

Chapter One: Reinterpreting p —A New Theory of How Individual Votes Contribute to Electoral Outcomes*

“If a clod be washed away by the sea, Europe is the less, as well as if a promontory were, as well as if a manor of thy friend’s or of thine own were...”

—John Donne (1959 [1623], 108)

“Everybody wants to rule the world.”

—Tears for Fears (1985)

1) INTRODUCTION

What is the probability that one vote will affect the outcome of an election? The conventional assumption—almost uniformly shared in the voting literature—is well known: The larger the electorate, the smaller the probability that any individual vote will affect the outcome, with that chance essentially equal to zero in all mass elections of modern democracies. From the individual-level point of view, therefore, participating in elections is said to be effectively meaningless, or more precisely, instrumentally useless. This chapter, however, argues that the conventional answer is wrong, or more precisely, that it represents a demonstrable mistake in normative reasoning. In its place is proposed an alternative interpretation of the probability of one vote affecting an election outcome, an interpretation that not only seems preferable from a normative perspective, but which is arguably also more empirically plausible than the conventional view.

This alternative interpretation of the efficacy of individual votes in large elections has crucial consequences for theoretical conceptions of voter turnout, in addition to having potentially broad relevance to ideas about the nature of rational behavior and the

* This chapter was published (with slight changes) as an article in *Election Law Journal* 14 (2): 111-135 (2015), available from Mary Ann Liebert, Inc., publishers: <http://dx.doi.org/10.1089/elj.2014.0263>.

motivations for participating in collective action. However, this is not just a matter of theory, for as discussed in the introduction to this dissertation, and as detailed below, the interpretation of how individual votes contribute to outcomes also has important implications for election law and policy. Normative concepts of the meaning and purpose of voting in a representative democracy inform the design of electoral institutions, and the rules and procedures for administering elections therefore reflect fundamental conceptions of democratic values (Thompson 2002, vii-viii). In short, elections put democratic theory into practice. Overall, this work argues that the conventional theory proclaiming the instrumental futility of electoral participation results in tremendous underestimation of the potential value of voting at the individual level. The alternative interpretation offered here thus helps constitute a defense of a more expansive participatory approach in the practice of elections. At the same time, this new interpretation highlights the importance of competitive elections for effective representation, and so it also lends support to electoral reforms based in the competitive school of democratic theory.

The probability of one vote affecting the outcome of an election is one of the most widely cited statistics in the voluminous voting literature. Anthony Downs was perhaps the first to formally consider “the probability that any one citizen’s vote will be decisive” to the outcome of a large election. While indicating that this probability is never exactly zero, Downs famously concludes that “under most circumstances, it is so negligible that it renders the return from voting...infinitesimal” (Downs 1957b, 146). Following Downs, Riker and Ordeshook’s (1968) canonical work codifies the individual decision to vote or abstain in the well-known calculus of voter turnout: $pB - C + D$. Formalizing the rational choice approach to the turnout decision, pB represents the probability that an individual’s vote will affect the election outcome (p), multiplied by the expected utility differential

between the candidates on the ballot (B), while the C term represents the costs of voting to the individual, and D represents any utility gained from the act of voting independent of the outcome—such as gratification from complying with a perceived duty to vote, or other “expressive” benefits.¹ The calculus designates that if the sum of these terms is positive, the individual votes, while if negative, the individual abstains (Riker and Ordeshook 1968, 25).

In this chapter, p is analytically isolated from the other elements of the calculus in order to focus on what exactly it means for an individual vote to have a causal effect on an election outcome. It is, however, important to note how the interpretation of p influences interpretations of the other elements of the calculus: The value of p affects B directly, but it can also affect how C and D are construed, for if pB is assumed to be negligible, yet individuals are observed voting, the assumption must be that C is low, or that D is high (or both), as discussed below. As mentioned in the introduction, the approach of this dissertation is not to focus on the calculus as an empirical model of the turnout decision, although there may be testable hypotheses associated with the new interpretation on p proposed below. Instead, the calculus is utilized just as a framework for analyzing the factors that influence individual decisions to participate or abstain from voting—here specifically, the expected probability of having a causal impact on the election outcome. The goal of this work is thus not empirical analysis, but rather normative evaluation of the concept of individual causation in elections, which leads to specification of a new formal model for this element of the turnout decision, and to implications for participatory norms and for election law and policy.

¹ The D term may also be said to represent the “consumption” benefits of voting, as opposed to the “investment” benefits represented by B (see Ferejohn and Fiorina 1974, 526).

While continuing empirical research on electoral institutions and voting behavior may be useful and important, the type of normative inquiry conducted here is arguably needed more urgently, as stronger foundations in democratic theory are essential to the development of more principled and coherent election laws and policies. There appears to be fairly wide agreement among legal scholars that the normative underpinnings of American electoral institutions are muddled and disorderly, particularly when viewed through the lens of U.S. Supreme Court decisions in this area (Karlan 1993; Gardner 1997; Issacharoff and Pildes 1998; Gerken 2002; Charles 2007; Post 2014). There are, however, differences of opinion on how stronger theoretical foundations could be constructed. Much of the debate has centered on whether constitutional issues in election law should be decided using traditional individual rights balancing tests (Hasen 2003), or whether a structuralist approach that explicitly considers core democratic goals and values is more appropriate (Pildes 2004b; Elmendorf 2008).² In broad terms, these debates trade on fundamental conceptions of democratic representation, including participatory and competitive variants of emphasis in democratic theory. This chapter does not take a definite side in these debates, but instead describes how the new interpretation of causation in voting has distinctive implications under these varying theoretical approaches.³

Section 2 below discusses the so-called paradox of voter turnout and reviews some of the previous attempts at reinterpreting the calculus to resolve this purported

² Others have argued that the rights versus structure debate is not particularly useful (Charles 2005), that it is not generally applicable across different types of election law cases (Fishkin 2011), or that it fails to capture important aspects of representational democratic theory (Stephanopoulos 2014). A few scholars have expressed skepticism about calls for more explicit theoretical foundations in election law jurisprudence, worrying that judicial decisions tend to unnecessarily “lock in” a specific approach to democratic theory (Cain 1999; Lowenstein 2002).

