department of Economics, Michigan State University, 110 Marshall-Adams Hall, 486 West Circle Drive, East Lansing, MI 48823, USA, eguia@msu.edu. Xefteris: Department of Economics, University of Cyprus, PO Box 20537, 1678 Nicosia, Cyprus, xefteris.dimitrios@ucy.ac.cy. We thank Tilman Börgers, Hari Govindan, Aniol Llorente-Saguer, Glen Weyl, Jacob Goeree, Nemanja Antic, Theodore Groves, John Ledyard, Yishu Zeng, Debasis Mishra, Debraj Ray, five anonymous reviewers, participants at the 2017 GLPETC, 2018 IMAEF, 2019 NSF/CEME Decentralization, EEA/ES Summer 2019, 2nd ETH Workshop on Democracy 2019, 2020 UC Riverside Theory conferences, the U. Michigan Theory Reading Group, and participants at talks at Columbia, Michigan State U., U. Illinois, U. Notre Dame, King's College, U. Mannheim, U. Maastricht, UNSW, U. Bath, U. Southampton, the Higher School of Economics, and the Indian Statistical Institute for their comments.

[1] Empirically, as the relative endowment of resources available to a social group or class increases, policies tilt toward favoring this group (Karabarbounis 2011).

mechanisms implements in large societies. Under a vote-buying mechanism, each agent can express her preference for an alternative by taking an action a in favor of this alternative, at a cost $c(a)$ that is increasing in the magnitude of action a. By taking a greater action, an agent can express any intensity of support, which —in an ordinal framework— is understood as the resources an agent is willing to give up to increase the probability that her preferred alternative is chosen. The collective choice is determined by the total sum of actions taken for each alternative.

Understanding how such mechanisms work is important not only due to their links with conventional political contests, but also to answer open institutional-design questions, such as how standard economic mechanisms would perform if applied to political contexts. While competitive markets for votes have been shown to implement the will of the most concerned individual (Casella, Llorente-Saguer, and Palfrey 2012), much less is known about the welfare effects of contest/auction-inspired procedures, like the ones that we study here.[2] Moreover, technological advances in the fields of encryption and data management have reinforced calls in mainstream media to embrace more sophisticated democratic procedures that endogenously redistribute political power among agents.[3] A number of organizations, such as Google (Hardt and Lopes 2015) and some political parties in Europe (Blum and Zuber 2016), have recently experimented with such procedures, and relevant innovations have even reached legislative institutions.[4] Theoretical results ought to anticipate these developments —or at least, to anticipate their expansion into the public arena— in order to inform any decisions on institutional reform, and to properly assess the effects of wealth inequality over policy and social welfare.

Our main contribution is to characterize the collective choice rule that is implemented in large societies by each vote-buying mechanism within a broad class, in the presence of preference and wealth heterogeneity. As it turns out, the family of implementable rules coincides with an important family, named after Bergson ([neé Burk] 1936 and 1938). Choice rules (or, welfare optima) in this family are neatly indexed by a single positive parameter ρ, which measures the influence of the agents' marginal willingness to pay, vis-à-vis the number of supporters of each alternative, over the collective choice. An agent's (marginal) willingness to pay is specifically defined as the agent's marginal rate of substitution of wealth for the likelihood of collectively choosing her preferred

[2]Traditionally, groups face a trade-off between political equality and the expression of strength of preferences. Political equality is arguably best protected by simple one-man one-vote procedures (Locke 1689; May 1952; Dahl 1989). But a binary choice determined by majority preferences over the two alternatives aggregates preferences too coarsely: it ignores preferences over trade-offs between the collective choice and other resources, so a majority of agents unwilling to sacrifice much for their cause prevails over a minority, no matter how much the minority is willing to sacrifice to prevail. This coarseness has been a constant source of criticism of one-person one-vote procedures, as acknowledged by Dahl (1989).

[3]See, for instance: John Richardson and Jed Emerson "eDemocracy: An Emerging Force for Change," in *Stanford Social Innovation Review* January 25, 2018; or Peter Coy "A new way of voting that makes zealotry expensive" in *Bloomberg Businessweek* May 1, 2019; or Rana Foroohar "Digital tools can be a useful bolster to democracy," in the *Finantial Times* February 1, 2020.

[4]In May 2019, the Democratic majority caucus in the Colorado House of Representatives used a vote-buying method with quadratic costs to select proposals for over $40 million dollars in discretionary spending. See Brian Eason, "$120 million in requests and $40 million in the bank. How an obscure theory helped prioritize the Colorado budget." *The Colorado Sun* May 28, 2019.

alternative.[5] Each individual's willingness to pay is taken to the power ρ to determine the weight of this individual preference over the collective decision, and the optimal alternative is the one with the greatest sum of such weights across all agents. If we label the two alternatives A and B, and we let N_A and N_B respectively denote the sets of agents who strictly prefer alternative A and alternative B, then the choice rule with parameter value ρ chooses alternative A over B if

$$\sum_{i \in N_A} |MRS^i|^\rho > \sum_{i \in N_B} |MRS^i|^\rho, \tag{1}$$

where MRS^i denotes agent i's marginal rate of substitution of wealth for probability of collectively choosing her preferred alternative.

A couple of observations are in order. First, low values of ρ indicate that willingness to pay matters little, and that the optimal alternative is mostly determined by the number of supporters of each alternative (the lower limit of this class, $\rho \to 0$, converges to majoritarianism); whereas, high values of ρ correspond to cases in which willingness to pay ("how much support?" rather than "how many supporters?") matters most. Second, if agents' marginal utility from wealth is not constant, the aggregate willingness to pay for each alternative depends not only on the distribution of preferences, but also on the distribution of wealth.

Denote the limit cost elasticity $\lim\limits_{a\to 0} \frac{c'(a)a}{c(a)}$ (limit of marginal cost over average cost) of a cost function c by $\kappa(c)$. We prove that every vote-buying mechanism under consideration whose cost function c has limit cost elasticity $\kappa(c) > 1$ asymptotically implements the social choice rule indexed by parameter $\rho = \frac{1}{\kappa(c)-1}$ as society becomes arbitrarily large. Note that we consider a general class of cost functions, including ones that do not even scale like power or other polynomial functions, and we still find that the welfare optima that they implement are in the class of power optima given by Expression (1). Essentially, the present analysis uncovers that vote-buying mechanisms admit a welfare-relevant ordering on the basis of a single intuitive parameter: as the elasticity of the cost function near its origin decreases, agents' willingness to pay matters relatively more for the determination of the outcome than the number of supporters of each alternative.

Moreover, the cost of individual equilibrium actions converges to zero as the society grows large. That is, the above result holds even if agents face tight budget constraints. This feature is in contrast with those of mechanisms for implementation in dominant strategies (Vickrey 1961, Clarke 1971, Groves 1973), or of a competitive market for votes (Casella, Llorente-Saguer, and Palfrey 2012), which require individuals to be able to afford substantial payments.

An example. For the sake of presenting some partial intuition behind the result, consider a large society such that a share n_A of the population supports A, and a share $n_B > n_A$ supports B. According to simple majority voting, B wins independently of each agent's willingness to pay to

[5] Bergson's aggregation rules are defined over profiles of individual values that are often interpreted as cardinal utilities (Moulin 1988). These rules are also applicable in an ordinal framework, as long as the individual values take on an alternative interpretation.

affect the outcome. Whereas, if society decides according to a vote-buying mechanism, then the outcome depends on the interplay among the mechanism's cost function, the agents' attitudes towards the social choice, and their wealth.

