Quadratic Voting and the Public Good: Introduction

By Eric A. Posner[1] and E. Glen Weyl[2]

Abstract: This introduction to the *Public Choice* special issue on "Quadratic Voting and the Public Good" provides an opinionated narrative summary of the contents and surveys the broader literature related to Quadratic Voting (QV). QV is a voting rule, proposed by one of us (Weyl, 2012; Lalley and Weyl, 2016) building off earlier work by Groves and Ledyard (1977) and Hylland and Zeckhauser (1980), where individuals buy as many votes as they wish by paying the square of the votes they buy using some currency. An appreciation of the history of research in the field suggests that QV is uniquely practically relevant compared to the other approximately Pareto-efficient mechanisms economists have proposed for collective decisions on public goods.

However, it faces a number of sociological and ethical concerns regarding how a political system organized around QV would achieve the efficiency aims stated in abstract theory and whether the pure aggregate income-maximizing definition of efficiency QV optimizes in its simplest form is desirable. The papers in this volume flesh out and formalize these concerns, but also provide important responses in two ways: by suggesting domains where they are unlikely to be applicable (primarily related to survey research of various kinds) and versions of QV (using an artificial currency) that maintain many of QV's benefits while diffusing the most important critiques. Together this work suggests both a practical path for applying QV in the near-term and a series of research questions that would have to be addressed to broaden its application.

Keywords: Quadratic Voting, collective decisions, survey research, welfare criteria, market design
JEL Classifications: B21, D47, D61, D63, D71

Determining a socially desirable means of making collective decisions is perhaps the oldest and largest open problem in the social sciences. While it was the central preoccupation in the Greek pre-history of the field, progress has been limited. Economists and much of the rest of society largely have reached consensus on the desirability of markets for allocating private goods. Yet deep dissatisfaction and pessimism are prevalent about nearly all existing mechanisms for collective decision-making and the allocation of public goods.

Perhaps the basic problem underlying existing collective decision procedures is that they rely on the principle of rationing (viz., every individual is rationed a single vote on each political contest or issue) rather than on the market principle of trade (viz., individuals can exchange influence on issues less important to them for influence on those more important to them). The conference published in this special issue explores a method of bringing the logic of the market to collective decisions proposed recently by one of us (Weyl 2012), Quadratic Voting (QV), from a variety of interdisciplinary perspectives to assess its promise for addressing this classic challenge.

The inspiration for this conference was a special issue published in *Public Choice* forty years ago on another proposed mechanism for collective decisionmaking called variously the "pivotal

[1] University of Chicago Law School, 1111 E. 60th Street, Chicago IL 60637; Email: eric_posner@law.uchicago.edu, http://ericposner.com/.
[2] Microsoft Research New York City, 641 Avenue of the Americas, New York, NY 10011 and Department of Economics, Yale University: 28 Hillhouse Avenue, New Haven, CT 06511; E-mail: glenweyl@microsoft.com, http://www.glenweyl.com.

mechanism", the "demand revealing process" or simply the "Vickrey-Clarke-Groves (VCG) mechanism" after its three discoverers, William Vickrey (1961), Edward Clarke (1971) and Theodore Groves (1973). Nicolaus Tideman and Gordon Tullock (1976) argued that this mechanism constituted "a new and superior process for making social choices". Despite this initial optimism, it is safe to say that the VCG mechanism essentially has had no impact in the 40 years since that time on the making of collective decisions in practice. One reason for its failure, we believe, is the practical weaknesses of VCG, particularly related to collusion and the necessity of using money, weaknesses that were recognized by the special issue's authors and later highlighted in experiments on the mechanism. The primary goal of this conference is to subject QV to analogous scrutiny and determine whether it suffers similar deficiencies that may limit its practical value or whether it may in fact make a useful contribution to promoting the public good.

These challenges, like those with VCG, fall into two broad categories: positive concerns about whether the mechanism operates as intended in face of potential collusion or manipulations and normative concerns about whether the efficient outcomes QV aims to implement are desirable and just. Some of the papers flesh out and define the nature of these challenges and others describe ways and domains in which QV can be implemented that avoid these concerns. Together these papers paint a rich picture of both the concerns about QV, a wide range of applications wherein these concerns are not central, and a range of modifications to the implementation of QV that could avoid these concerns more broadly. From this narrative emerges not just a picture of the promise of QV in the near term, but an agenda for building off QV and making it the basis for improving collective decision in the most important and problematic national and international arenas.

1 Background

The first paper in this issue tracks the development of economists' study of collective decision making from the Second World War to the late 1980s. Beatrice Cherrier and Jean-Baptiste Fleury highlight the frustrations and disappointments economists repeatedly confronted in the post-war period as they grappled with the problem of collective decisionmaking. Prior to and during the war, economists largely either confined their attention to positive analysis or took as given a set of objective utilitarian goals in the spirit of Bentham and Sidgwick (Petit 2012). However, as Tuck (2012) highlights, the latter approach was abandoned by economists who tried to found utilitarian aggregations in individual *preferences* rather than a more objective utilitarian notion of welfare. Such efforts forced Samuelson and other "new welfare economists" to confront the lack of a clear basis for the welfare functions they worked with in individual preferences. Reviving and extending the insights of the Marquis de Condorcet from the late 18th century, Arrow (1951) provided a sense in which such a foundation is impossible so long as the only available information is individuals' preference orderings over outcomes.

Arrow's work was the most famous of a series of pessimistic conclusions economists reached mid-century about the possibility of meaningful collective decision-making. Bowen (1943) highlighted the tyranny of the majority in democratic voting: a majority may favor, for example, banning gay marriage, but a ban on gay marriage would be Pareto-inefficient if the supporters of gay marriage would be willing to compensate its opponents so that no one is made worse off. Samuelson (1954) demonstrated the free rider problem and the failure of linear pricing to achieve reasonable quantities of public goods. A variety of other results, both empirical and theoretical, suggested that the failures of majority rule might be extreme.

This pessimism, however, was interrupted briefly by the discovery of efficient mechanisms for collective decision-making in the 1970s by Clarke and Groves. Their work was eventually tied

back to that of Vickrey ten years earlier and spurred a resurgence of interest in whether satisfactory collective decisions could be reached using this VCG mechanism or an alternative procedure proposed by Groves and John Ledyard (1977a). However, pessimism returned, partly as a result of the special issue mentioned above, where the practical relevance of the mechanisms largely was dismissed by experts, and a series of negative experimental and theoretical results related primarily to the VCG, but also to some extent to the Groves and Ledyard mechanism, discouraged economists from pursuing the applications of these mechanisms. At the same time, the acceleration of deregulation and privatization during the 1980s and 1990s, as well as the work on the design of private markets without money by Alvin Roth, increasingly drew the attention of applied mechanism designers to private goods problems.

What exactly were these mechanisms and what were the challenges that undermined their plausibility in applications? Nicolaus Tideman and Florenz Plassmann cover this intellectual prehistory of QV in their paper. We leave the formalism to them and here summarize briefly and verbally the operation of these mechanisms and the concerns they raised.

While the mechanism may be applied much more broadly, the basic concept behind VCG can be seen in the example of a binary collective choice. We use the example of a referendum on the legality of gay marriage. Every individual i announces how many dollars b_i she would be willing to pay to see gay marriage legalized (positive) or banned (negative). The alternative with the greatest total willingness to pay is chosen (gay marriage is legalized if $\sum_i b_i \geq 0$), and any individual who is *pivotal* in the sense that the outcome would have been different had she not participated is forced to pay the minimum amount she would have had to announce in order to win. For example, imagine that the total willingness-to-pay in favor of gay marriage absent some individual i is \$500,000. If that individual announces opposition in the amount of \$600,000, gay marriage is banned and she must pay \$500,000. If instead she announces \$400,000 of opposition, gay marriage is legalized and she pays nothing.