³ This is perhaps consistent with the intermediary approach of Daniel Farber, who indicates that election law issues generally implicate both structural and individual rights concerns (Farber 2004).

problem. Section 3 articulates the critique of the conventional interpretation of p leading to the paradox, and then introduces a new causal logic of voting based primarily on theoretical insights from Richard Tuck's innovative work, *Free Riding* (2008). Section 4 develops and elaborates on Tuck's basic theory by proposing a formal model and simulating how this new interpretation of p could be calculated in actual elections. Section 5 responds to initial critiques of this alternative approach and considers some normative implications of how individual causation in elections is conceived. Section 6 discusses practical implications of this new model, reexamining a few central issues in American election law and policy in light of the new understanding of how individual votes contribute to outcomes. Section 7 concludes and points out directions for future research in this area.

2) THE PARADOX OF VOTER TURNOUT

Riker and Ordeshook formally define p as the probability that an election is expected to result in a tie, which would allow a single individual to cast a tie-breaking vote that is decisive—or pivotal—to the outcome,⁴ and they show how the value of p is higher the closer the outcome is expected to be (Riker and Ordeshook 1968, 31-32). Owen and Grofman further specify p as the “subjectively estimated probability that [one’s] vote will change the election outcome from what it would have been had [one] not voted” (Owen and Grofman 1984, 312). They formally model p as a function of both the size of the electorate and the expected closeness of the outcome, and they

⁴ This assumes the total number of voters, including the individual voter, is odd. If the total is even, p represents the probability of an individual vote causing (rather than breaking) a tie. Riker and Ordeshook model p for this case as well, but an odd total can be assumed for simplicity, given any substantial number of voters (see e.g. Owen and Grofman 1984, 312). The terms “decisive” and “pivotal” are used interchangeably.

demonstrate how in a two-candidate race forecasted as a virtual dead heat, the size of the electorate would have to be less than 10,000 for p to rise above 0.01, while for an evenly split electorate of 1 million voters, p equals approximately 0.0008 (315). Owen and Grofman thus conclude, “no matter how close an election, it is almost inconceivable that one vote will prove decisive” (318).⁵

Empirical studies of actual election results echo the conclusion that the value of p is always negligible in any large election. Mulligan and Hunter calculate the average probability of casting a pivotal vote in a U.S. House or state legislative race to fall somewhere between $2/n$ and $1/n$, where n is the number of voters in the electoral district (Mulligan and Hunter 2003, 51). Similarly, Gelman, Katz, and Bafumi find that the chance of casting a pivotal vote in a U.S. presidential election is on the order of $1/n$, where n is the size of the state electorate (Gelman et al. 2004, 669). Such studies utilize complex statistical methods to derive the value of p , but their conclusions are clear and unequivocal in devaluing the instrumental efficacy of an individual vote. Gelman, Silver, and Edlin thus compare the act of voting to buying a lottery ticket with about a 1 in 10 million chance of winning (Gelman et al. 2012, 324; see also Jankowski 2002). Thomas Schwartz describes the upper limits of p in the closest of elections, and he concludes, “Saying that closeness increases the possibility of being pivotal...is like saying that tall [people] are more likely than short [people] to bump their head on the moon” (Schwartz 1987, 118). Likewise, Paul Meehl grimly asserts that the “chances of determining who

⁵ Owen and Grofman’s method assumes that the expected closeness of the election is known—or at least estimated—as an exact figure. Fischer (1999) describes an alternate method of calculating p that utilizes polling data to create a margin of error around the expected outcome, which yields even smaller probabilities. For example, in a population of 1 million voters who appear—based on a sample of 300—to be evenly split, Fischer’s method would put the chance of casting a decisive vote at about 0.000014. Thus, the chance of one vote being pivotal is arguably even more “inconceivable” than Owen and Grofman estimate. Fischer attributes his method to Good and Mayer (1975), while Owen and Grofman’s method builds on the work of Beck (1975) and others.

becomes president are of about the same order of magnitude as [the] chances of being killed driving to the polls—hardly a profitable venture” (Meehl 1977, 11).

These are the types of descriptions that form the background for what is known as the paradox of voter turnout. Why—at least from a perspective of instrumental rationality—would so many individuals bother to engage in an activity with only a negligible chance of having any causal efficacy? This is the problem that led Riker and Ordeshook (1968) to introduce the *D* term into the calculus, representing the non-instrumental motivations that seem necessary to explain why (relatively) large numbers of people do turn out to vote in mass elections.⁶ Such motivations may also justify the decision to vote in terms of “collective rationality,” where an individual’s group membership produces social pressure that makes participation rational notwithstanding a lack of individual-level instrumental efficacy (Uhlener 1989; Morton 1991).⁷ More generally, the *D* term represents expressive motivations for voting, which may have social-psychological or “existential” foundations (Schuessler 2000), and it also includes moral impulses toward cooperative behavior, which may be associated with a “rule-utilitarian” approach to the turnout decision (Harsanyi 1980; Feddersen and Sandroni 2006a). Furthermore, explanations of turnout that rely on the motivation to increase a preferred candidate or party’s political “mandate” may also be included in the *D* term,

⁶ Downs originally proposed a somewhat similar solution to the paradox, suggesting that many individuals decide to vote after considering that the democratic system would collapse if no one participated in elections (Downs 1957a, 261-262). However, this solution runs back into the paradox, since individuals should reason that their individual vote would be highly unlikely to be pivotal to saving democracy (Fiorina 1976, 392). Riker and Ordeshook address the paradox more directly by defining *D* in terms of ethical, or social-psychological gains in utility. They thus list several types of personal satisfaction that an individual might derive from the act of voting, even indicating that certain voting costs could be perceived by some individuals as benefits (Riker and Ordeshook 1968, 28).

⁷ Note that leaders of large groups, or others with wide spheres of influence (including parties and candidates), might conceivably influence enough votes to have a non-negligible chance of being pivotal to an election outcome (Uhlener 1989, 402; see also Shachar and Nalebuff 1999). However, such attempts to influence outcomes through exertions of social pressure are not directly encompassed by the voting calculus, which addresses the individual motivation to turn out and cast a single vote.

since any expected utility in this case is derived independently of the electoral outcome (Guerrerro 2010; Mackie 2014).⁸

To some, however, the *D*-term solution to the paradox is not a sufficiently rational explanation of the turnout decision, and it was criticized early on as a tautological and non-predictive model (Barry 1970, 13). Moreover, the assumption that the decision to vote is characterized primarily by expressive benefits raises issues beyond the question of whether a *D*-term solution is theoretically satisfying. The absence from the voting decision of any direct concern for an instrumental effect on the outcome can be said to open the way for “electoral irrationality of the most basic kind” (Brennan and Buchanan 1984, 199), and expressive voting could lead individuals to vote for “morally unsavory” policies they would not choose if they thought their vote might actually impact the result (Brennan and Lomasky 1985, 204). Less perniciously, but perhaps more practically relevant, a purely expressive motivation for voting counsels always casting one’s ballot in favor of the most preferred candidate or party, regardless of that candidate or party’s chances of winning, given that one vote could never conceivably affect the outcome (Owen and Grofman 1984, 322). There may thus be serious consequences if voting is characterized as only expressively—and never instrumentally—rational.