Assume that each citizen's utility is given by $v \times 1_A + u(w)$ if she ends up with wealth w, where 1_A is equal to one if A wins, and zero otherwise, and u is a (weakly) concave function. A positive valuation ($v > 0$) indicates a supporter of A, and a negative valuation ($v < 0$) indicates a supporter of B. Furthermore, suppose that all supporters of A [B] have a common valuation $v_A > 0$ [$v_B < 0$] and initial wealth w_A [w_B], and that all agents share a common belief that the marginal pivotality is exogenous, i.e., that each vote/unit of the costly action for A or B increases the probability that this alternative is chosen by $\phi > 0$.[6]

A supporter of alternative A who takes an action $a_A > 0$ in favor of A has marginal gain equal to $\phi \times v_A$ and marginal cost equal to $c'(a_A) \times u'(w_A - c(a_A))$. This individual maximizes her expected utility by choosing $a_A > 0$ such that $\phi \times v_A = c'(a_A) \times u'(w_A - c(a_A))$. Similarly, a supporter of B, chooses $a_B > 0$ such that $-\phi \times v_B = c'(a_B) \times u'(w_B - c(a_B))$. As the pivotality becomes small —for instance, when the society grows large— we have that the optimal actions of both individuals (and hence their costs) converge to zero, so the shape of the cost function matters only near its origin. Specifically, it follows that

$$\text{(2)} \qquad \frac{|MRS^A|}{|MRS^B|} \approx \frac{c'(a_A)}{c'(a_B)},$$

where $|MRS^A| = |v_A/u'(w_A)|$ and $|MRS^B| = |v_B/u'(w_B)|$. If the cost function has a constant elasticity equal to $\kappa > 1$ —i.e., if it is of the form $c(a) = |a|^\kappa$, for some $\kappa > 1$— the above equation can be written as $(|MRS^A|/|MRS^B|)^{\frac{1}{\kappa-1}} \approx \frac{a_A}{a_B}$. Notice that the sum of actions in favor of A, $n_A \times a_A$, is larger than the sum of actions in favor of B, $n_B \times a_B$, if $\frac{a_A}{a_B} > \frac{n_B}{n_A}$, or, $(|MRS^A|/|MRS^B|)^{\frac{1}{\kappa-1}} > \frac{n_B}{n_A}$. Since the latter can be written as $n_A \times |MRS^A|^{\frac{1}{\kappa-1}} > n_B \times |MRS^B|^{\frac{1}{\kappa-1}}$, and the elasticity near the origin, $\kappa(c)$, is equal to the constant elasticity κ, the outcome aligns, generically, with the choice rule in the above family characterized by $\rho = \frac{1}{\kappa(c)-1}$.[7] In-

[6] In our theory, pivotality is endogenous and varies across actions. But (as we prove), in large societies the marginal pivotalities become approximately the same for all agents, and the intuition illustrated in this example holds if rational agents correctly compute their endogenous pivotality. The exogenous-pivotality approach has been followed as well by Lalley and Weyl (2018). The concave shape of the utility from wealth captures the fact that wealthy individuals might be willing to sacrifice more resources for the social choice than poorer ones. Taking no action in favor of any alternative is free. Our analysis is valid independently of whether or not the undertaken costs are redistributed among the agents.

[7] Alternative A is more likely to win if the A supporters are many, their valuations and wealth are high, and —conditional on their willingnesses to pay being higher than those of the B supporters— if the elasticity of the cost

terestingly, the same result holds even if the cost elasticity is not constant (i.e., if c is not a power function), establishing a two-way mapping between a broad class of vote-buying mechanisms and a diverse, but tightly characterized, family of social choice rules.[8]

The impact of willingness to pay on the social choice decreases with $\kappa(c)$: if $\kappa(c)$ converges to one (its lowest bound), A wins if $|MRS^A| > |MRS^B|$ and B wins otherwise —i.e., the supporters' count does not matter at all— and if $\kappa(c)$ takes arbitrarily large values, A wins if $n_A > n_B$ and B wins otherwise —i.e., willingness to pay does not matter at all.

We depart from most existing approaches by relaxing the simplifying assumption of a linear utility over wealth —i.e., the absence of wealth effects— to study collective choice by agents with weakly diminishing marginal utility of wealth.[9] That is, the agents' willingness to pay for their preferred collective choice increases in wealth irrespective of their attitudes towards the collective choice. This more general setup poses additional analytical challenges compared to the standard linear one, but it allows us to provide an accurate description of the welfare effects of vote-buying mechanisms.

A positive, descriptive interpretation of our main result is that if the political competition in a society is a contest such that the process that determines political outcomes amounts to a vote-buying mechanism, then this society's choices follow the power welfare optimum that corresponds to such vote-buying mechanism, whether or not this rule aligns with society's values. Since these mechanisms aggregate willingnesses to pay, they essentially overweigh the preferences of the wealthy, who are willing to pay more than poorer agents to influence the choice. Hence, economic inequality undermines political equality.[10] At the same time, individuals who intrinsically care more about the collective decision are able to exert more influence on the outcome than others who care less. Our work identifies the exact way in which true preferences and wealth effects interact in shaping collective outcomes if the society engages in such a contest, and it highlights the role of the elasticity of the relevant cost function in the aggregation process.[11]

function is low. For instance, if $n_A = \frac{1}{4}$, $n_B = \frac{3}{4}$, $v_A/u'(w_A) = 10$, and $v_B/u'(w_B) = -1$ —i.e., the supporters of A are few but either care a lot about the election or they are very wealthy (or both)— A wins if the elasticity of the cost function, κ, is below 3.1, and B wins otherwise.

[8] Since the reasoning supporting this more general finding requires more subtle formal arguments, we present it after our main theorem.

[9] Kazumura, Mishra, and Serizawa (2020) study dominant strategy incentive compatible mechanisms without assuming that preferences are quasilinear, and they survey the "*small*" literature on mechanism design without quasilinearity.

[10] Indeed, assuming that $u(w) = \ln w$, and given the 20th and 80th percentile total incomes of respectively \$28,000 and \$142,000, it follows that $u'(28,000)/u'(142,000) = \frac{142}{28}$. So, agents in the 80th percentile would be, by Expression (2), about 5 (1.7) times as influential as those at the 20th percentile if $c(a) = |a|^2$ ($c(a) = |a|^4$). These income values are from the distribution of US household total incomes in 2019, as reported by Semega, Kollar, Shrider and Creamer (2020).

[11] From a normative standpoint, one could consider the institutional design problem in a large society in which the designer wishes to implement a given social choice rule using a mechanism that does not require large transfers from citizens, is budget-balanced, and does not use detailed information regarding voters' preferences that may be unavailable to the mechanism (Matsushima 2008). Our results imply that a society whose value system is compatible with any given power welfare optimum over willingnesses to pay —and only such a society— could asymptotically implement its desired social choice rule by using the appropriate vote-buying mechanism.

We next present our model and main result in Section I. We review the literature in Section II. While we regard our work as primarily basic research and we do not aspire to address all the practical issues that might arise in real settings, we discuss in our concluding remarks (Section III) a number of concerns that would likely be salient in applications, including caveats regarding exogenous costs of participation, voters' information and sophistication, and collusion. A mathematical formalization and all proofs are in the Online Appendix.