Numerous concerns have been raised regarding the practical value of this mechanism, many related to its complexity, but we highlight two that we believe had the largest impact on the evolution of the debate, stood the test of time most clearly, were lucidly highlighted by Groves and Ledyard in their contribution to the special issue on VCG (1977b), and have close parallels in the discussion of QV. First, the mechanism is sensitive to even small collusive schemes or manipulation. Second, the mechanism relies heavily on individuals being risk-neutral and having unlimited stores of cash to draw upon.

First, consider collusion or fraud. Suppose that two individuals who want to see gay marriage defeated both report a willingness to pay \$1 billion to see it defeated. Supposing this amount is larger than any plausible margin of victory in the election among all other individuals, neither of these individuals *individually* will be pivotal. Thus, they will not have to pay anything and they will get the outcome they both desire. In fact, this situation is an equilibrium as there is no incentive for any other individual, or the conspirators, to change their behavior. Thus, any two individuals can in VCG achieve any outcome they desire and make no payments. VCG is extremely sensitive to collusion and in experiments such schemes have proven quite common (Attiyeh et al. 2000). A similar problem arises if a single individual can misrepresent herself successfully as two people.

Second, VCG leads individuals very often to make no payments at all … but very infrequently make extremely large payments, up to the full amount of their willingness to pay to change the outcome. This may create major challenges for individuals facing budget constraints, who are risk averse or, most severely, if income effects come into play. Groves and Ledyard (1977b) show that in this last case VCG tends to be extremely unstable and Budish (2011) discusses the

problems such income effects may create with VCG even in simpler settings of allocating private goods.

While income effects arguably are small in many real-world settings with reasonably well-off individuals (Willig 1976), as we will discuss further in Section 4 below, many people are deeply suspicious of real money being used to make collective decisions because of equity and legitimacy concerns, as well as taboos against vote buying. This means that, at least initially, practical implementations of efficient collective decision mechanisms are likely to be applied primarily to allow trade among a relatively narrow set of collective choices rather than to allow the wholesale purchasing of collective decisions using the currency universally applied to private goods. When the scope of trade is thus limited, however, income effects become extremely strong because each decision consumes a large fraction of the total currency and thus VCG becomes essentially impossible to implement.

For both of these reasons, Groves and Ledyard despaired of the potential of using VCG for collective decisions and instead proposed a mechanism that applied only to continuous social choices, such as over the provision level of a public good valued in a differentiable and concave manner by the participants. We describe an altered version of the Groves-Ledyard (GL) mechanism that we believe highlights its essential features better and focus on the case of a single public good. We also present the mechanism in a quasi-linear valuation setting where all value can be denominated in units of dollars for simplicity of exposition, though a key feature of the GL mechanism is that it extends to allow income effects and limited budgets as we discus below. See the Tideman and Plassman contribution for a more faithful representation of the original GL mechanism.

Every individual i announces the amount of the public good $q_i^\star$ she expects to be produced and the increment Δ_i she would like to make to this public good. If all individuals agree that this expectation is $q^\star$, then this quantity plus the sum of the reported increments $q^\star + \sum_i \Delta_i$ is implemented and every individual pays some constant (which we will normalized to one) multiplied by the square of her reported increment Δ_i^2 and receives back the average amount paid in by others. If there is a single dissident about the expected amount, this individual is punished harshly. If multiple dissidents emerge, increments are solicited again.

GL assume that there is complete information about the distribution of preferences and thus that individuals can be relied upon to report the correct and same amount of the public good to be created; we will return to this point shortly. But for the moment suppose this condition is satisfied and consider the incentives created by the rule in that case.

The part of individual $i's$ utility under her control is $2U_i\left(q^\star + \sum_j \Delta_j\right) - \Delta_i^2$, where $2U_i$ is her utility over the public good. The benefit each individual derives by increasing her reported increment a bit higher is the marginal value of the public good $2\frac{\partial U_i}{\partial q}$. The marginal cost is twice the size of this increment $2\Delta_i$. A maximizing individual will equate the marginal benefits and costs and thus set the size of her increment equal to her marginal utility: $\Delta_i = \frac{\partial U_i}{\partial q}$. If the expected amount of the public good is the socially optimal amount, then $\sum_i \Delta_i = 0$ so that the planner will implement exactly $q^\star$. But this result implies that $\sum_i \frac{\partial U_i}{\partial q} = 0$, which is the condition for socially optimal provision of public goods from Samuelson (1954). Thus, at the point where the increments sum to zero, we have achieved the socially optimal allocation.

Clearly the assumption that $q^\star$ is common knowledge is unrealistic in most settings. A more realistic procedure, and most experimental implementations of the GL procedure, involve gradual iteration towards discovery of $q^\star$ through an auction-like process. This eliminates the need for

participants to announce $q^\star$ while retaining their announcements of Δ_i and the rule for determining when equilibrium has been reached. Hylland and Zeckhauser (1980) propose a similar idea, but using an artificial currency to trade off multiple public goods, rather than using the same money used to purchase private goods. Each individual can demand at each stage a movement in the public good vector in any direction, but the sum of the square of the changes requested must be a constant. In the simplest version, each individual must choose a vector of unit length for the movement and equilibrium occurs when the sum of all vectors cancel. Hylland and Zeckhauser also discuss more explicitly the process of iterating towards the equilibrium.

The GL and Hylland-Zeckhauser (HZ) mechanisms attracted some theoretical and even experimental attention, but to our knowledge were not used in applications. An important reason likely is the complexity involved. Unlike VCG, these mechanisms do not offer a direct solution to the binary collective decision problems to which standard voting commonly is applied. In essence, QV adapts the core insight of GL to this binary context, wherein no iteration is necessary as there is a single decision to be made.

In QV individuals buy votes v_i (positive or negative) on the issue at hand. People use their votes to influence the probability of this issue being decided one way or the other rather than to influence the level of a continuos public good; individuals in QV do not therefore announce any level of the public good. The decision is made in the direction of aggregate votes. Otherwise the rules are as in the version of GL we described above. This change may seem small, but significantly expands the range of applications. Additionally, the version of GL we describe above incorporates elements innovated by HZ and is more suggestive of QV, which shares elements of both of these mechanisms.

A final difference between QV, on the one hand, and GL/HZ on the other is that no formal results have ever been proven about the efficiency in the presence of optimizing strategic agent behavior in GL or HZ when information is incomplete and iteration towards the optimum thus is needed. It can easily be shown (see, e.g., the Benjamin et al. contribution in this volume) that in finite populations such strategizing can lead GL/HZ to choose an outcome that is not Pareto optimal. In the analogous context of general equilibrium theory, Roberts and Postelwaite (1976) show that incentives exist to manipulate the path towards a Walrasian equilibrium, though these vanish in large populations. This suggests that GL/HZ might be efficient under strategic behavior in large populations. However, this result has not been shown formally, as the existing analysis considers only the complete information setting or the final iteration at equilibrium. The strategic incentives under GL/HZ in realistic contexts are thus conjectural.

By contrast, Lalley and Weyl (2016) (LW) prove that QV achieves welfare arbitrarily close to the first-best in Bayes-Nash equilibrium of a standard independent private values environment as the population grows large. The proof involves quite detailed statistical analysis even in this simpler binary collective decision context, which may account for the lack of a rigorous strategic analysis of GL/HZ in previous literature.

While these clear, if somewhat superficial, differences exist between GL, HZ and QV, all share a similar logic and set of contrasts to VCG. Why might GL, HZ and especially QV be more practically useful than VCG? One reason, emphasized in the Tideman and Plassmann contribution, is simplicity: in these systems every individual knows a simple rule that turns her report into a payment regardless of others' behavior and even in GL and HZ individuals report only a series of increments rather than their full value function for the public good as is required in the corresponding version of VCG; see LW for a further discussion of simplicity.

However, the focus of the Weyl contribution in this volume is a close cousin of that simplicity: robustness. Let us return to the two concerns we highlighted, echoing Groves and Ledyard, about VCG above.