More generally, however, addition of the *D* term to the calculus may be useful to symbolize how the turnout decision can incorporate both instrumental and expressive elements (Fiorina 1976, 393). In the real world it might be difficult or impossible to distinguish such motivations from one another (Fischer 1996, 172), and such is not the present purpose. The intention here is rather to demonstrate how voting could indeed have instrumentally rational motivations, because the value of pB may in fact be non-

⁸ However, such explanations presumably remain subject to the paradox, as individuals should reason that their one vote would make no appreciable difference in the value of any political mandate.

negligible, and the D term therefore need not hold all the “action” in the calculus, as often assumed (Fiorina 1976, 393; see e.g. Copeland and Laband 2002). This is not to imply that addition of the D term is the only avenue of escape from the paradox. To the contrary, John Aldrich’s oft-cited conclusion is that turnout is best explained as a “low-cost, low-benefit decision,” and that voting is therefore not a genuine collective action problem at all (Aldrich 1993, 265). Furthermore, others have suggested that including altruistic or social benefits in the value of B could yield a utility high enough to make voting instrumentally rational notwithstanding an infinitesimally small value for p (Margolis 1982; Jankowski 2002; Edlin et al. 2008).

Moving beyond solutions based on interpretations of B , C , or D , there have also been more direct attempts to avoid the paradox by reinterpreting or modifying the understanding of p . In addition to their D -term explanation, Riker and Ordeshook also suggest that voters might simply be greatly overestimating their probability of being pivotal due to widespread “propaganda” about the importance of individual votes in close elections (Riker and Ordeshook 1968, 39).⁹ Others have suggested that voters either fail to understand or even consider their probability of casting a pivotal vote, instead relying on broad predictions of the chance that their preference will prevail (Hinich 1981; Peters 1998). These simplified approaches may or may not have empirical support,¹⁰ but they

⁹ Similarly, Gregory Brunk (1980, 550) writes, “[T]he reason why so many people vote is because democratic societies systematically feed their citizens false information about the utility each individual personally gains from the franchise.” Downs actually suggests that the value of p could be “significant if [one] thinks the election will be very close” (Downs 1957b, 146), but it is not clear whether he thinks this valuation would be a statistical error, and he does not mention it resulting from propaganda.

¹⁰ There is in fact some evidence that individuals either overestimate or fail to consider the probability of casting a pivotal vote (Blais 2000, 62-70; Blais et al. 2000, 191; Klor and Winter 2006; Esponda and Vespa 2010). In general, there appears to be little empirical support for pivotal thinking outside of small-group experiments, and some contradictory evidence even at that level. Experimental studies that raise doubts about pivotal theory include: Tyran (2005), Klor & Winter (2007), Coate et al. (2008), Esponda & Vespa (2010), Großer & Schram (2010), and Morton & Tyran (2012). Experiments finding evidence for pivotal voting in smaller groups include Fischer (1996), Levine & Palfrey (2007), Duffy & Tavits (2008), and Feddersen et al. (2009).

clearly fall short of a theoretically grounded resolution of the paradox that remains consistent with an instrumentally rational approach (Dowding 2005, 452).

Other p -based solutions to the paradox do not rely on assumptions of miscalculation or misunderstanding on the part of voters. One such approach is Ferejohn and Fiorina's (1974) minimax regret model, which frames the turnout decision as taking place under conditions of generalized uncertainty (rather than specifiable risk), where the prevailing assumption is that it is rational to minimize the chance for an occurrence that would cause the maximum amount of regret. This approach effectively eliminates the need to consider p , as minimax regret logic appears to reduce the calculus to just the B and C terms (Ferejohn and Fiorina 1974, 528).¹¹ However, the minimax model is not generally accepted as a viable solution to the paradox (Dowding 2005, 449; Geys 2006b, 21; Blais et al. 1995). For one thing, election outcomes are generally not completely—or even very highly—uncertain, as probabilities are quite often assigned to expectations about results. More fundamentally, however, the logic of pivotal theory in fact finds its way back into the minimax calculus: Maximum regret is said to arise from abstaining when one's preferred candidate loses by one vote (or ties), but the chance of that occurrence remains infinitesimal in any large election. As Aldrich concludes, “[T]he positive prediction of turnout in minimax regret swings on the same set of (still just as wildly implausible) circumstances as in the calculus of voting” (Aldrich 1997, 381).

Aldrich reaches a similar conclusion regarding attempts to rationalize turnout by reinterpreting p in a strategic context, indicating that “game theoretic models of turnout are apparently ‘driven’ by the same basic features as the individual decision-making models” (Aldrich 1997, 383). Other proposed solutions model the turnout decision as an

¹¹ Ferejohn and Fiorina are skeptical regarding the need for adding the D term to the calculus, asserting that “it is rational for many citizens to vote even if they neither distort their individual impact nor place a direct value on the act of voting” (Ferejohn and Fiorina 1974, 526).

adaptive learning process grounded on previous voting experiences, thereby interpreting p in a way that makes participation more likely to be instrumentally rational (Kanazawa 1998), or eliminating the need to calculate p altogether (Bendor et al. 2003). However, these models depart from the central assumption of forward-looking utility maximization in the calculus. Moreover, the main assumption of these models—that the turnout decision is based on perceptions of positive or negative outcomes associated with past actions—generally fails to correspond with empirical evidence regarding the motivating factors of turnout (Dowding 2005, 451; Guinjoan et al. 2014).