I. The Theory

We first describe the model, and then we present and explain our result.

A. *The agents*

A society with n agents must collectively choose one of two alternatives A and B. Agents care about their wealth and about which alternative is chosen. Each agent i's ordinal preferences over combinations of her wealth outcome w_i^O and the chosen alternative are representable by a utility function that takes the form

$$u(w_i^O) + v_i \text{ if } A \text{ is chosen and} \\ u(w_i^O) \text{ if } B \text{ is chosen,} \tag{3}$$

where u is a strictly increasing, weakly concave, bounded utility function over wealth, and v_i is i's "valuation" over alternative A. The related literature typically assumes that u is linear, and in order to relate to this literature, we allow for and discuss this special case. However, the main substantive interest lies in applications in which marginal utility over wealth is decreasing (u is strictly concave), so we focus on the case with a strictly concave u, and on the implications of wealth inequality in this case.

Each agent i is characterized by her initial wealth w_i and by her valuation v_i. Agent i's type $\theta_i = (w_i, v_i)$ is privately and independently drawn from $[w_{\min}, 1] \times [-\gamma, \gamma]$ according to a smooth probability measure P with positive density p, where $w_{\min} > 0$ and $\gamma > 0$. Note that we allow for any correlation between wealth and valuations under distribution P. We assume that agents are rational expected utility maximizers.

B. *The mechanisms*

A vote-buying mechanism is defined by a cost function c. Agent i chooses action a_i and pays $c(a_i)$. Positive actions are in favor of A and negative actions promote alternative B. The social decision is stochastic: the probability that alternative A is chosen is $G\left(\sum_{i=1}^n a_i\right)$, where the outcome function $G : \mathbb{R} \longrightarrow [0, 1]$ is strictly increasing in the sum of actions. Each payment $c(a_i)$ is redistributed equally among all other agents. We consider a class of admissible (well-behaved) vote-buying mechanisms and an environment with the following features:

i. Neutrality: The cost function satisfies $c(a_i) = c(-a_i)$ for any $a_i \geq 0$, and the probability that A is chosen given $\sum_{i=1}^n a_i$ is the same as the probability that B is chosen given

$\sum_{i=1}^{n} -a_i$ for any vector of actions $(a_1, ..., a_n)$.[12]
ii. Responsiveness: The probability that A is chosen converges to one as $\sum_{i=1}^{n} a_i$ becomes arbitrarily large.
iii. Voluntary participation: There is no cost to taking a null action, i.e., $c(0) = 0$.
iv. Smoothness: The cost function is continuously differentiable everywhere and twice continuously differentiable for any non-zero action, and the elasticity of the cost function (the marginal cost over average cost) has a bounded limit at zero. Further, the outcome function G is twice continuously differentiable and its first derivative denoted g has vanishing but "fat" tails. Formally, $\lim_{x\longrightarrow\infty} g(x) = 0$ (vanishing tails), and for any ε sufficiently close to zero, $\lim_{x\longrightarrow\infty} \frac{g'(x+\varepsilon)}{g(x)}$ is bounded ("fat" tails).
v. Affordability: For any strictly positive willingness to pay per unit, some strictly positive quantity of action can be purchased at that price, i.e., $c'(0) = 0$.
vi. Monotonicity: Greater actions are more expensive, i.e., $c'(a_i) > 0$ for any $a_i > 0$.
vii. Convexity: The limit of the elasticity of the cost function as the action goes to zero is greater than one.

Features i, ii, iii, v, and vi are substantive features with normative appeal. Feature vii coincides with a standard definition of strict convexity for cost functions that are power functions with functional form $c(a_i) = |a_i|^{\rho}$ (for these functions it is equivalent to $\rho > 1$).[13] Smoothness is a technical assumption: its continuity requirements guarantee existence of equilibrium, and as we discuss below, the assumption of "fat" tails is sufficient, but not necessary, for our main result. Any polynomial cost function, and many non-polynomial ones, satisfy these conditions.[14]

With regard to the outcome function G, it can take, for instance, the functional form $\frac{e^{\zeta x}}{1+e^{\zeta x}}$ for some $\zeta > 0$; if so, as ζ increases, the outcome function becomes arbitrarily close to the deterministic step function. Or it can take other functional forms, such as those of the logistic cumulative distribution, or of any Student-t cumulative distribution.[15]

Agents choose an action with common knowledge about the probability distribution P over types, and about the mechanism, and with private information about their type. So, given a probability distribution P, and given a mechanism, a pure strategy s is a function from types to actions.[16] Our solution concept is symmetric Bayes Nash Equilibrium.

[12] We relax neutrality, considering mechanisms that favor one of the two alternatives, in an extension in the Online Appendix, Section 4.

[13] For more general functions, a limit elasticity greater than one implies strict convexity, and strict convexity implies that this limit elasticity is at least one, so feature vii only rules out the corner case in which strict convexity limits to weak convexity.

[14] We formalize all technical details in the Online Appendix.

[15] In the Online Appendix we generalize this framework, allowing for the mapping G from the net sum of actions to the probability over the social choice to be unknown to the designer, so that she only knows that this mapping belongs to an arbitrary class of mappings that satisfy the requirements in features i, ii, and iv. In this more general environment, a cost function c must obtain the desired results for any outcome function in this class. The case with a fixed mapping known to the planner, or chosen by the planner, is a special, simpler case.

[16] Admissible vote-buying mechanisms are "bounded" in the sense of Jackson (1992), but they are not "strategically simple" in the sense of Börgers and Li (2019).

C. *The social choice rules and the implementation concept*

A social choice rule *SCR* maps any vector of types $\theta_{N^n} = (\theta_1, ..., \theta_n)$ to the alternative or alternatives $SCR(\theta_{N^n}) \subseteq \{A, B\}$ that are deemed normatively desirable or socially preferred according to this rule. To make all its collective decisions according to rule *SCR*, a society needs to follow a choice mechanism such that for any collective choice problem —characterized by the probability distribution P from which preferences are drawn— for every realization of types θ_{N^n}, and for every equilibrium of the game induced by the mechanism, the equilibrium outcome is an alternative chosen by rule *SCR*.

The mechanism cannot condition on the vector θ_{N^n}, because valuations are private information, nor can it condition on the probability distribution P. Our preferred interpretation of the inability to condition on P is that society's institutions (its mechanism) are adopted at an earlier constitutional stage, and agents only learn the defining parameters (the distribution P) of the specific collective choice problem at hand after the mechanism is in place. If so, to attain its desired outcome in any collective choice problem it may face, society needs a mechanism that works across different distributions. In the absence of a mechanism that always works, we settle for finding a mechanism that almost always works for large societies. To find it, we use an asymptotic notion of (full) implementation.

DEFINITION 1: *We say that a mechanism c implements the social choice rule SCR for a given probability distribution P if, for any sufficiently large society size: i) a symmetric equilibrium exists, and ii) in any symmetric equilibrium, the ex-ante probability that the alternative chosen in this equilibrium is an alternative chosen by rule SCR is arbitrarily close to one.*

We say that a mechanism c generically implements rule SCR if c implements SCR for almost all smooth distributions with positive density.