First, while collusion certainly can be effective against these mechanisms, it is not as devastating as it is under VCG where any two individuals may obtain any outcome they like at no cost to themselves. Under QV, by contrast, individuals may gain by partly overcoming the convexity of the quadratic function through collusion. For example, instead of buying 10 votes for $100 one could recruit 99 friends willing to pay $1 each and get 100 votes for the same expenditure. However, this is a quite large-scale fraud and increases the votes one manages to have by less than the number of fraudulent colluders. In this sense, collusion under QV may not pose much greater problems than the analogous phenomenon, vote buying, under standard one person, one vote (1p1v) rules. Second, and we return to this in much greater detail below, these mechanisms can easily be implemented with tight budget constraints and large income effects because each participant knows precisely how much she is spending when choosing her votes and these expenditures typically will be small in a reasonably large population as no individual can have much influence. This allows budget-constrained individuals readily to plan their purchases of influence and not wait for the unpredictable realization of other individuals' choices.

Weyl's contribution studies QV as in LW, though his methods use approximations and are less fully rigorous than those of LW. He focuses on the first issue, as well as on two other robustness concerns: those surrounding the information available to voters and the way that voters make their choices (their motivation and rationality). We start with the collusion, as that issue is most related to distinguishing QV from VCG.

LW show that two regimes arise in QV equilibrium. When the election is close, all individuals vote roughly proportionally to their values, with this approximation getting arbitrarily good as the population grows. When the election is not close, almost all individuals vote in proportion to their values, but with a tiny probability "extremist" individuals buy so many votes that they unilaterally cause a tie to be quite likely. It is this risk of an extremist "stealing" the election that keeps others voting. Weyl's contribution shows that the efficacy and barriers to collusion are quite different in these two cases.

When the election is close, the principal bar to collusion under QV is the fact that, to be efficacious in moving the election requires either a very large number of participants or for the participants to buy votes that are very large relative to what unilaterally is optimal for them to buy. This encourages participants to defect and makes it unlikely that a conspiracy small enough to avoid detection could have much impact.

When the election is not close, however, a conspiracy aimed at extremist behavior is more self-reinforcing: because a tie is very unlikely unless an extremist conspiracy forms, the conspirators who know it has formed have a stronger incentive to buy votes because they are more likely to be pivotal than other members of the population. However, if others in the population realize that even if a chance of such a conspiracy exists, they will believe that a tie is much more likely. This will lead them to buy more votes and make the conspiracy partly self-defeating. Because of this, only collusive groups that are very large relative to the population, and thus likely to be detected by authorities, can be effective.

Weyl also studies the performance of QV when values are not drawn independently and identically from a fixed value distribution. In that case, the outcome of the vote is uncertain even in large populations. He shows that this uncertainty may result in some inefficiency because of an "underdog effect" such that supporters of the likely loser of the election believe that a tie is more likely and thus vote more heavily than supporters of the favorite. However, this can occur only if the favorite remains favorite and thus cannot cause severe inefficiency. Weyl finds that at most this

leads to an efficiency loss of a few percentage points, while 1p1v in many cases has inefficiency as high as 100%.

Finally, he considers the possibility that voters are not the perfectly rational and instrumental automata that the basic theory assumes by allowing voters to a) overestimate the chance they are pivotal, b) wish to use their votes to express themselves or c) vote partly to influence the vote total and the signal it sends. So long as the strength of these motives relative to values are not correlated with the values themselves, none interferes with efficiency when the election is not close. In fact, such considerations often accelerate the rate of convergence to efficiency by increasing voting and thus deterring extremists.

When the election is close, however, these motives can introduce noise that can cause inefficiency. Whether 1p1v or QV is superior then depends on the extent to which this noise is greater than the heterogeneity of intensity of preference QV allows to be communicated. This is closely analogous to a market economy, where rationing is superior to trade if and only if individuals are sufficiently irrational that the rationed allocation is superior to the trade facilitated by the market.

While these results suggest the robustness of QV to variations in some of the most unnatural assumptions of the baseline analysis, violation of other assumptions may create problems for QV, as the next pair of papers highlights.

2 Positive challenges

The first set of concerns are raised by the contribution from John Patty and Maggie Penn. They focus on purely rational manipulations of QV through the choice of the initiatives on the collective decision-making agenda. Their concern is that issues that are highly divisive will increase spending on votes and thereby generate revenue that is refunded to the participants. Especially for voters who simply do not care about the issue, their only interest in what issues are on the agenda is to generate maximum revenues. Such agenda-setters may even wish to make the overall quality of a proposal on the agenda worse simply to make it more controversial.

An important contribution of the Patty and Penn piece is to provide some of the first closed-form solutions for QV in finite populations; the first such solution to our knowledge appeared for the case of normally distributed mean zero values in Goeree and Zhang (Forthcoming). This allows them to study these concerns analytically. However, it also somewhat limits the relevance of their analysis to a case quite different from that on which we have focused thus far. QV was intended primarily by Weyl to be a mechanism for fairly large populations wherein bargaining is unlikely to lead to efficient solutions (Mailath and Postelwaite 1990). In large populations, the refund received by any individual will be small relative to the work necessary to influence the agenda and thus it is unlikely that the concerns Patty and Penn raise will be among the leading issues even in determining how agendas are set.

In contrast to Patty and Penn's focus on exact solutions of a perfectly rational model in a small society, the contribution by Louis Kaplow and Scott Duke Kominers is much less formal in its analysis and focuses on problems that arise particularly in very large societies, comprising voters with a broad range of motivations and strategies. They follow up on the motivation behind Weyl's departure from purely rational and instrumental voter behavior, noting that such behavior cannot account for voting patterns at present and thus may not be a good model of voting if QV were to be adopted, nor a good benchmark to compare QV against under 1p1v. Because in very large populations the chance of being pivotal necessarily is quite small, purely rational and instrumental behavior is very unlikely to explain current voting patterns, suggesting that under QV much voting may be driven by factors other than purely instrumental rationality.

However, rather than taking the relatively constrained perspective on deviations from rational agent behavior on which Weyl focuses, wherein all voters are assumed to behave according to one particular model, possibly with some tightly specified heterogeneity independent of values, Kaplow and Kominers explore what may occur when different voters differ in fundamental ways in their motivations and social structures. For example, they consider the possibility that many voters will follow rules of thumb that lead them to buy, say, exactly one vote. This may depress the vote among other, more rational voters by reducing their chances of being pivotal. This may entirely drive out rational voting as well as depressing the weight put on rational voters' preferences. If these voters have substantive preferences different from the rest of the population, this may lead to bad outcomes under QV in a way similar to differential turnout rates in standard voting. In fact, all of these effects are more severe if voting is not compulsory and rational voters thus may choose not to turn out. On the other hand, if all voters behave according to heuristics, QV may behave basically like 1p1v as everyone will choose to buy one vote. This will not make QV worse than 1p1v, but may undermine many of its benefits.

On the other hand, if these possible behaviors are combined with clever strategies on the part of social movements, the effects may be more positively harmful. While social movement have limited ability to manipulate outcomes without genuinely changing minds under 1p1v, primarily restricted to encouraging turnout when voting is voluntary, under QV they might try to impact the number of votes purchased. To some extent, their ability to do so would be constrained by the preferences of voters, in which case the votes still would be roughly proportional to values, but might be skewed by the organizational ability of social movements in a way that might bias outcomes. However, in some cases social movements' abilities to instill misconceptions, superstitions and the like, may fundamentally undermine any relationship between QV's outcome and efficiency in a way that is harder to achieve under 1p1v. In these cases, QV may significantly underperform 1p1v.

Perhaps the most important takeaway from the Kaplow and Kominers analysis is that the real-world performance of QV likely will depend very heavily not merely on economic or even psychological factors, but on the sociology and political organizations that form around it. These organizations are critical to the actual effects of present political arrangements based on 1p1v. It is quite difficult to predict, certainly without modeling that goes far beyond anything scholars have undertaken so far, the nature of the organizations and political structures that would form in a society whose collective decisions were based on the QV concept. As a result, it seems hard to imagine forming accurate predictions about these effects without large scale experimentation. However, such experimentation could be quite costly, especially if these factors end up having some of the extreme negative repercussions that Kaplow and Kominers highlight.