There have also been some notable attempts to resolve the paradox by stepping outside conventional causal logic. In place of the individualized expected utility maximization employed by the causal decision theory of the calculus, Robert Grafstein (1991) suggests an “evidential” decision theory in which potential voters base their actions on expectations regarding the behavior of other similarly situated individuals. Accordingly, if one believes that most members of one’s social group are likely to vote for the same candidate or party, this increases the “conditional” expected utility of one’s vote, notwithstanding the fact that one’s decision to vote has no causal impact on anyone else’s actions. Grafstein explains this as an application of “stochastic dependence,” which he identifies with generalizing from one’s own behavior in a manner suggestive of the Kantian imperative, asking, “What if everyone did that?” (Grafstein 1991, 1006). This logic, however, has been faulted for irrationally confusing causal with “diagnostic,” or “correlational” contingencies, which Quattrone and Tversky refer to as “the voter’s illusion,” in which individuals mistakenly believe that their own decision to participate has some effect on the turnout decisions of others (Quattrone and Tversky 1988, 733-734). Jon Elster similarly criticizes this approach as an erroneous application of “everyday Kantianism” that rests on a form of “magical thinking” (Elster 1989, 195).

Grafstein's theory does indeed seem problematic to the extent that it relies on assumptions that cannot be supported within the conventional rational choice framework of the calculus (Dowding 2005, 450).¹² Nevertheless, his approach is noteworthy for recognizing the potential role of collective interdependence in the individual turnout decision, and it perhaps points toward a more ethically grounded theory of the decision to participate.

In sum, while there have been many attempts at resolving the paradox of voter turnout, the conventional interpretation of p as the probability of casting a pivotal vote continues to pose a major obstacle to theoretical modeling of the turnout decision. Hence, Keith Dowding's comprehensive review of proposed solutions to the paradox describes these efforts as a "pathological quest for the holy grail of individually rational turnout" (Dowding 2005, 442). Dowding expresses doubts about whether an individual contribution can ever be conceived as having a "useful effect" on the outcome of a large election, maintaining that any perceived effect is likely "an expressive value masquerading as an instrumental one" (452). He also offers a practical critique of pivotal theory, indicating that no candidate would ever strive toward a one-vote margin, not just for prudential reasons, but because it would represent the weakest possible victory in terms of political mandate (Dowding 2005, 451-452). It does indeed seem clear that politicians and campaign strategists have incentives to challenge pivotal theory and mobilize supporters by arguing that their individual contributions could actually affect the outcome of the election. The question is, should prospective voters view this as mere propaganda, or might these urgings in fact allude to another way of interpreting the causal effect of individual votes?

¹² In subsequent work, Grafstein concludes that his notion of conditional expected utility ultimately "disables the core notion of free choice and self" (Grafstein 1995, 78), which certainly seems inconsistent with rational choice theory.

3) TOWARD A NEW CAUSAL LOGIC OF VOTING

As indicated, pivotal voting theory assumes that an individual's participation affects the result of an election only when it directly changes the outcome. This implies a counterfactual—or “but for”—logic of conditional causation, meaning that a given outcome is caused by an individual only if it would not have occurred without that individual's participation. However, it is also possible to conceive of causation taking place in the absence of a strictly counterfactual relationship. In tort law, for example, if two defendants separately started fires that each independently would have burned down the plaintiff's property, either one can be held fully responsible for causing the damage, even though neither was a “but for” cause of the result (Wright 1985, 1776). Another example is a firing squad, where no one member of the squad is a “but for” cause in the execution, but moral responsibility may still be assigned to each participant individually (Goldman 1999, 205). These situations may be referred to as cases of *overdetermined*, or redundant, causation.

Derek Parfit points out that overdetermined causation often generates normative dilemmas, as “mistakes in moral mathematics” can emerge from concentrating only on the effects of an individual act, while “ignoring the effects of sets of acts” (Parfit 1984, 70). The examples of the fire-starters and the firing squad are cases where each individual acting alone would have been fully sufficient to bring about the outcome, but Parfit also gives examples of overdetermination arising in contexts where collective action is required to achieve a certain goal (72).¹³ There may thus be reason to doubt whether the

¹³ In connection with voting, Parfit disputes the claim that “below some threshold, extremely small chances have no rational or moral significance” (Parfit 1984, 73). However, he invokes the conventional interpretation of p —assuming it to equal about one in a hundred million—and suggests that voting may still

counterfactual causal logic of pivotal theory is truly the most appropriate way to model the turnout decision. Nonetheless, hardly anyone seems to have questioned the assumption that instrumentally rational voters should only value being a counterfactual cause of the outcome, and more specifically, that they should discount their expected benefits by the probability that their vote will be pivotal.¹⁴ It is worth emphasizing at this point that defining the boundaries of rational action is a fundamentally normative endeavor. In the words of John Harsanyi, the concept of rationality is normative simply because “it points to what we *should* do in order to attain a given end or objective” (Harsanyi 1986, 83). Pivotal theory can thus be seen as setting a specific normative standard for rationality—a standard that, incidentally, has implications not just for voting, but for other collective action situations as well.

In one of the rare critiques of pivotal theory, Patrick Dunleavy pointedly asks, “What is rational about wanting to be a unique swing voter in contexts where large numbers of actors are involved?” (Dunleavy 1997, 56). Tracing the origins of pivotal theory to the works of Downs (1957a) and Olson (1965) applying the expected utility analysis pioneered by von Neumann and Morgenstern (1947), Dunleavy admits that it makes sense to assume that rational actors in a collective action situation should discount their expected benefits by some probability factor (Dunleavy 1997, 57-63). However, he forcefully rejects the assumption that this discount factor should be the probability of being pivotal to the preferred outcome, and he ultimately characterizes the desire to be pivotal within a large group as not only “not rational,” but even “pathological, almost psychotic” (81). His particular arguments against pivotal theory may not be very

be instrumentally rational given the expectation of an extremely large public benefit, thus relying on a *B*-term solution to the paradox (74-75).

¹⁴ Of course, many scholars have criticized rational choice theory as a model of political or social behavior (e.g. Green and Shapiro 1994; Udehn 1996). Few, however, appear to question the specific assumption that the desire to be pivotal is inevitably associated with rational choice in a collective action situation.

convincing,¹⁵ but Dunleavy makes at least one point that seems novel and noteworthy, and which suggests—albeit obliquely—a different way of thinking about what it means to have a causal effect on an election outcome.