The mathematical formalization of this notion of asymptotic full implementation for a given probability distribution P is in the Online Appendix (Definition 4). The idea is that while there may be some type profile realizations for which the probability of choosing an undesirable alternative is substantial, such type profiles are vanishingly rare if society is sufficiently large. Thus, in a sufficiently large society, and in any equilibrium, the probability that the choice is as desired by the implemented rule is arbitrarily close to one. The genericity in this definition refers to the shape of the distribution P over types; it means that for almost any such shape, the result holds, or in other words, that the choice problems for which it may not hold are knife-edged. We precisely define the topological sense of genericity we use here in the Online Appendix (Definition 5).

Our main result is to characterize the class of social choice rules that are generically implementable by vote-buying mechanisms. Let N_A denote the set of agents who prefer A, and N_B the set of agents who prefer B. We find that vote-buying mechanisms can generically implement any Bergson social choice rule defined over marginal willingness to pay. Any such rule chooses A if

$$\sum_{i \in N_A} \left| \frac{v_i}{u'(w_i)} \right|^{\rho} > \sum_{i \in N_B} \left| \frac{v_i}{u'(w_i)} \right|^{\rho}, \tag{4}$$

and chooses B if the inequality is reversed, for some parameter $\rho > 0$ that defines the rule and captures the importance of willingness to pay.

Let SC_ρ denote the rule in this family with parameter ρ. Each rule SC_ρ first weighs each agent's valuation v_i by the inverse of the agent's marginal utility of wealth. The fraction $\frac{v_i}{u'(w_i)}$ is the agent's marginal rate of substitution between the social decision and wealth, or, put differently, the agent's marginal willingness to pay to change the decision from B to A. Rule SC_ρ then takes all the marginal willingnesses to pay to power ρ, where parameter ρ captures the importance of the willingness to pay. The limit case with parameter $\rho = 0$ corresponds to majoritarianism; the special case with $\rho = 1$ is the rule that chooses the alternative that maximizes aggregate willingness to pay; and the limit case with parameter $\rho = \infty$ corresponds to choosing the alternative preferred by the agent with the greatest willingness to pay.[17]

D. The main result

For each admissible vote-buying mechanism c, let $\kappa(c)$ denote the limit as a goes to zero of the elasticity of c, that is, $\kappa(c) = \lim_{a \longrightarrow 0} \frac{ac'(a)}{c(a)}$. This elasticity alone determines the asymptotic results obtained under each admissible vote-buying mechanism

THEOREM 1: *Any admissible vote-buying mechanism c generically implements rule* $SC_{\frac{1}{\kappa(c)-1}}$.

That is, in large societies, vote-buying mechanisms implement the class of social choice rules of the form given by Expression (4). Given that c generically implements $SC_{\frac{1}{\kappa(c)-1}}$, setting $\frac{1}{\kappa(c)-1} = \rho$ and solving for $\kappa(c)$, it follows that for any $\rho > 0$, any c with $\kappa(c) = \frac{1+\rho}{\rho}$ generically implements SC_ρ. Rules that diverge from those given by Expression (4) with more than vanishing probability as society gets sufficiently large cannot be implemented by any admissible vote-buying mechanism. Since for any $\kappa \in (1, \infty)$, the limit elasticity of a power cost function $c(a) = |a|^\kappa$ is $\kappa(c) = \kappa$, a simplifying corollary follows.

COROLLARY 1: *If a social choice rule is generically implemented by an admissible vote-buying mechanism c with limit elasticity* $\kappa(c) = \kappa$*, then this rule is generically implemented by the power vote-buying mechanism* $\hat{c}(a) = |a|^\kappa$*. In particular, for any* $\rho > 0$, SC_ρ *is generically implemented by* $\hat{c}(a) = |a|^{\frac{1+\rho}{\rho}}$.

[17] Casella, Llorente-Saguer, and Palfrey (2012) have shown that in a quasilinear environment, the equilibrium outcome with a competitive (Walrasian) market for votes coincides with the outcome chosen by choice rule SCR_∞. Dekel, Jackson, and Wolinsky (2008) propose a related theory of vote-buying, in which two leaders buy a majority of votes at an all-pay auction.

Therefore, the class of social choice rules generically implementable by the class of admissible vote-buying mechanisms is the same as the class generically implementable by the subclass of power function vote-buying mechanisms.

E. The underlying intuition

The intuition behind our characterization result is as follows. First note that the continuity of the utility function, of the cost function, and of the mapping from actions to outcomes, together with the other assumptions in the model, allows us to invoke existence results from Milgrom and Weber (1985) and Reny (1999), which, combined, guarantee that a symmetric Bayes Nash equilibrium exists (see lemmas 1 and 2 in the Online Appendix). Let $\Pr[A|a_i, s]$ denote the probability that the outcome is A given that agent i plays action a_i and all other agents follow strategy s. With this notation, in an equilibrium in which agents play strategy s, each individual agent i with type $\theta = (w, v)$ chooses equilibrium action $s(\theta)$ such that:

(5)
$$v \times \left.\frac{\partial \Pr[A|a_i, s]}{\partial a_i}\right|_{a_i = s(\theta)} = c'(s(\theta)) E\left[u'\left(w - c(s(\theta)) + \frac{\sum_{j=1}^{n} c(s(\theta_j)) - c(s(\theta))}{n-1}\right)\right].$$

The left-hand side of Equality (5) is the marginal benefit, which is equal to the agent's valuation multiplied by her marginal probability of affecting the outcome by increasing her action, or "marginal pivotality." The right-hand side is the marginal (utility) cost, which is equal to the marginal monetary cost $c'(s(\theta))$ multiplied by the expected marginal utility of wealth, where the expectation is over the realization of the types of other agents.

As society grows large (as n diverges to infinity), the probability that an extra vote changes the social decision vanishes. For any given action, marginal pivotality —and with it, the marginal benefit— then vanish as well for every agent. It follows that in equilibrium, marginal costs —and hence equilibrium actions— must converge to zero as well, so the expectation over the marginal utility of wealth on the right-hand side of Equation (5) converges to simply $u'(w)$. Since equilibrium actions converge to zero, asymptotically only the shape around zero of the cost function c matters for equilibrium behavior.[18]

Further, as n diverges to infinity and equilibrium actions converge to zero, the ratio of equilibrium marginal pivotalities of any two different types converges to one.[19] This is a key intermediate result: it implies that the ratio of the marginal benefit for an agent with type $\theta = (w, v)$ over the marginal benefit for an agent with type $\hat{\theta} = (\hat{w}, \hat{v})$ converges

[18] Note that the convergence of actions to zero is slower than that of $\frac{1}{n}$, so that the expectation over the magnitude of the net sum of actions $\left|\sum_{i=1}^{n} a_i\right|$ diverges to infinity.

[19] The assumption that types are independently drawn is necessary at this step of the proof. With correlated types, each agent's belief about the behavior of the rest of the society, and hence about her marginal pivotality, depends on her own type.

to simply $\frac{v}{\hat{v}}$.[20]

Since the ratio of marginal (utility) costs converges to $\frac{c'(s(\theta))}{c'(s(\hat{\theta}))}\frac{u'(w)}{u'(\hat{w})}$, if we divide the equilibrium Equation (5) for type θ over the same equation for type $\hat{\theta}$, and we rearrange terms, we obtain the asymptotic condition

$$\frac{|MRS_\theta|}{|MRS_{\hat{\theta}}|} \approx \frac{c'(|s(\theta)|)}{c'(|s(\hat{\theta})|)}, \tag{6}$$

where, for each $\theta = (w, v)$, $|MRS_\theta| = |\frac{v}{u'(w)}|$ is type θ's marginal rate of substitution of wealth for probability of collectively choosing her preferred alternative, and $|MRS_{\hat{\theta}}| > 0$.