That said, it is important to recognize that such speculative sociological, political, and psychological concerns could be—and have been—lodged against all voting reform proposals. Authoritarian governments have, for centuries, made similar arguments against democratic restructuring, which they have argued, with some justice, may lead to even weaker protection of minorities than exists in nondemocratic countries. No voting rule can be implemented mechanically; legal safeguards will need to be introduced as society learns about its weaknesses and observes how people take advantage of them.

As a result, it is natural to try to identify both ways to rectify some of these potential harms as well as domains in which QV can be tested where the potential risks from these factors are not relatively small. This would allow a gradual process of incremental experimentation, learning and improvement to permit the exploitation and eventual maximization of QV's benefits while containing both the scale and scope of the potential harms it could create. It is to this task that the next pair of contributions turn.

3 Solutions and domains robust to positive challenges

The contribution by Sunoo Park and Ron Rivest in the main addresses the concerns about centralized "attacks", such as collusion or fraud, with which Weyl's contribution is concerned, rather than the "softer" manipulations that Patty, Penn, Kaplow and Kominers study. However, their design seems important for building public confidence in QV in almost any context and thus would likely be an important baseline for beginning to confront those concerns.

Experts on voting security, Park and Rivest develop new protocols that would protect the integrity (immunity to fraud) and secrecy (an important component of the robustness against collusion) of a QV election either in a physical voting location or online, using either very primitive and thus transparent physical technologies or using cutting-edge cryptographic techniques. All of these approaches have a common set of goals, which they are able to achieve to greater or lesser degrees depending on the setting. These are:

a. Verifiability: any member of the public with enough time and patience can verify that the election was carried out honestly.
b. Secrecy: no one can know how much anyone else voted.
c. Robustness against false accusations: no one falsely can claim fraud when it did not occur.
d. Usability: the system is fairly painless for participants.

To achieve these goals, Park and Rivest propose a series of clever ideas in both the physical and digital realms that allow voters to challenge potentially compromised systems without giving them the opportunity to either manipulate the system or prove to outsiders that they have voted in a particular way. They also propose rules that keep track of precisely the necessary information to determine the total vote and money raised, while disguising everything else about the identity of the contributors to the outcome so as to avoid collusion. Furthermore, they manage to extend many, though not all, of these benefits to surveys that are taken online, from remote locations. These ingenious designs seem destined to be of practical relevance to a variety of applications of QV, though a very important direction for future research will be to extend their insights to the artificial vote currency versions of QV that we describe below and appear to be more practical in the near term in many settings.

Park and Rivest also offer an interesting speculative investigation into the types of refund schemes that could be used under QV without impairing its theoretical properties. In particular, they consider the class of refund rules, both randomized and deterministic, that would not only avoid giving voters distorted unilateral incentives but also avoid collusive schemes that could be problematic. While we suspect that in practice most designs will adopt only the simplest schemes, such as proportional rebates or simply using the funds raised for defraying the costs of public projects, this articulation of the boundaries of the efficiently feasible offers a good benchmark for further analysis.

Jonathan Masur's contribution provides a clearer answer to the challenges raised by Patty, Penn, Kaplow and Kominers. He describes a domain in which QV could be applied directly to public policy choices in their current form without raising, or at least without raising very strongly, the sorts of concerns they highlight. In particular, he discusses the use of QV in a setting where economics and economic-based mechanisms already are the standard: benefit-cost analyses (BCA) undertaken by executive administrative agencies.

BCAs, which are common for health, environmental and safety regulations and have played critical roles in determining regulatory policy since the early 1980s, draw on three categories of data.

The most solid typically are scientific data that tie down the physical effects of regulations. However, these data are not sufficient because the physical effects must be translated into dollar values to offer a full account of benefits and costs. This is usually done in one of two manners: either using values already existing in marketplaces (e.g., if a certain amount of farmland is encumbered by a regulation, its encumbrance is valued at the reduced market value of the output) or using surveys to judge the value of public goods not already priced in the market (e.g., how much is it worth to save a species of whale from extinction). While the techniques for the former sort of valuation have serious limitations as they reveal only marginal and not infra-marginal market valuations, techniques for the latter category of valuation are far more controversial and it is fair to say that no existing methodology has widespread support. Masur thus focuses on the second issue.

The most common technique used in this setting is the so-called contingent valuation survey. In a contingent valuation survey, respondents are asked to place a monetary value on the cost of extinction. The average value reported is then used to determine the aggregate cost of that outcome. The trouble is that the surveys provide no incentives for people sincerely to reveal their valuations. If you believe that you are more (less) environmentally concerned than an average respondent, you want massively to overstate (understate) your value to move the average toward your preferred point. The more strategic participants are, the wilder valuation reports should get. Moreover, as Hausman (2012) highlights, even when agents apparently do not act strategically, they give responses that are extremely strange and misleading, subject to a range of framing effects and confusion. In short, this method largely is discredited among economists.

Masur proposes an alternative approach based on QV. A benefit-cost analyst would determine a threshold for cost that would have to be passed for the relevant policy to be implemented. Respondents would then be asked to participate in a quadratic vote on whether the value was above or below this threshold, effectively determining the policy outcome. This would put either a lower or an upper bound on the value of extinction that could be applied in future policies, determining some of their outcomes. However, when the bound does not give a clear answer, further surveys would have to be held, or, if many such questions were expected to arise, a series of surveys could be conducted upfront to narrow the valuation range. When multiple public goods must be valued to make the ultimate determination, a series of surveys could be held.

The advantage of this approach is that, unlike contingent valuation, it gives truthful incentives to participants as they are implicitly participating in a quadratic vote on the relevant policy, assuming that they trust the other analyses of the administrative agency. In fact, one can view the choice of the value assigned to the public good at hand as essentially a continuously valued public good, with participants voting on whether to move upward or downward the bounds on its value. This approach offers an interesting alternative path towards convergence to efficiency in a GL setting with a single public good: rather than individuals purchasing the sizes of increments they could simply purchase votes on whether any given level of the public goods is an upper or lower bound for the amount that should be provided. The procedure offers a version of GL that is even more closely related to QV.

Furthermore, such a setting is ideal for diffusing the concerns raised by Patty, Penn, Kaplow and Kominers. First, it is of large enough scale, and the agenda is set clearly by someone with no direct interest in the funds raised, that little concern arises about referendums being designed to extract money from the respondents. Second, because the number of respondents is relatively small, respondents are fairly likely to be pivotal and unlikely to associate the survey response with standard preconceptions on which they have above voting. As a result, the heuristic behavior Kaplow and Kominers are concerned with is less likely to be a problem, though the sort of limited rationality and expressive motivation Weyl considers could still arise. Third, and what is perhaps most important, it will be hard for interest groups to organize around the outcome of such a survey as they do not

know who will participate or when. Thus, BCA seems like perhaps the most promising large-scale public policy application of QV in the near term, until these other concerns are resolved.

4 Normative challenges

The concerns about QV we discussed above relate to whether it is likely in practice to implement the outcome that maximizes the total willingness-to-pay in the population. Another set of questions, raised in the contributions by Ben Laurence, Itai Sher and Josh Ober, is whether this is a desirable outcome as well as whether the outcome it implements alone is a sufficient basis for judging the legitimacy of using QV to make collective decisions.

Laurence and Sher investigate the conditions under which the outcome maximizing total willingness-to-pay (WTP) is the outcome that maximizes aggregate utility and under what conditions the preference of the majority implemented by 1p1v may achieve a more socially desirable outcome. The basic weakness they highlight with the WTP-maximizing outcome is that WTP is not the same as the utility gain individuals experience from a collective decision. Instead, it is (roughly) the ratio of that utility gain to that individual's marginal utility of money, a quantity that declines with income. Thus, the WTP-maximizing outcome overweights the preferences of the wealthy, who can afford to buy more votes. If all individuals have the same wealth or if preferences over the issue in question are uncorrelated with the marginal utility of wealth, then then WTP-maximizing outcome also will maximize social welfare. However, if an issue is polarized along income lines, the WTP-maximizing outcome may be inferior to its alternative, which in turn may be chosen by 1p1v. Thus, the relative merit of QV and 1p1v turns, in this analysis, on whether heterogeneity in preferences given wealth or heterogeneity in wealth driving heterogeneity in WTP is a more important determinant of the divergences between QV and 1p1v on the range of issues confronted in practice.