Dunleavy mentions an obvious consequence of pivotal theory that seems to have gone largely overlooked: It is that no individual can ever be *solely* pivotal to an election outcome. The reasoning is simple: In any election decided by a margin of one vote (or tied), *everyone* who cast a vote for the winner (or either side if tied) would “simultaneously be equally pivotal” (Dunleavy 1997, 79). In other words, the crowning achievement of pivotal theory, the goal toward which all rational actors should presumably be striving—casting a vote that individually decides the election—is by definition always a shared experience. Dunleavy argues that the fundamental ideal of pivotality is thus basically unrealizable, since “many others voting or abstaining would have an equal claim to being the decisive actor” (79). Strictly speaking this could be mistaken, as one might still insist that it is rational to participate only to the extent that one’s vote will be pivotal, even if this distinction must always be shared with other voters. However, the requirement of always having to share any claim to pivotal causation hints at an alternate way of thinking about causing an election outcome.

Perhaps the pivotal situation is best framed simply as a representation of the closest possible election. Yet elections can of course be very close without being as close

¹⁵ Dunleavy argues that the concept of pivotality is problematic because it is subject to formal ambiguity, since there could be reasonable disagreement, even after the fact, about whether an individual’s vote was decisive to an outcome. The reason, he argues, is that a pivotal situation could also arise from the perspective of an abstainer whose participation might have affected the result, or from a voter contemplating voting against preference (Dunleavy 1997, 75). (He admits that voting against preference seems implausible, but he implies that it could happen as a result of ambivalence or uncertainty.) In the latter case, a pivotal situation could arise from a two-vote difference, as a changed vote by anyone who would otherwise have voted for the winner (assuming a two-candidate race) would create a tie. Dunleavy’s objections, however, do not challenge the basic assumption of counterfactual causation employed in pivotal theory, and the formal ambiguities he points out seem trivial in large electorates.

as mathematically possible. In fact, when a very large election is decided by 2 votes, or 10 votes, or 100 votes (or even say 537 votes), the outcome may be statistically indistinguishable from the pivotal situation. It is thus arguably a mistake to draw what is virtually an arbitrary line at the ultimate degree of closeness—i.e., a one-vote margin (or a tie)—and then to assume that participation becomes utterly ineffectual, from an individual perspective, at any greater margin. Instead, what seems missing is a way to define and measure how an individual vote could still represent a causal factor as an electoral outcome diverges from the pivotal situation.

As indicated, when an election is decided by one vote, everyone who voted for the winner—bracketing for now those who voted for the loser—was strictly pivotal to the outcome. When an election is decided by *two* votes, the conventional assumption is that *no one* was pivotal. Note how this is tantamount to asserting that each individual voter—holding all other voters constant—could have abstained without altering the outcome. However, while this may be true in a formal sense, practically speaking it seems problematic, for given a two-vote margin the outcome would have changed if any more than one person voting for the winner had abstained. More realistically perhaps, instead of saying that no one was pivotal in an election decided by two votes, one might instead assert that *everyone* (who voted for the winner) was pivotal, *except for one voter*. Moreover, since votes are perfectly fungible and effectively simultaneous, there is no way of knowing which voter cast the one non-pivotal vote. Therefore, if X equals the number of votes for the winner, everyone voting for the winner had a $1/X$ probability of having cast the non-pivotal vote, or an $(X-1)/X$ probability of having been pivotal to the outcome.

Admittedly, the term “pivotal” is being stretched beyond its intended meaning here. A vote in the true pivotal situation is not only individually *necessary* to an election

outcome, it is also, in some sense at least, individually *sufficient* to bring about the outcome.¹⁶ In the case of an election decided by a margin of two votes, there exists a set of $X-1$ votes that were each individually necessary, and that were collectively—though not in any sense individually—sufficient to cause the outcome. Although in practice no one voter can be identified with certainty as having been part of that set, nevertheless, everyone who voted for the winner has an $(X-1)/X$ probability of having been *individually necessary* to a set of votes that were *collectively sufficient* for the outcome. In the terminology of Richard Tuck, we can identify an “efficacious set” of votes, “with each vote in the efficacious set having true causal efficacy in bringing about the result” (Tuck 2008, 44). In a plurality voting system, this set of votes will always contain exactly one vote more than the number of votes for the second-place finisher, since this is the number of votes needed to win.¹⁷ It is thus simple to calculate the *ex post* probability that an individual vote was in the efficacious set by dividing the number of votes in this set—the second-place finisher’s tally plus one—by the number of votes the winner actually received. The probability of being in the efficacious set can then be said to represent the chance that an individual vote causally contributed to the election outcome, in the sense of having been individually necessary to that outcome.

While the above calculation may be quite simple, the underlying concept is another matter. At issue here is the essential nature of the collective action problem posed by opportunities for participation in very large groups. In fact, Tuck’s pathbreaking work,

¹⁶ One vote, even if pivotal, can never truly be individually sufficient, since an election by definition requires collective action. Nevertheless, from a *ceteris paribus* perspective, a pivotal vote can be seen as individually sufficient to bring about the desired outcome. This is related to the fact that being pivotal is always a shared experience, yet each individual claims counterfactual causal power over the outcome.

¹⁷ This also holds true for majority-rule elections with only two candidates, and even with more candidates the two top finishers may be said to win—in the sense of advancing to a runoff—by obtaining at least one vote more than the third-place finisher. The idea of an efficacious set of votes under a system of proportional representation (PR) is more complicated, and calculating values for p under PR is a complex problem beyond the current scope (see Blais et al. 2014).

Free Riding (2008), poses a direct challenge to Mancur Olson’s well-established theory, in *The Logic of Collective Action* (1965) regarding the negligibility of individual contributions to very large enterprises. Tuck explains that voting is a type of collective action problem that is not actually a good example of a “genuinely Olsonian problem,” because there is always a specific threshold that determines the outcome of an election (Tuck 2008, 44). The real Olsonian problem, according to Tuck, is a situation portrayed by the paradox of the sorities, an ancient riddle asking when a collection of grains of wheat becomes—or ceases to become—a “heap.” The paradox, of course, is that one can never identify a threshold, for at any point it could be argued that one grain more or less would make no difference (67). Tuck explains how in a true sorities situation with no threshold, one might argue—as does Olson—that any individual contribution has only a negligible effect, since there is no identifiable point when an additional contribution clearly matters. In cases like elections, however, where there is always a definite threshold determining the outcome, Tuck argues that the collective action problem described by Olson never arises, for rational individuals in such situations have incentives to coordinate their activity to ensure meeting the threshold (48).¹⁸ Since elections do not represent a genuine Olsonian problem, Tuck concludes that “any analysis of voting which presumes that no individual vote has causal power over the result unless it is pivotal is mistaken” (44).¹⁹

¹⁸ Such coordination should in fact be relatively easy in the electoral context, as the essential organizational structure for collective action is fully in place, at least for the major political parties, and all individuals must do is obtain a ballot and cast a vote for their preference.