As illustrated in the introduction, if c is a power function $c(a) = |a|^\kappa$, for some $\kappa > 1$, then our result directly obtains. But what if c is not a power function? To answer this question, we first observe that for any admissible cost function c with limit elasticity $\kappa(c) > 1$ we have that

$$\frac{c'(y)}{c'(\hat{y})} \approx \left(\frac{y}{\hat{y}}\right)^{\kappa(c)-1} \tag{7}$$

if both $y > 0$ and $\hat{y} > 0$ become arbitrarily small.[21] Since the equilibrium actions, $s(\theta)$ and $s(\hat{\theta})$, of any pair of types $(\theta, \hat{\theta})$ vanish to zero as the number of agents diverges, it follows that in a large society,

$$\frac{c'(|s(\theta)|)}{c'(|s(\hat{\theta})|)} \approx \left(\frac{|s(\theta)|}{|s(\hat{\theta})|}\right)^{\kappa(c)-1}. \tag{8}$$

Combining Condition (6) with Condition (8), we get

$$\frac{|s(\theta)|}{|s(\hat{\theta})|} \approx \left(\frac{|MRS_\theta|}{|MRS_{\hat{\theta}}|}\right)^{\frac{1}{\kappa-1}}. \tag{9}$$

[20] The assumption of fat tails of g is not necessary for our result to hold, but we use it to ease this step of the proof. If g has fat tails, the ratio $g'(x + \varepsilon)/g(x)$ is bounded, which implies that for any sufficiently small action a_i^*, $g\left(\sum_{j\neq i} a_j + a_i^*\right)\Big/ g\left(\sum_{j\neq i} a_j\right)$ is arbitrarily close to one (see the Online Appendix). Numerical examples — available from the authors— show that the ratios of pivotalities converge also to one with an alternative g without fat tails (for instance, with g taking the functional form of a Normal density function).

[21] For any fixed $\hat{y} > 0$, it is easy to check that $\ln \frac{c'(y)}{c'(\hat{y})} / \ln \frac{y}{\hat{y}}$ takes the same value as $\frac{c''(y)y}{c'(y)}$ when it is maximized or minimized with respect to y, and also asymptotically as y approaches zero or $\hat{y}$. Moreover, since $\frac{c''(y)y}{c'(y)}$ converges to $\kappa(c) - 1$ when y vanishes to zero (indeed, $\kappa(c) = \lim_{y\to 0} \frac{c'(y)y}{c(y)} = \lim_{y\to 0} \frac{c''(y)y}{c'(y)} + 1$), $\ln \frac{c'(y)}{c'(\hat{y})} \Big/ \ln \frac{y}{\hat{y}}$ must also approach $\kappa(c) - 1$ when both y and $\hat{y}$ become arbitrarily small, so Condition (7) follows. A fully fledged argument can be found in the proof of Lemma 11 in the Online Appendix.

That is, the equilibrium actions of types θ and $\hat{\theta}$ become proportional to the willingness to pay raised to a power that depends solely on the limit cost elasticity.[22] Therefore, the sign of the net vote total aligns, generically, with the Bergson choice rule $SC_{\frac{1}{\kappa(c)-1}}$ as defined in Expression (4).[23] Notice that the above argument does not depend on the specific form of the outcome function —i.e., how the net vote total maps to the probability that A is chosen— nor on other characteristics of the cost function besides its elasticity near the origin. For any admissible vote-buying mechanism, vote acquisitions become proportional to power functions (determined by the limit cost elasticity) of the willingnesses to pay. Hence, the implemented social choice rules cannot substantially differ from the described Bergson ones for generic probability measures.[24]

F. *Linear preferences over wealth*

Under the restriction to a quasilinear preference subdomain, the utility over wealth u is linear, the wealth terms in Expression (4) cancel out, and the class of social choice rules generically implementable by admissible vote-buying mechanisms simplifies to those that, for some $\rho > 0$, choose A if and only if $\sum_{i \in N_A} |v_i|^\rho \geq \sum_{i \in N_B} |v_i|^\rho$. This class of Bergson rules allows for a normatively more compelling interpretation, since they aggregate individual valuations unmediated by wealth (see Roberts 1980).

The generic implementation of utilitarianism by a quadratic vote-buying mechanism under quasilinearity (Lalley and Weyl 2019) follows as a corollary for the special case $\kappa(c) = 2$.[25] Goeree and Zhang (2017) and Lalley and Weyl (2019) provide the following intuition for the asymptotic implementation of utilitarianism by a quadratic vote-buying mechanism: since the marginal benefit of acquiring votes is almost constant in the quantity of votes acquired, agents infer that the marginal benefit of acquiring votes is approximately linear in their valuation. Given a mechanism $c(a)$ with derivative $c'(a)$ that is linear in a, agents equate marginal benefit and marginal cost by acquiring votes

[22] Condition (8) holds even if different players of the same type are allowed to use different actions, as long as these actions are vanishing. The ratio of marginal pivotalities converges to one and, hence, Condition (6) holds in every equilibrium (symmetric or asymmetric) with vanishing actions. Since in large societies individual actions converge to zero also in asymmetric equilibria (if these exist), it follows that all equilibria are asymptotically symmetric (i.e., in every sequence of equilibria, the ratio of the actions of two players of the same type converges to one as society grows large).

[23] Because the shape of utility function u only matters through the derivative of u at wealth w, our result generalizes to an environment in which agents' utility functions are heterogeneous, privately observed, and independently and identically drawn from a finite set of admissible utility functions. In this more general case, since the type of each agent i is given by $\theta_i = (u_i, w_i, v_i)$, alternative A is implemented (asymptotically) if $\sum_{i \in N_A} \left|v_i/u_i'(w_i)\right|^{\frac{1}{\kappa(c)-1}} > \sum_{i \in N_B} \left|v_i/u_i'(w_i)\right|^{\frac{1}{\kappa(c)-1}}$, and B wins otherwise.

[24] Our proof strategy is valid except for the knife-edge case in which the expectation of $\left(v_i/u'(w_i)\right)^{\frac{1}{\kappa(c)-1}}$ is zero. We have no reason to believe that implementation of the corresponding Bergson rule fails even in such cases. In fact, for the most salient subset of these knife-edge cases —the "neutral" measures, in the sense that for any given w, the density of valuation v is symmetric around zero— we can show that each admissible vote-buying mechanism c with elasticity $\kappa(c)$ asymptotically implements social choice rule $SC_{\frac{1}{\kappa(c)-1}}$ in neutral Bayes Nash equilibria; "neutral" in the sense that if agent with type (w, v) plays a, then agent with type $(w, -v)$ plays $-a$.

[25] Lalley and Weyl's (2019) model and our own are not nested: while they also use a continuous —hence, stochastic— outcome function, they do not assume fat tails. This assumption simplifies the proof and allows one to work with general cost functions and non-quasilinear preferences.

approximately in proportion to their valuation, which leads to utilitarian efficiency.