Arguments of this general sort are the most common objections we hear to the use of QV in the form described originally by Weyl and advocated by us in Posner and Weyl (2014). While in principle we acknowledge the force of these concerns, we believe that in practice they are unlikely to be very important counterweights to the benefits of QV for several reasons. First, QV could be adopted in conjunction with other reforms that would compensate the less-wealthy for any loses they would incur, as well as making issues polarized by wealth less important. We advocate such a coupling (with a system of making most property commonly held) in a forthcoming book titled *Radical Markets*. Second, our society rewards the accumulation of wealth because of the taxes and other social spillovers (like jobs) it generates despite the inequality it produces. We see no reason that this policy should not be extended to at least some collective decisions, such as those over local public good provision or zoning policies, even if not to the most fundamental issues like constitutional design or redistribution. Finally, votes could be made more expensive for wealthy voters, perhaps priced as a percentage of their taxable wealth or income rather than in absolute dollars. While Laurence and Sher argue this would be complicated and politically challenging, we don't see it as much more difficult than implementing any other income-based tax, and certainly less challenging than gaining political support for QV in the first place.

However, we acknowledge that most people do not find these arguments persuasive. To make matters worse for QV, Laurence and Sher, as well as Ober, fault it for disregarding even more basic senses of fairness and legitimacy. Laurence and Sher argue that democratic legitimacy requires that citizens have equal opportunities to influence political outcomes. Standard 1p1v satisfies this requirement because, on its face, it treats individuals equally. On the other hand, QV inherently gives greater opportunities for influence to the wealthy. Thus, QV is democratically illegitimate.

However, we would note that there are that QV attends to many important values much more effectually than 1p1v does. One is the importance, embodied in the principles of freedom of

speech and assembly, of allowing individuals to choose how strongly they wish to express their political views. Recent decisions by the Supreme Court have affirmed the importance of allowing very free use of private resources for political participation, even if these may imply inequalities in the abilities of individual to participate politically. As Hirschman (1982, p. 104) put it, "the 'one man one vote' rule gives everyone a *minimum* share in public decision-making, but it also sets … a maximum or *ceiling*; for example, it does not permit citizens to register the different intensities with which they hold their respective political convictions and opinions." QV, by allowing individuals truthfully to express this intensity, expands the freedom of individuals to participate in the political process. Even if this added freedom may be outweighed by the lesser equality allowed by QV, it seems to us a basic error to exclude it from the balance.

Ober makes a related series of claims, but draws more heavily on the social meaning of voting and its roots in the classical Greek tradition than on the sort of absolute liberal position adopted by Laurence and Sher. First, Ober argues that for many political issues, all individuals in the public have equal interests. He highlights national defense and security as such an area. He argues that in such an area, QV cannot be applicable because it is concerned with aggregating different intensities of preference, which inherently are the same on many issues of common interest. This claim strikes us as implausible for two principal reasons. We find it hard to imagine *any* issue on which interests truly are identical. Even regarding national defense, individuals are likely to differ fundamentally in the values involved (the appropriateness of violence and intervention in another nation's affairs, the value of property, the degree to which foreign countries should be seen as a menace at all, and so on). As Mouffe (1999) argues in her classic critique of deliberative democracy, the concept that there exist any fully shared interests relies on an extremely constricted understanding of acceptable political outlooks in a pluralistic society.

Moreover, QV, unlike 1p1v, allows not just for the expression of different degrees of *interest* but also different degrees of *expertise and knowledge* about issues. One might therefore expect it to perform better than 1p1v in such common interest settings when degrees of expertise differ across individuals and issues. In fact, Ober (2013) argues explicitly that in common interest settings it is important to put greater weight on the views of experts, in their relevant domain of expertise, than on the weights of non-experts. While Ober offers a variety of both objective and voting-based means to arrive at these expertise weights, QV potentially offers an incentive-compatible way for individuals to *reveal* their own expertises. In fact, in a pure common interest setting QV always will outperform 1p1v (at least in the best-case equilibrium) because it allows each individual to express more and, on top of that, the Pareto-dominant equilibrium of common interest voting always is one that maximizes social welfare subject to the constraints of what the voting language allows to be expressed (McLennan 1998). However, no formal analysis of QV in the setting with partially shared interests and differing information exists to date because of the complex game theoretic issues that arise in this context (Federsen and Pesendorfer 1996, 1997, 1998). Thus, at present the relative merits of QV over 1p1v in the settings Ober is interested in are more suggestive than they are definitive. However, it seems odd based on this to say that QV simply is inapplicable; simple mechanisms are often useful in many settings beyond the stylized models that are used to motivate them, as Weyl's robustness contribution highlights.

Second, Ober argues that QV basically is inconsistent with the democratic concept of giving equal respect to all individuals and that violating this condition may lead to civil strife and political instability. We again differ from Ober here on two points. We are skeptical that the superficial egalitarianism of the 1p1v rule carries the weight that Ober attributes to it. History is replete with formally egalitarian rules, such as "separate but equal", with dramatically inegalitarian consequences that, when replaced with formally inegalitarian rules, such as affirmative action, both promoted and testified to a commitment to equality more clearly. While we cannot here make a definitive

argument that QV would be more egalitarian in effect (for that see Posner and Weyl 2015), we do find the rejection of QV *prima facie* based on its failing to fully enforce equality in influence overly superficial. Further, we note that Ober himself seems to endorse this perspective in his work on weighting expertise (Ober, 2013). However, in that work he relies on esteem of peers or some objective imposition of expertise weights to determine how individual opinions should be weighted in aggregation. It strikes us that, relative to this external set of weights, the manner in which QV arrives at expertise weights, based on individual expressions of willingness to sacrifice resources for, say, on an issue of common interest, actually testifies more clearly to the possibility that regardless of reputation any individual may turn out on a particular issue to be expert. That is, among weighted voting schemes, QV is most egalitarian in allowing the self-, rather than external, expression of expertise weights.

Third, Ober argues that by introducing money in some form into the political process QV may remove all restraints on its use in any forum and thus encourage extra-system vote-buying and collusion that will undermine the operation of QV. This is a serious concern. While QV is somewhat robust against collusion, as Weyl shows, and would be more so using the ingenious schemes of Park and Rivest, a general expansion in manipulation of the rules would doubtless harm the operation of QV. Whether or not it is possible to build a culture around QV in opposition to collusion as strong as the one existing in democracies against vote buying remains to be seen. That said, most market economies have quite strong norms against collusion among competitors in marketplaces. Perhaps this ethos could be extended to QV. This, once again, is one of the important sociological issues, like those raised by Kaplow and Kominers, that is only likely to be worked out in the process of small-scale experimentation with QV.

5 Solutions and domains robust to normative challenges

Laurence, Sher and Ober all find versions of QV using artificial currency that can be spread over multiple decisions unobjectionable even in large-scale political choices. This is a reaction we have encountered frequently among academics and members of the public alike. As a result, we suspect that in the near-to-medium term most applications of QV are likely to use artificial currencies that allow relatively narrow tradeoffs among defined sets of collective choices rather than exchanges between private and public goods.

Such applications also avoid most of the concerns we discussed in Section 2. Because no real money is used in these versions of QV, no incentive exists to manipulate the agenda in order to receive refunds. Furthermore, because all uses of the artificial currency will be on issues where the chance of pivotality is low – and it seems plausible that most voters will exhaust their budgets (as we will discuss further below) – concerns about some voters being induced by interest groups to vote more than is rational seem unlikely to cause significant problems. Perhaps because of the range of objections they avoid, three of the contributions to this volume concern precisely this style of application.

The contribution by Daniel Benjamin, Ori Heffetz, Miles Kimball and Derek Lougee lays the theoretical foundations. It discusses the HZ mechanism, which they call Normalized Gradient Addition (NGA), and its relationship to QV. Some of our discussion in Section 1 derives from their analysis, though we differ from their interpretation in some subtle ways. In particular, they identify ways in which QV and NGA may each be seen as applications or adaptations of each other.