¹⁹ One might argue that elections actually do represent an instance of the sorities paradox, at least in practical terms, because if an initial result is close enough it will often elicit a recount and/or litigation, making the actual threshold for winning uncertain. Given this reality, combined with the practical limitations of election administration, any large election with a close enough result might be described as “a statistical tie,” with no identifiable threshold for when a result becomes that close (see e.g. McCaffery et al. 2004a, 5). This is not a problem for Tuck’s theory, however, for he eventually concludes that even in a true sorities situation, rational behavior entails acting as if there *is* in fact a definite threshold, although it cannot be precisely identified (Tuck 2008, 95, 208; see Runciman 2008). In any event, an electoral outcome is

While Tuck provides the theoretical foundation for a new understanding of the collective action problem of voting, his analysis nevertheless stops short of a comprehensive account of the implications of efficacious set causation in the context of elections. Tuck broadly associates his position with an article by Alvin Goldman (1999) entitled, “Why Citizens Should Vote: A Causal Responsibility Approach” (see Tuck 2008, 51). Goldman echoes the idea that overdetermined causal effects can be meaningfully ascribed in cases where counterfactual causation fails, and he seems to have independently arrived at the general idea of an efficacious set of votes existing within the larger set of votes for a winning candidate (Goldman 1999, 205-207). However, neither Tuck nor Goldman fully integrates the *probabilistic* context of voting into their causal theory. Goldman indicates in a footnote, “Causation can take place even in chancy situations, where merely probabilistic laws hold sway,” yet he puzzlingly concludes, “In the context of voting, however, we do not need to worry about probabilistic causation. Wherever an electoral outcome occurs, some set of votes is sufficient for the outcome” (Goldman 1999, 208).

Similarly, while Tuck originally references the “probability that [one] vote was part of the efficacious set” (Tuck 2008, 44), he later appears to abandon this probabilistic focus in favor of an emphasis on the rationality of “bandwagon” incentives that justify participation in cases of clear overdetermination. He thus concludes, “[I]t is precisely in the situation where it looks on the standard modern view as if my vote is unnecessary that I have a good reason to vote” (60).²⁰ Strangely, he appears to entirely ignore elections in

always at least formally determined by an exact threshold, and moreover, the practical uncertainty surrounding close elections can be modeled in a way that mirrors an exact threshold determination (Gelman et al. 2004, 674).

²⁰ Tuck again references the probability of being in the efficacious set in explaining that when one’s preference is an overwhelming favorite to win, the probability of being in that set might be small enough to justify abstention (Tuck 2008, 61). Since his primary focus is on an instrumental argument for voting in cases where one’s preference is expected to win, Tuck associates efficacious set causation with being

which the outcome may be uncertain, instead highlighting how individuals can have an instrumentally rational reason to vote even when they fully expect their preference to prevail. Tuck's point is an important theoretical innovation of its own, but it arguably misses another key consequence of the theory of efficacious set causation—the motivation it provides for individuals to participate in *close* elections. Furthermore, Tuck also neglects to discuss how his theory might apply in cases where an individual's preferred choice is expected not to win but rather to *lose*. A more comprehensive explanation of how efficacious set causation functions in these cases is needed in order to more fully integrate the probabilistic context of voting as *ex ante* decision-making. This more comprehensive explanation provides the basis for the new interpretation of *p* in the voting calculus.

4) DEVELOPING THE NEW INTERPRETATION OF *P*

The probability of one vote being necessary to an election outcome by forming part of an efficacious set can be illustrated with the result of a very famous close election: the 2000 U.S. presidential race in Florida. The official final tally was 2,912,253 votes for Al Gore, and 2,912,790 votes for George Bush. Leaving the Gore voters aside for the moment, the *ex post* probability that any individual Bush voter was in the efficacious set is $2,912,254/2,912,790$, or about 0.9998. Prior to the election, assuming it was uncertain who would prevail in Florida, with polls showing the race to be within the margin of error (and thus “too close to call”), any prospective voter should rationally have assumed, *ex ante*, that their vote had an effectively 100 percent chance of being necessary to their

sufficient, but not *necessary* to the outcome (101-102). Sufficiency, in Tuck's usage, seems to refer to collective sufficiency, not to the individual (*ceteris paribus*) sufficiency associated with casting a pivotal vote, while his version of necessity seems to imply strict—not probabilistic—necessity.

preferred outcome. In other words, every vote mattered, at least prospectively. This is presumably how campaign strategists think about close elections, and it's not clear why voters—at least those with clear preferences—should think much differently. Pivotal theory leads of course to a very different conclusion, as a negligible value for p effectively eliminates instrumental benefits in even the closest elections. Under this new interpretation, however, p can essentially be ignored in these cases.

The theory of efficacious set causation can thus justify instrumental participation in cases of close elections, as just illustrated, as well as when one's preferred choice is a favorite to win, as Tuck demonstrates in his emphasis on the rationality of bandwagon incentives. What of the case where one's preferred choice is expected to lose? If *ex ante* information predicts that one's favored candidate is essentially certain to lose (e.g., Nader voters in the 2000 election), then prospective voters should indeed rationally conclude that votes for that candidate would have zero probability of affecting the outcome. However, as the probability that one's preferred choice might prevail begins to increase, the logic of efficacious set causation takes over: Suppose that pre-election polls indicate the outcome of a two-candidate race is expected to be 45 percent for candidate X, and 55 percent for Y, leaving aside for now any margin of error. A prospective voter who prefers X might have a rational incentive to participate in this situation, since X is not certain to lose, even if that result appears more likely than not. What then is the *ex ante* probability that this voter's participation will be necessary for X to win? It seems clear that in this situation the chance that a vote for X will be necessary to the preferred outcome is 100 percent. Every vote for candidate X can be expected to be in the efficacious set in this case, because every vote is expected to be necessary in order for X to win. This is true even though *ex post*, if X happens to lose, every vote for X will have had a zero probability of forming part of an efficacious set, and will in fact have been instrumentally

useless. *Ex ante*, however, the fact that the odds appear to be somewhat—though not overwhelmingly—in Y’s favor means that the votes of every one of X’s supporters are needed if X is to have a chance at victory, so prospective voters who prefer X should interpret p as equal to 1.