This intuition can also be used to easily identify the choice rules implemented by other (non-quadratic) power cost mechanisms. However, it does not provide helpful insights regarding the welfare effects of more general cost mechanisms. Consider, for example, a non-power mechanism such as $\hat{c} \in C$ defined by

$$\hat{c}(a) = (\cos(|a|) - 1)(2\ln(|a|) - 3) = 1, \text{ for any } a \in [-2, 2].$$

Notice that $\hat{c}(a)$ and $c(a) = |a|^2$ are generically unequal. In fact, $\lim_{a \longrightarrow 0^+} \frac{\hat{c}'(a)}{c'(a)} = +\infty$, ($c'$ converges to zero arbitrarily faster than $\hat{c}'$). The marginal cost $\hat{c}'(a)$ is a (cumbersome) trigonometric function, so according to the above approximate reasoning, one would not infer that mechanism $\hat{c}$ also implements utilitarianism; and yet, it does (indeed, it is easy to check that $\kappa(\hat{c}) = 2$).[26] Put differently: implementation of the utilitarian optimum does not hinge on the marginal cost being linear but, rather, on the limit cost elasticity being equal to two.

While the case of linear utility over wealth is prominent in the related literature, we regard as substantively more relevant the case in which agents' marginal utility from wealth is decreasing, so u is strictly concave. With strictly concave utilities over wealth, since richer agents are willing to pay more for a given valuation, and since vote-buying mechanisms aggregate willingnesses to pay, it follows that these mechanisms overweigh the preferences of rich agents in their aggregation of valuations. This link between economic power and political power violates the normative principle that justice (and power) should be distributed differently across separate spheres of social interactions (Walzer 1983).[27]

II. Literature Review

Vote-buying mechanisms are rooted in three theoretical foundations: first, the theories of political competition as a costly contest; second, the literature on electoral competition with heterogeneous voting costs; and third, classic implementation theory.

The first and most direct motivation to study vote-buying mechanisms is to interpret

[26] We observe that both the quadratic and the trigonometric mechanism not only implement the utilitarian optimumim asymptotically, but often align with it with high probability even in finite societies. For instance, if the probability that alternative A is chosen given $x = \sum_{i=1}^{n} a_i$ is $\frac{e^x}{1+e^x}$, and P assigns probability $\frac{3}{4}$ to valuations in a small neighborhood of -1 and probability $\frac{1}{4}$ to valuations in a small neighborhood of 10, and the utility over money is linear, then for any $n \geq 10,000$, the probability of implementing the utilitarian optimum is greater than 99.9% if society uses a vote-buying mechanism with a quadratic cost function, and greater than 99.5% if society uses the trigonometric cost function $\hat{c}$; (whereas, it is less than 0.01% under simple majority voting). We further illustrate the convergence toward our asymptotic result through numerical examples with alternative power cost functions and preferences in Section 5 of the Online Appendix.

[27] In theory, wealth effects over the collective choice could be eliminated if the cost function were allowed to depend on the wealth of each agent. For instance, if the cost of taking action a for an agent with wealth w were $c_w(a) = c(a)/u'(w)$, then the outcome would coincide (asymptotically) with a Bergson choice rule defined solely over valuations. Hybrid systems, such that the net sum of actions matters only if neither alternative is supported by a sufficiently large supermajority of agents, could also impose bounds on the effects of economic inequality on collective decisions (examples of such mechanisms are available from the authors). Of course, these interventions are not always plausible in settings of applied interest.

them as a positive description of a competitive political process in which supporters and opponents of any given candidate or policy take costly actions to advance their cause, resulting in a collective choice as a function of these costly actions.[28] Under this interpretation, the "mechanism" or process that maps actions to costs and outcomes is an exogenously given societal characteristic. Such mechanisms have been extensively studied in applications to lobbying (Dixit, Grossman, and Helpman 1997; Campante and Ferreira 2007; Kang 2016), political competition over abstract issues or ethnic interests (Esteban and Ray 1999 and 2011), or two-candidate electoral competition (Stromberg 2008; Bouton, Castanheira, and Drazen 2018). Closest to our findings, Esteban and Ray (2006) show that if agents compete for scarce government permits through lobbying, and if lobbying costs are lower for wealthier agents, then wealth inequality distorts the efficient allocation of permits. In equilibrium, the allocation of permits is determined by agents' willingness to lobby, which —just like willingness to pay in our theory— correlates both with agents' valuations of the permits and with wealth.

The theories on electoral competition with costly voting constitute a second stream of work underpinning vote-buying mechanisms. These theories are a special class of collective choice with costly actions in which the actions available to agents are to vote or to abstain. Turning out to vote is costly, and the collective choice is determined through aggregation of all votes cast. In such a voting game, agents who face idiosyncratic stochastic turnout costs express their strength of preferences through their decision to turn out to vote. Welfare evaluations trade off the optimality of the outcome against the aggregate turnout costs.[29] With a small electorate, turnout can be above (Börgers 2004) or below (Krasa and Polborn 2009) the level that maximizes aggregate utilitarian welfare. In sufficiently large societies with costly turnout, turnout rates —and with them per capita turnout costs— vanish, and asymptotically, welfare hinges only on the collective choice. If the lower bound on the stochastic cost of turnout is zero, and if the realization of the turnout cost is uncorrelated with the preference over the alternatives, simple majority with voluntary voting asymptotically maximizes utilitarian aggregate welfare (Krishna and Morgan 2011 and 2015). Alas, if the distribution of turnout costs does not have a lower bound at zero, the utilitarian optimum is not achieved. If so, Krishna and Morgan (2011) propose to endogenize the cost of participation in the election to bring the lower bound to zero.[30] Vote-buying mechanisms further endogenize the cost of participation: they expand the set of actions available to each agent, and assign a cost to each action, so that agents can express their exact willingness to pay for or against an alternative by choosing one of these costly actions.

[28]Seen as a model of collective choice with costly actions, vote-buying mechanisms are hardly an original class of procedures. Rather, they are an application of one of the most popular allocation mechanisms within the economic literature —a winner-take-all contest— to a binary decision. Indeed, a vote-buying mechanism is just a cost function that maps actions to costs incurred, paired with an outcome function that probabilistically maps the vector of actions to an outcome. The term "vote-buying" is interpreted here broadly to include not just casting a ballot, but also campaigning or donating to a campaign, or taking any other costly action in favor of an alternative.

[29]For a welfare analysis of incentive compatible rules without turnout costs, see Schmitz and Tröger (2012).

[30]If the distribution of costs is uncorrelated with preferences, such endogenous cost of voting restores utilitarian efficiency to majority voting. If the distribution of exogenous costs is correlated with preferences, a designer would need to further endogenize the stochastic costs of voting to make their draws uncorrelated.

Vote-buying mechanisms are also rooted in classic implementation theory with transfers. “VCG” mechanisms (Vickrey 1961; Clarke 1971; Groves 1973; Groves and Loeb 1975; Green, Kohlberg, and Laffont 1976; Green and Laffont 1977a) are dominant strategy incentive compatible, and they choose the utilitarian optimum given all agents’ truthful reports. To achieve this, they extract a payment from each agent equal to the value of the effect of the agent’s report on other agents’ aggregate utility.[31] “AGV” mechanisms (Arrow 1979; D’Aspremont and Gerard-Varet 1979) charge agents the expected externality that their reports cause on others, and are interim (Bayesian) incentive compatible with utilitarian efficiency in quasilinear contexts. To compute the expected externality, it is necessary to know population parameters such as the distribution from which preference profiles are drawn.[32] If the distribution of preferences does not favor any alternative, as society becomes large, the functional form of the expected externality converges toward the square of the valuation expressed by the agent (Goeree and Zhang 2017). Lalley and Weyl (2019) show that in fact, (in a quasilinear environment) a vote-buying mechanism with a quadratic cost function approximates the utilitarian optimum in all equilibria in large societies.[33]

III. Conclusion

We characterize the family of social choice rules that are implementable by a broad class of vote-buying mechanisms in large societies with wealth inequality. Under standard theoretical assumptions, such as zero exogenous costs of participation in the mechanism, common knowledge about the prior distribution from which types are drawn, independently drawn types, and perfectly rational individual optimization, we find that the family of implementable social choice rules coincides with the class of Bergson choice rules defined over willingnesses to pay. Relaxing these assumptions affects our result, as follows.