To understand this relationship, it is first necessary to define what "QV with artificial currency" is. We do not know who exactly to attribute this concept to, as it has been discussed in a variety of literature since the proposal of QV, but perhaps the first extended discussion was ours

(Posner and Weyl 2015). The basic concept is that if multiple issues are up for vote, each voter may buy as many votes as she wishes on each issue *subject to the constraint that the sum of the squares of her votes equals a constant* (perhaps the same constant, if we wish to be egalitarian across voters). Benjamin et al. extend this concept to a context with continuous public goods by making the change in the continuous public good (its increment) proportional to aggregate votes. They then relate this to NGA, in which each individual reports a vector along which they would like to move the public goods, these vectors are normalized (their lengths shortened or lengthened to all be of standard Euclidean length one), added and the public good is moved in this summed direction. These mechanisms are the same as long as the QV budgets are equal because Euclidean length is determined by the sum of the squared lengths in each dimension, so that exhausting equal QV budgets is the same as expressing a normalized vector.

However, as Benjamin et al. highlight, far less is known analytically about such formulations of QV than about the basic binary decision version with money.[3] Even with a series of binary decisions, we do not know if some version of Lalley and Weyl's results about efficiency carry over to a setting with a series of referendums, votes on which are traded using an artificial currency. If they did, it would have to be based on some more narrow notion of Pareto efficiency among the set of public goods, as such a construction does not allow for external trade; however, even this more limited claim has not been analyzed.

Matters get subtler still regarding the setting with continuous goods as Benjamin et al. point out. One possibility is to consider, as HZ do, a dynamic process wherein the mechanism repeatedly is applied taking small steps to reach an equilibrium where all vectors cancel out. In this case, even if real money were used, there exists no analysis of when, perhaps in large populations, the incentives to manipulate this convergence process are small or approximately aligned with the social interest in achieving (Pareto-) efficient outcomes. Another possibility is to apply only a single step of the mechanism, as Benjamin et al. describe in their contribution. However, if such a step is very small, it achieves little good and if it is not very small it runs the risk of running into the limits of allowed policies or failures of the local approximation on which the linearity of utility in the increment is based. In short, while the NGA/HZ mechanism is intriguing with a strong motivation based on existing results on QV, it is one about which our theoretical knowledge remains quite limited.

Despite incomplete formal theory, the final two contributions to the special issue consider more practical deployments of QV. The first, by Posner and Nicholas Stephanopoulos, takes up the challenges posed by Laurence, Sher and Ober directlt. In particular, they consider the

---

[3] Benjamin et al. also consider whether the NGA can be adapted to the goals of the basic form of QV: to make a binary decision allowing trade between real money and influence on that decision. Their motivation, beyond relating the mechanisms to each other theoretically, is to derive a version of QV in which individuals may report their preferences directly rather than having to buy votes in proportion to those preferences. This may be useful because it avoids individuals having to estimate their chances of their being pivotal. The only existing form of QV that is "direct" in this sense was proposed by Goeree and Zhang (Forthcoming) and applies only to the non-generic case when an election ends in an exact tie. Benjamin et al. use NGA to derive a version that works in the generic case when one alternative clearly is superior to the other. Like the Goeree and Zhang mechanism, however, this requires the designer to have very fine knowledge of the value distribution in order to choose the parameters appropriately. By contrast, QV requires no such details, even if it does demand from (very rational) participants speculation about their chances of being pivotal. However, as Weyl shows in his contribution, even if participants do not all share common views about this, are not very rational in forming their beliefs about it or are not even motivated by their chances of being pivotal, QV can perform quite well. These more direct mechanisms, by contrast, are simply ill-defined in such cases and thus cannot be applied. This is why we share the assessment of Benjamin et al. that this is "is not an implementation of NGA that we would actually recommend."

implementation of QV (which they call "modified QV" or mQV) based on artificial and equally allocated currency for elections in the United States. Their analysis grapples with the concerns of Laurence, Sher and Ober in two ways. First, it proposes a means of implementing QV at large scale to which Laurence, Sher and Ober's objections do not apply and which they are openly sympathetic towards. Second, most of the paper is devoted to analyzing the extent to which this implementation would promote values other than the welfarist aims that motivated QV, goals of competitiveness, participation, minority representation and individual liberty that have legal foundations related to the non-utilitarian values Laurence, Sher and Ober invoke to critique basic QV.

In Posner and Stephanopoulos's vision, all voters would be allocated artificial currency, say, at the age of 18 and would receive some periodic replenishments of their holdings of it. This currency could be used for any US election at any level of government to buy votes quadratically on a range of issues, such as referendums, local elections and national elections. While Posner and Stephanopoulos outline how this system might work, they leave many details to be determined. How persistent would the currency be and how often replenished? What form would the electoral system take interacting with this periodically replenished bank (e.g., would legislators be elected in single-member districts or at large)? Would QV be used by legislators themselves in legislative bargaining?

Instead, most of their piece focuses on two questions. First, would such a system, with some configuration of these details, be constitutional? Second, to what extent would it promote judicially recognized values, beyond social welfare. On the first question, they argue that while very strong practical problems would be raised about applying QV with real money to US elections, mQV would pose few if any constitutional issues. In their view, mQV in no way impairs the ability of the financially poor to exercise the franchise. Indeed, mQV has compelling benefits that the Supreme Court would be likely to acknowledge as a counter-weight to most potential constitutional objections. Most importantly, voting systems that would seem to cause strictly greater constitutional problems related to 1p1v's principles have been upheld in the past by the Supreme Court. In particular, the system of cumulative voting, under which voters have a budget that they can allocate linearly across different candidates, was upheld despite theoretical (Mueller 1973, 1977; Laine 1977) and empirical (Haley and Case, 1979) findings that such systems lead individuals to put all their votes on a single issue or race and thus violate 1p1v much more radically than would occur under mQV. Furthermore, Posner and Stephanopoulos argue that in some areas, such as primaries, not even this minimal concern about applying mQV would arise as it would be the purview of parties rather than governments.[4]

Most of the Posner and Stephanopoulos contribution therefore focuses on the extent to which QV would promote non-welfarist values around elections. First, they argue that mQV could improve the representativeness of districts by reducing the efficacy of strategic gerrymandering of districts. Because QV accounts for the intensity as well as the direction of preferences, it would be dangerous and likely ineffective for parties to construct districts that narrowly favor themselves. Because intensity may vary from election to election, the parties would run considerable risk of losing control and even when they did succeed would be likely to elect more moderate candidates in these less intense districts. Second, they argue that mQV likely would increase the competitiveness of elections, or at least the degree to which this competitiveness is unpredictable and varies across

[4] However, as Posner and Stephanopoulos highlight, the party and primary systems might change significantly under mQV. The current two-party US system is to some extent an outgrowth of the necessity of forming viable coalitions created by the current first-past-the-post plurality voting system (Duverger 1959). In a system like mQV that would be less sensitive to various paradoxes of 1p1v's rules, parties might be less necessary or at least more fluid and primaries thus might play a less central role.

time, by incorporating the additional uncertain variable of district intensities over and above partisan affiliations within districts. This would likely lead to greater turnover and accountability.

Third, they consider the effects of mQV on minority representation. Because mQV allows minorities to express their political preferences more than proportional to their size, mQV would allow them to at least occasionally win fights with the majority in favor of candidates they strongly support or against candidates they strongly oppose. In districts comprising a majority of non-minority voters, this would very likely lead to a larger number of minority representatives. However, the reverse could happen in "majority minority" districts. Thus, while QV likely would help minority voters secure their most highly valued interests, it would not necessarily or always directly promote minority representation. This highlights a broader theme of the Posner and Stephanopoulos argument: QV often promotes non-welfarist goals, but does so to the extent that these non-welfarist goals align with welfarist aims and especially in the cases where the greatest welfare would be lost by failing to affirm the non-welfarist goals. As such, its effects on these aims *per se* may be only slightly positive, but is likely to be very positive if they are seen as subsidiary goals that eventually aim to achieve broader welfare gains, such as the protection of the valued interests rather than just the literal representation of minorities.