In sum, the theory advanced here suggests that a prospective voter with a clear preference between two candidates should go through a two-stage decision-making process with regard to the p term: First, the individual asks: Is there a “realistic” chance that my preferred candidate can win?²¹ If the answer is no, there is indeed no instrumental motivation to vote. If the answer is yes, the individual then proceeds to ask: What is the chance that my vote will be necessary to that outcome by forming part of an efficacious set? If the preferred candidate appears more likely to lose, or if the election is essentially “too close to call,” then p is equal to one. If, on the other hand, the preferred candidate is expected to win, the probability of forming part of an efficacious set declines in proportion to the strength of that expectation. Tuck’s theory of bandwagon incentives indicates that calculation of this probability may continue well past the point where the preferred candidate is perceived to have any realistic chance of losing, but at some point that probability may become small enough that there could again cease to be any instrumental motivation to participate (Tuck 2008, 61).²²

²¹ Asking whether one’s preferred candidate has a “realistic” possibility of winning implies a behavioral assumption relating to a prospective voter’s beliefs apart from objective estimations of the probability (or probability distribution) associated with the expected outcome. In other words, the answer to this question will be idiosyncratic, as discussed below.

²² This implies another behavioral assumption regarding a prospective voter’s beliefs, not about whether the preferred candidate has a realistic possibility of losing, but rather about the idiosyncratic value of forming part of an efficacious set of votes for a winning candidate.

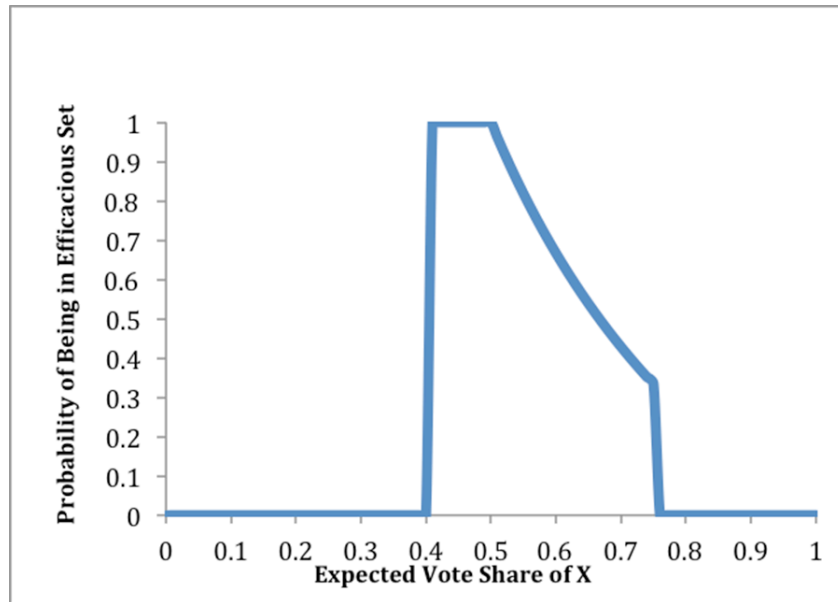


Figure 1.1: Value of p as Expected Outcome Changes

Figure 1.1 illustrates a simulation of the value of p in a large electorate as the expected share of votes for candidate X moves between 0 and 1. When X's expected share is small enough that the candidate is expected to lose for certain, p equals 0. However, there is a point where the expected loss of candidate X becomes uncertain, and a prospective voter begins to perceive that X has a realistic chance at winning. At this point there is a discontinuity, and p jumps from 0 to 1. Where exactly this occurs may be completely idiosyncratic, but Figure 1.1 assumes that it happens when the expected share of X rises above 40 percent. The value of p then remains at 1 until X becomes more likely to win, at which point it begins to descend, with p at each point equal to the expected number of votes for Y, plus one, divided by the expected votes for X.²³ As

²³ One might wonder whether p should instead continue to equal 1 past the expected vote share of 0.5 until the point where X's losing ceases to be a realistic possibility. This point might be assumed to fall somewhere around 0.6, in parallel to the discontinuity point on the losing side. However, the argument here is that p should be interpreted as 1 on the prospective losing side (at all points where there is still a realistic chance of winning) only because each vote is 100 percent certain to be necessary for X to win. On the prospective winning side p is attributed a distinct value, and there is no reason to assume a p of 1. Nevertheless, p will of course be very close to 1 when X is only a slight favorite.

indicated, there is presumably another point where the expected share is so high that voting for X begins to seem instrumentally useless, and p drops to 0 even though there remains a non-zero probability of forming part of the efficacious set. Note, however, the asymmetry: Where X appears more likely to lose, p shifts from 0 to 1 when a prospective voter perceives a realistic possibility of X winning—here when the expected share rises above 0.4. Where X is expected to win, however, a prospective voter might rationally pursue a chance at joining an efficacious set up to the point where bandwagon incentives run out, which Figure 1.1 assumes to occur when X 's vote share reaches 0.75, after which p returns to 0.²⁴

A key observation in distinguishing this interpretation of p from pivotal theory is that under this new approach the size of the electorate is effectively irrelevant in all but the very smallest elections: In other words, Figure 1.1 looks almost exactly the same whether the electorate has 100 or 100 million voters. In contrast to pivotal theory, where both closeness and size determine the value of p , what matters under this interpretation is essentially only expected closeness. Under efficacious set causation, p still equals 0 in some cases, but it may equal 1, or close to 1, when an election is expected to be highly competitive. Therefore, the perceived benefit of one's preferred choice prevailing in the election— B in the calculus—need not be reflexively discounted to zero. This resolves the paradox of turnout, at least for cases of relatively close elections.