Consider first that participation in a vote-buying mechanism with cost function c entails an additional exogenous cost $\bar{c} > 0$ of participation, so that an agent pays a cost $\bar{c} + c(a)$ if she takes any action $a \neq 0$, and pays no cost if she takes a zero action. As in standard costly voting models with a common fixed cost, if the society is large, only agents with the greatest willingness to pay (agents with maximal wealth and maximal magnitude of their valuation) participate, so only their preferences are aggregated.[34]

[31] Beyond utilitarianism, Roberts (1979); Carbajal, McLennan, and Tourky (2013); and Marchant and Mishra (2015) characterize the set of choice rules that are implementable in dominant strategies.

[32] Ledyard and Palfrey (2002) show that if the shape of the distribution from which individual preferences are drawn is known and the only information missing is its exact support, then a binary referendum mechanism asymptotically recovers utilitarian efficiency in sufficiently large societies.

[33] On quadratic costs mechanisms, see as well the theoretical contributions by Groves and Ledyard (1977), Hylland and Zeckhauser (1980), Patty and Penn (2017), Kaplov and Kominers (2017), and an experiment by Casella and Sánchez (2020) on a hybrid mechanism with features of storable votes (Casella 2005) and quadratic voting.

[34] As suggested by Krishna and Morgan (2011), if a central planner incentivizes turnout by absorbing the fixed costs of participation, then the results from Theorem 1 are restored. Alternatively, one can recover the same result by making participation compulsory for a randomly drawn subset of the society that is diverging in size, yet tiny as a fraction of the whole population (e.g., a subset of size $\sqrt{n}$). However, these solutions either break budget balance, or may be infeasible (or undesirable) in many applications.

If we weaken the assumption of common knowledge about the environment, we find that Theorem 1 generalizes if agents have an incorrect but common belief about the distribution of payoff-types, or if they have heterogeneous beliefs that are uncorrelated with their payoff types.[35] Whereas, if beliefs are correlated with preferences, then the results are distorted: agents who believe that they are more likely to be pivotal take actions that are greater in magnitude. A similar distortion applies if the environment is common knowledge but types are not independently drawn. In this case, agents' beliefs about the distribution of others' types, and hence about marginal pivotality, depend on the realization of their individual preferences, and thus agents whose preference realization is correlated with greater marginal pivotality will take actions greater in magnitude than agents whose preferences lead them to infer that their marginal pivotality is lower. Theorem 1 is also robust to some departures away from the assumption that the agents are perfectly rational. For instance, if voters mistakenly believe that the marginal pivotality is constant or if they disagree on the rate at which it converges to zero, then our result once again holds as long as the beliefs regarding the marginal pivotality are uncorrelated with payoff types.[36] Therefore, while our main finding does require a substantial level of sophistication, it seems to be, at least in part, pertinent in certain environments with some degree of imperfect rationality.

Finally, it is clear that vote-buying mechanisms are susceptible to collusion.[37] Collusion with commitment and with side trades within a coalition allows members of this coalition to jointly optimize by equating the marginal cost of an action to its total marginal benefit aggregated across all coalition members. Joint optimization increases the magnitude of the coalition's actions relative to individual optimization, and hence it enhances the coalition's influence over the social decision.

Our overall understanding of these mechanisms and the outcomes that they lead to will improve as we complement theoretical studies with additional methodological approaches. Experimental investigations to assess the comparative performance of alternative vote-buying mechanisms seem particularly apt. Clearly, such empirical testing is beyond the scope of the current analysis. Our formal results, though, provide a theoretical foundation to inform future empirical research and institution designers about the interplay between citizens' preferences, the costs of participation in the political process, and the distribution of wealth, as they jointly determine the collective choices made by society.

[35] See Bergemann and Morris (2013) or Carroll (2019) for an overview of implementation theory in environments in which agents lack common knowledge about the environment.

[36] As first noted by Green and Laffont (1977a and 1977b) in the context of VCG mechanisms with sampling of a subset of agents, individual incentives for exact optimization (including incentives to learn one's own type) vanish in arbitrarily large societies under a large class of mechanisms that includes VCG mechanisms, AGV mechanisms (D'Aspremont and Gerard-Varet 1979; Arrow 1979), majority voting with exogenously costly voluntary turnout (Krishna and Morgan 2011 and 2015), and vote-buying mechanisms. If agents optimize imperfectly by taking actions that follow an equilibrium strategy distorted by an additive noise, then our result holds if the magnitude of the noise is smaller than the magnitude of the strategy, but not otherwise.

[37] Most of the closely related mechanisms are also susceptible to collusion, including VCG mechanisms (Bennett and Conn 1977; Green and Laffont 1979), AGV mechanisms (Safronov 2018), and voting mechanisms with exogenous turnout costs.

REFERENCES

Arrow, Kenneth J. “The property rights doctrine and demand revelation under incomplete information.” In Michael Boskin, ed., *Economics and Human Welfare*, New York: Academic Press, pp. 23–39, 1979.

Bennett, Elaine, and David Conn. “The group incentive properties of mechanisms for the provision of public goods.” *Public Choice* 29: 95–102, 1977.

Bergemann, Dirk, and Stephen Morris. “An introduction to robust mechanism design.” *Foundations and Trends in Microeconomics* 8(3): 169–230, 2013.

Blum, Christian, and Christina I. Zuber. “Potentials, problems and perspectives.” *Journal of Political Philosophy* 24(2): 162–182, 2016.

Börgers, Tilmann. “Costly voting.” *American Economic Review* 94(1): 57–66, 2004.

Börgers, Tilmann, and Jiangtao Li. “Strategically simple mechanisms.” *Econometrica* 87(6): 2003–2035, 2019.

Bouton, Laurent, Micael Castanheira, and Allan Drazen. “A theory of small campaign contributions.” NBER w.p. 24413, 2018.

Burk, Abram. “Real income, expenditure proportionality, and Frisch’s ‘New Methods of Measuring Marginal Utility’.” *Review of Economic Studies* 4(1): 33–52, 1936.

Burk, Abram. “A reformulation of certain aspects of welfare economics.” *Quarterly Journal of Economics* 52(2): 310–334, 1938.

Campante, Filipe R., and Francisco H.G. Ferreira. “Inefficient lobbying, populism and oligarchy.” *Journal of Public Economics* 91(5–6): 993–1021, 2007.

Carbajal, Juan Carlos, Andrew McLennan, Rabee Tourky. “Truthful implementation and preference aggregation in restricted domains.” *Journal of Economic Theory* 148: 1074–1101, 2013.

Carroll, Gabriel. “Robustness in mechanism design and contracting.” *Annual Review of Economics* 11: 139–166, 2019.