Fourth, Posner and Stephanopoulos argue that mQV raises the value voters derive from voting and thus is likely to raise voter turnout, engagement and effective access to the franchise. This cuts quite directly against the grain of Kaplow and Kominers's arguments that QV may induce "headaches" because of the complexity of the voting process and thus deter voting. Clearly one difference is that mQV does not involve real money, but it is not clear whether this would raise or lower the complexity. As we discuss shortly, existing empirical evidence about engagement with mQV systems does not appear to show a strong effect in one direction or the other, but these reaults are in settings where individuals are not impacting outcomes directly. This suggests that QV's headaches are probably unlikely, but it will be hard to know if Posner and Stephanopoulos's claims about greater engagement hold up until QV is deployed in higher-stakes settings.

Finally, Posner and Stephanopoulos suggest that QV, beyond its tendency to encourage campaigns to focus more on shifting preference intensities, may actually offer a new method for public finance. They argue that a system of taxes and/or subsidies on campaign contributions could be implemented so that the amount received by candidates is proportional to the square root of contributions made, rather than linearly proportional to these contributions. Such a funding scheme would both have some of the welfare benefits of QV itself and would help address a number of concerns about corruption and inequality in the ability to exercise free speech that exist at present because it would lead to the magnification of small contributions while dampening the impact of large contributions. The system, at least in the form with taxes, doubtless would face serious constitutional challenges. However, the authors offer another example of how QV, far from being the vector for inequality that Laurence, Sher and Ober suggest, may actually promote a more egalitarian political process.

While Posner and Stephanopoulos's vision makes the case that a socially acceptable form of QV can be deployed at large scale, it seems likely that even mQV is many years from an electoral debut. In the intervening time, we have been working to use mQV on a more limited scale in the private sector to experiment, learn and improve, as well as create economic value. The vehicle for this practical deployment has been Collective Decision Engines (CDE), a start-up we co-founded with Kevin Slavin. CDE has produced proprietary software that implements mQV through a user interface, "weDesign", which has proven easy for participants, even those with very limited education, to navigate successfully. CDE is using this software for a variety of private sector applications, but our primary focus has been on an area where a strong demand for determining individuals' truthful preferences already exists: polling and market research.

The final contribution to this volume, by David Quarfoot (our chief data scientist), Douglas von Kohorn (our head of engineering), Kevin Slavin, Rory Sutherland (who has no affiliation to CDE), David Goldstein (our Chief Executive Officer) and Ellen Konar (an advisor) reports the results of our first large-scale deployment of mQV in the field. In addition to the specific results obtained, their work is of great interest because it is the first deployment of any form of QV for substantive research and social purposes of which we are aware, rather than just in the laboratory (see Goeree and Zhang Forthcoming) or classroom (Cárdenas et al. 2014). As such, it offers the first test of whether any form of QV can offer substantive applied value.

The motivation for the application is to consider whether weDesign offers a useful alternative to the standard method of eliciting preference intensity in surveys using the standard Likert scale. In this method, participants are asked to express an opinion ranging from strong disagreement to strong agreement on (usually) a seven-point scale. While Likert has a long track record, an important and consistent limitation is that it does a poor job of separating out the views of respondents with very strong preferences, as these tend to be "top-coded" at the extreme ends of the scale, leading to preference distributions with an unnatural W shape (Hamilton, 1968). Our deployment of QV was based on the expectation that, by making the expression of extreme opinions not only costly but differentially so (because of the quadratic nature of costs), we would recover a richer portrait of the intensity of respondent preferences across the full range of intensities.[5] To do this, Quarfoot et al. constructed a survey of Americans based on ten recent politically sensitive issues and assigned participants randomly to groups wherein they were asked to express their opinions on these issues using Likert, weDesign or both.

The first finding that Quarfoot et al. report relates to Kaplow and Kominers's concerns about QV headaches. Quarfoot et al. find that survey completion rates for participants randomly assigned to weDesign arms of the study were indistinguishable from those assigned to the Likert and joint Likert-weDesign segments; all were around two-thirds of those who began the survey. They also find that weDesign participants spent about a third longer engaging with the survey than those assigned to Likert. Furthermore, they find that this time was devoted to revising vote allocations to proposals already considered, suggesting a greater thoughtfulness in the nature of engagement with the weDesign survey. Together, these findings suggest that Posner and Stephanopoulos's hope that QV would create greater engagement and incentive to vote seems to have some limited support and the idea of QV headaches seems to be refuted, at least tentatively. These results also suggest that, from a purely practical perspective, weDesign can collect a set of responses successfully.

The second finding is that weDesign largely accomplishes the goal of replacing W distributions of preferences with bell-shaped preference distributions, avoiding top-coding of preferences and teasing out their full distribution. This can be seen in the histograms of responses under the two methods. Even more interesting were the changes in the responses of individuals who gave very extreme responses on all or nearly all questions under Likert. These respondents showed a wide range of behaviors under weDesign: some continued to be quite strong on all questions while others allocated their budgets to one issue disproportionately. Whether these more finely grained preferences reflect meaningful information that was being missed by Likert, or whether they simply reflect the constraints imposed by the method itself, is less immediately obvious.

Quarfoot et al.'s third finding aims to provide some evidence to resolve this question. They asked respondents to indicate at the end of the survey the issues on which they would like to receive

[5] Other budgeted methods, based on linear budgets, have been shown to give rise to extreme behavior similar to Likert, though in many cases even worse as respondents simply place none of their budget on issues other than those of most interest to them (Haley and Case, 1979).

additional information. Given that the survey also asked for an email address, this can be seen as a proxy for taking action on the issue because they reasonably could expect to receive more information. Quarfoot et al. find that the number of votes cast in weDesign, even within the levels typically assigned by those top-coded in Likert, continues to provide significant predictive value on the probability of clicking that box. While certainly an imperfect proxy, this finding suggests that weDesign is picking up some meaningful signal about preference intensity among these strong supporters or opponents of particular issues.

Finally, Quarfoot et al. found that opinions among groups using the different survey designs were quite consistent across the two methods, *outside of the extreme top-coded range*. That is, weDesign gives a portrait similar to Likert of the opinions of all respondents except for those that are very extremist in their opinions. However, among those with extreme views, weDesign revealed some information that seems consistent with commonly held intuitions that were hidden in Likert. For example, while Likert implies that Hispanics and non-Hispanics hold identical views about deportation of illegal immigrants, the weDesign survey revealed significant differences *at the extreme tails of the preference distribution*. Many more non-Hispanics are strong supporters of deportations and many more Hispanics are strong opponents, but this difference is masked by top-coding in Likert. Another interesting finding is that Likert implies that opinions about women's pay equality is much more divergent among Democrats of different party affiliation strengths, while under weDesign preference intensities on this issue were quite uniform among Democrats regardless of intensity of party affiliation. This finding is consistent with pay equality being an issue that signals social commitments strongly, but not one on which respondents are willing to sacrifice influence on issues more important to them.

In sum, weDesign appears to be a practical, engaging and thought-provoking instrument for survey participants that reveals some additional information about the preferences of extreme respondents that is hidden by Likert, while retaining the informational content of Likert among more moderate respondents. This result suggests that in areas where the preferences of extreme individuals are important in determining policies positions (such as trying to reach compromise in a divided country) or product design (such as trying to avoid product failures based on concentrated opposition), weDesign may provide significant added value at limited cost. In other cases, where the general weight of opinion among those who are not that passionate is of greater moment, the benefits of weDesign may be less. Thus, the usefulness of weDesign may to a large extent depend on the number of cases for whic the intensity of extreme support or opposition is relevant, a question that is being explored at present through a number of applications of weDesign CDE currently is undertaking.