Casella, Alessandra. “Storable votes.” *Games and Economic Behavior* 51(2): 391–419, 2005.

Casella, Alessandra, Aniol Llorente-Saguer, and Thomas R. Palfrey. “Competitive equilibrium in markets for votes.” *Journal of Political Economy* 120(4): 593–658, 2012.

Casella, Alessandra, and Luis Sánchez. “Democracy and intensity of preferences. A Test of storable votes and quadratic voting on four California propositions.” *Journal of Politics*, forthcoming.

Clarke, Edward H.. “Multipart pricing of public Goods.” *Public Choice* 11(1): 17–33, 1971.

Dahl, Robert Alan. *Democracy and its Critics.* Yale University Press, 1989.

d’Aspremont, Claude and Louis-André Gérard-Varet. “Incentives and incomplete information.” *Journal of Public Economics* 11(1): 25–45, 1979.

Dekel, Eddie, Matthew O. Jackson, and Asher Wolinsky. “Vote buying: General elections.” *Journal of Political Economy* 116(2): 351–380, 2008.

Dixit, Avinash, Gene M. Grossman, and Elhanan Helpman. "Common agency and coordination: General theory and application to government policy making." *Journal of Political Economy* 105(4): 752–769, 1997.

Esteban, Joan, and Debraj Ray. "Conflict and distribution." *Journal of Economic Theory* 87(2): 379–415, 1999.

Esteban, Joan, and Debraj Ray. "Inequality, lobbying and resource allocation." *American Economic Review* 96(1): 257–279, 2006.

Esteban, Joan, and Debraj Ray. "A model of ethnic conflict." *Journal of the European Economic Association* 9(3): 496–521, 2011.

Goeree, Jacob K., and Jingjing Zhang. "One man, one vote bid." *Games and Economic Behavior* 101: 151–171, 2017.

Green, Jerry, Elon Kohlberg, and Jean-Jacques Laffont. "Partial equilibrium approach to the free-rider problem." *Journal of Public Economics* 6(4): 375–394, 1976.

Green, Jerry, and Jean-Jacques Laffont. "Characterization of satisfactory mechanisms for the revelation of preferences for public goods." *Econometrica* 45(2): 427–438, 1977.

Green, Jerry, and Jean-Jacques Laffont. "Imperfect personal information and the demand revealing process: A sampling approach." *Public Choice* 29: 79–94, 1977.

Green, Jerry, and Jean-Jacques Laffont. "On coalition incentive compatibility." *Review of Economic Studies* 46(2): 243–254, 1979.

Groves, Theodore. "Incentives in teams" *Econometrica* 41(4): 617–631, 1973.

Groves, Theodore, and Martin Loeb. "Incentives and public inputs." *Journal of Public Economics* 4(3): 211–226, 1975.

Groves, Theodore, and John Ledyard. "Optimal allocation of public goods: A solution to the free rider problem." *Econometrica* 45(4): 783–810, 1977.

Hardt, Steve, and Lia C.R. Lopes. "Google Votes: a liquid democracy experiment on a corporate social network." Working paper, 2015.

Hylland, Aanund, and Richard Zeckhauser. "A mechanism for selecting public goods when preferences must be elicited." Kennedy School of Government Discussion Paper D, 70, 1980.

Jackson, Matthew O. "Implementation in undominated strategies: A look to bounded mechanisms." *Review of Economic Studies* 59(4): 757–775, 1992.

Kang, Karam. "Policy influence and private returns from lobbying in the energy sector." *Review of Economic Studies* 83(1): 269–305, 2016.

Kaplov, Louis, and Scott D. Kominers. "Who will vote quadratically? Voter turnout and votes cast under quadratic voting." *Public Choice* 172(1–2): 125–149, 2017.

Karabarbounis, Loukas. "One dollar, one vote." *Economic Journal* 121(553): 621–651, 2011.

Kazumura, Tomoya, Debasis Mishra, and Shigehiro Serizawa. "Mechanism design without quasilinearity." *Theoretical Economics* 15(2): 511–544, 2020.

Krasa, Stefan, and Mattias K. Polborn. "Is mandatory voting better than voluntary voting?" *Games and Economic Behavior* 66(1): 275–291, 2009.

Krishna, Vijay, and John Morgan. "Overcoming ideological bias in elections." *Journal of Political Economy* 119(2): 183–211, 2011.

Krishna Vijay, and John Morgan. "Majority rule and utilitarian welfare." *American Economic Journal: Microeconomics* 7(4): 339–375, 2015.

Lalley, Steven P., and E. Glen Weyl. "Quadratic voting." *American Economic Association Papers and Proceedings.* 108: 33–37, 2018.

Lalley, Steven P., and E. Glen Weyl. "Nash equilibria for quadratic voting." arXiv:1409.0264v6, 2019.

Ledyard, John O., and Thomas R. Palfrey. "The approximation of efficient public good mechanisms by simple voting schemes." *Journal of Public Economics* 83(2): 153–171, 2002.

Locke, John. *Two Treatises of Government.* Cambridge U. Press, 1988 [1689].

Marchant, Thierry, and Debasis Mishra. "Mechanism design with two alternatives in quasi-linear environments." *Social Choice and Welfare* 44(2): 433–455, 2015.

Matsushima, Hitoshi. "Detail-free mechanism design in twice iterative dominance: Large economies." *Journal of Economic Theory* 141(1): 134–151, 2008.

May, Kenneth O. "A set of independent necessary and sufficient conditions for simple majority decision." *Econometrica* 20(4): 680–684, 1952.

Milgrom, Paul R., and Robert J. Weber. "Distributional strategies for games with incomplete information." *Mathematics of Operations Research* 10(4): 619–632, 1985.

Moulin, Hervé. *Axioms of Cooperative Decision Making*. Cambridge U. Press, 1988.

Patty, John W. and E. Maggie Penn. "Uncertainty, polarization, and proposal incentives under quadratic voting." *Public Choice* 172(1–2): 109–124, 2017.

Reny, Philip J. "On the existence of pure and mixed strategy Nash equilibria in discontinuous games." *Econometrica* 67(5): 1029–1056, 1999.

Roberts, Kevin W. "The characterization of implementable choice rules." In: *Aggregation and revelation of preferences* (Editor J.J. Laffont) 12(2): 321–348, 1979.

Roberts, Kevin W. "Interpersonal comparability and social choice theory." *Review of Economic Studies* 47(2): 421–439, 1980.

Safronov, Mikhail. "Coalition-proof full efficient implementation." *Journal of Economic Theory* 177: 659–677, 2018.

Schmitz, Patrick W., and Thomas Tröger. "The (sub-) optimality of the majority rule." *Games and Economic Behavior* 74(2): 651–665, 2012.

Semega, Jessica, Melissa Kollar, Emily A. Shrider, and John F. Creamer. *U.S. Census Bureau, Current Population Reports, P60–270, Income and Poverty in the United States: 2019*. U.S. Government Publishing Office, 2020.

Stromberg, David. "How the Electoral College influences campaigns and policy: the probability of being Florida." *American Economic Review* 98(3): 769–807, 2008.

Vickrey, William. “Counterspeculation, auctions and competitive sealed tenders.” *Journal of Finance*, 16(1): 8–37, 1961.

Walzer, Michael. *Spheres of Justice: A Defense of Pluralism and Equality*. New York : Basic Books, 1983.