One additional encouraging result in the Quarfoot et al. test relates to Masur's contribution. While Masur proposes to use a version of QV with money, his application context (public surveys to determine valuations) is otherwise quite close to those tested by Quarfoot et al. This indicates the feasibility of Masur's application. Furthermore, like the other applications we consider in this setting, we suspect that Masur's would not raise serious concerns of the sort highlighted by Laurence, Sher and Ober. Because dollar valuations already are in use in BCAs, applying QV to truthfully elicit that information does not seem to raise ethical problems about expanding the scope of money's influence beyond its appropriate sphere.

6 Summary and path forward

Where does all of this leave us regarding the practical promise of QV in promoting the public good? The contributions to this volume raise probing and challenging questions about QV on both the practical and normative dimensions. However, because of the simplicity and flexibility of QV, there

appear to be a variety of areas to which all authors agree so that it can be applied in the near-term at both low risk along all of those dimensions and with the expectation of at least some appreciable gain. Leading among these are survey research (in both the public policy and product design domains) and contingent valuation surveys. A surprising consensus also emerged about the desirability of applying mQV to a broader set of elections and traditional QV to corporate governance and other domains where money already is in use. This suggests a rich and promising set of avenues for experimentation from which we expect to learn much both about the potential benefits of QV and about how the challenges raised here might be overcome.

However, many issues arose that both researchers and practitioners will have to work out before QV can reach its potential. Many of these came directly out of the challenges raised by contributions to this volume.

First, despite being apparently the most practical applications of QV, very little is known about mQV or even about QV in the setting Masur analyzes wherein it is applied repeatedly to put ever narrower bounds on valuation of a public good whose price may vary over time. Similarly, a natural application of QV, only briefly discussed by Posner and Stephanopoulos, is to settings with multiple mutually exclusive candidates or where several candidates must be elected from a larger set of options. While we proposed a potential rule for this in Posner and Weyl (2015), no results have been proven about this rule or any other variant of QV in this setting, though Weyl now is pursuing such analysis with Nicole Immorlica and Katrina Liggett.

Second, no analysis of QV at present allows for partially common interests about which individuals have varying expertise and information, an issue that Ober rightly highlights as of central importance. The game theoretic issues involved are extremely subtle, but we are hopeful that researchers will be able to analyze them in the future.

Third, while weDesign reveals clearly some information hidden by Likert, it was surprising to us how closely the information revealed by Likert tracked that in weDesign, at least among moderates. This suggest that mechanisms now exist that express preference intensity with some fidelity and that they already may be having some important effects on collective decision-making. Further empirical investigation of the extent to which our political and social system incorporates preference intensity effectively, and determination of the scope for further improvement using QV, will be crucial to pinning down the domains in which QV has something important to add.

Finally, and most importantly, the contributors to this volume raised a rich set of questions regarding the sociological implications of QV, the political organizations that would be built around it and the implications it would have for the culture of collective decision-making. In our view, these questions are the most challenging, uncertain and likely most consequential surounding the mechanism. They are not ones we know of any simple way to address short of large-scale experimentation (or perhaps thought experimentation by applying artistic imagination). We hope that if initial applications of QV continue to be successful, both social researchers and social activists will engage with these questions and help build a set of social institutions around, as well as social critiques of, QV ensuring that it serves the public good most effectually. If so, such thinkers and activists will find a rich store of questions to consider in this volume.

Arrow, K. (1951). Social choice and individual values. New York: John Wiley & Sons.

Attiyeh, G., Franciosi, R., & Isaac, R. M. (2000). Experiments with the pivot process for providing public goods. Public Choice, 102(1–2), 95–114.

Bowen, H. R. (1943). The interpretation of voting in the allocation of economic resources. The Quarterly Journal of Economics, 58(1), 27–48.

Budish, E. (2011). The combinatorial assignment problem: Approximate competitive equilibrium from equal incomes. Journal of Political Economy, 119(6), 1061–1103.

Cárdenas, J. C., Mantilla, C., and Zárate, R. D. (2014). Purchasing votes without cash: Implementing quadratic voting outside the lab. http://www.aeaweb.org/aea/2015conference/program/retrieve.php?pdfid=719.

Clarke, E. H. (1971). Multipart pricing of public goods. Public Choice, 11(1), 17–33.

Duverger, M. (1959). Political parties: their organization and activity in the modern state trans B. and R. North. London: Methuen.

Feddersen, T., & Pesendorfer, W. (1996). The swing voter's curse. The American Economic Review, 86(3), 408–424.

Feddersen, T., & Pesendorfer, W. (1997). Voting behavior and information aggregation in elections with private information. Econometrica, 65(5), 1029–1058.

Feddersen, T., & Pesendorfer, W. (1998). Convicting the innocent: the inferiority of unanimous jury verdicts under strategic voting. American Political Science Review, 92(1), 23–35.

Goeree, J. K., and Zhang, J. (2016). One man, one bid. Games and Economic Behavior, ttp://dx.doi.org/10.1016/j.geb.2016.10.003.

Groves, T. (1973). Incentives in teams. Econometrica, 41(4), 617–631.

Groves, T. and Ledyard, J. (1977a). Optimal allocation of public goods: A solution to the free rider problem. Econometrica, 45(4), 783–810.

Groves, T., & Ledyard, J. (1977). Some limitations of demand-revealing processes. Public Choice, 29(2S), 107–124.

Haley, R. I. and Case, P. B. (1979). Testing Thirteen Attitude Scales for Agreement and Brand Discrimination. Journal of Marekting, 43(4), 20-32.

Hamilton, D. L. (1968). Personality Attributes Associated with Extreme Response Style. Psychological Bulletin. 69(3): 192-203.

Hausman, J. (2012). Contingent valuation: From dubious to hopeless. The Journal of Economic Perspectives, 26(4), 43–56.

Hirschman, A. O. (1982). Shifting involvements: Private interests and public action. Oxford: Martin Robertson.

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

Laine, C. R. (1977). Strategy in point voting: A note. The Quarterly Journal of Economics, 91(3), 505–507.

Lalley, Steven P., and Weyl, E. Glen (2016). Quadratic Voting. Available at: https://papers.ssrn.com/sol3/Papers.cfm?abstract_id=2003531.

Mailath, G. J., and Postlewaite, A. (1990). Asymmetric information bargaining problems with many agents. Review of Economic Studies 57(3), 351–367.

McLennan, A. (1998). Consequences of the Condorcet Jury Theorem for beneficial information aggregation by rational agents. American Political Science Review, 92(2), 413–418.

Mouffe, C. (1999). Deliberative democracy or agonistic pluralism? Social Research, 66(3), 745–758.

Mueller, D. C. (1973). Constitutional democracy and social welfare. The Quarterly Journal of Economics, 87(1), 60–80.

Mueller, D. C. (1977). Strategy in point voting: Comment. The Quarterly Journal of Economics, 91(3), 509.

Ober, J. (2013). Democracy's wisdom: An Aristotelian middle way for collective judgment. American Political Science Review, 107(1), 104–122.

Petit, P. (2012). Analytical philosophy. In R. Goodin, P. Pettit, & T. Pogge (Eds.), A companion to contemporary political philosophy (pp. 5–35). Oxford: Blackwell-Wiley.

Posner, E. and Weyl, E. G. (2014). Quadratic voting as efficient corporate governance. University of Chicago Law Review, 81(1), 251–272.

Posner, E. and Weyl, E. G. (2015). Voting squared: Quadratic voting in democratic politics. Vanderbilt Law Review, 68(2), 441–499.

Roth, A. E. (2008). What have we learned from market design? Economic Journal, 118(527), 285–310.

Samuelson, P. A. (1954). The pure theory of public expenditure. Review of Economics and Statistics, 36(4), 387–389.

Tideman, T. N., and Tullock, G. (1976). A new and superior process for making social choices. Journal of Political Economy, 84(6), 1145–1159.

Tuck, R. (2012). History. In R. Goodin, P. Pettit, & T. Pogge (Eds.), A companion to contemporary political philosophy (pp. 69–87). Oxford: Blackwell-Wiley.

Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. The Journal of Finance, 16(1), 8–37.

Weyl, E. G. (2012). Quadratic vote buying. http://goo.gl/8YEO73.

Willig, R. (1976). Consumer's surplus without apology. American Economic Review 66(4), 589–597.