²⁴ The chosen discontinuity points of 0.4 and 0.75 are admittedly somewhat arbitrary. The hypothesis here is that these discontinuities should exist, but obviously, empirical research is needed to verify their existence and location. On the more likely to lose side, the discontinuity might be hypothesized to fall somewhere between 0.40 and 0.45, assuming that most prospective voters would likely agree that a candidate polling over 45 percent has a “realistic” chance at winning, and the reverse at less than 40 percent. On the likely to win side, the discontinuity might be subject to greater variation based on how individuals value a chance at joining an efficacious set of votes. Note that the locations of these discontinuities could be seen as true instances of the sorities paradox, since they lack definite thresholds. Tuck's theory, however, would entail acting as if they in fact do have clear thresholds, even if they cannot be precisely identified (see *supra* n. 19).

A simple approach to calculating p as the probability of being in the efficacious set could also incorporate a margin of error from a pre-election poll, as follows: A prospective voter who favors candidate X over candidate Y (in a two-person race) would first note whether the margin of error crosses the 50 percent point. For example, polls might indicate that candidate X is expected to receive 54 percent of the vote, ± 5 percent. In this case, the probability of being in the efficacious set should be estimated at 1, since the expected outcome appears to be a statistical tie. If candidate X's predicted likelihood of winning is beyond the margin of error, the chance of a vote for X being in the efficacious set would be projected downward from 1 in proportion to the (mean) expected outcome, up to the point where the higher discontinuity falls outside the margin of error, and p drops to 0. Conversely, if X appears more likely to lose, p would equal 1 at all points where the lower discontinuity point falls within the margin of error. To illustrate, X may be expected to obtain only 38 percent of the vote, again with a margin of error of 5 points, meaning that X's expected share could really be as high as 43 percent. Arguably, it would not be beyond the realm of "realistic" possibility for a potential voter to conclude that X has a chance of prevailing in this situation.²⁵ In other words, it could well be worth the effort of voting for X, after considering expected costs and benefits, given that any vote for X in this situation is essentially certain to be necessary if X is to win.²⁶

²⁵ It might first be noted that most polls have a 95 percent confidence interval—meaning there is a 5 percent chance that the true value will fall beyond the margin of error (of about two standard deviations from the mean) in repeated sampling. More importantly, pre-election polls of "likely" voters measure only expressed intentions at the time surveyed, which are subject to various sources of bias, including the fact that respondents may not be very good at predicting their own likelihood of voting (Rogers and Aida 2014). As Fischer (1999, 273) explains, polls may have "non-sample errors, mainly because people do not vote the way they say they will." For these reasons, it arguably makes sense to construct another margin of error around the formal margin of error when using pre-election polls in *ex ante* instrumental turnout decisions.

²⁶ Interestingly, there could be an interaction between the expected benefits of an individual's preferred candidate winning (B in the calculus) and the location of the discontinuities in the estimate of p . If B is perceived to be very high—for instance, when one thinks the opposing candidate will cause great societal

A more complex estimation of p can be derived by using the predictions of pre-election polls to construct a probability distribution for the expected outcome, and then summing up the probabilities for each possible outcome discounted by its associated value of p . Hypothetically, suppose a poll sampled 100 voters out of an electorate of 1000, finding that 52 individuals preferred candidate X, and 48 favored Y. In this case, candidate X would be expected to receive 52 percent of the vote with a standard deviation of about 5 percentage points.²⁷ Applied to the population of 1000 voters, the mean expectation for candidate X would be 520 votes, with a standard deviation of 50 votes, yielding the normal distribution shown in Figure 1.2:

harm—one might consider participating even when the chance that one’s preferred candidate will win is relatively small. This perhaps suggests a logic reminiscent of minimax regret: The greater the chance for regretting a decision to abstain, the more one might want to minimize that risk by participating. When a candidate loses by a relatively small margin, supporters who abstained might be seen as having more causal responsibility for the adverse outcome (see Goldman 1999). (This of course assumes the rejection of pivotal theory, so that one might experience regret even with a margin greater than one.) The level of regret for abstaining would then presumably be greater the closer one’s preferred candidate came to winning, and the greater the strength of one’s preference for that candidate. Prospectively, therefore, one would want to minimize the possibility of “maximal” regret by regarding one’s vote as potentially necessary and participating even in cases where one’s candidate seems more likely to lose, particularly where the B term is perceived to be high. In formal terms, this could be represented by a leftward shift in the location of the lower discontinuity in the value of p . When a highly preferred candidate is favored to win, one might also experience some regret at not being part of an efficacious set, but it would not seem to be as “maximal” a regret as abstention when one’s preferred candidate loses.

²⁷ This estimation method draws from Fischer’s (1999) approach to calculating the value of p under pivotal theory. As Fischer (1999, 270) notes, in a poll for a two-candidate election (a binomial distribution), the standard deviation (in numbers of voters) is calculated by multiplying the sample size by the expected vote share of each candidate and taking the square root of the product: $\sqrt{100 * 0.52 * 0.48} \approx 5$.

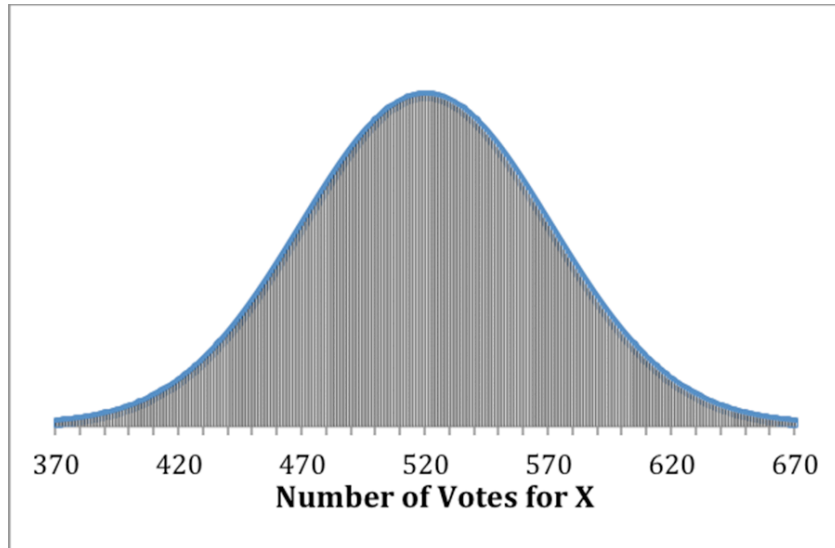


Figure 1.2: Probability Distribution of Votes for X in Electorate of 1000 ($n=100$)

The probability that candidate X will fail to win (losing or tying) is the area under the curve representing 500 votes or less, which is calculated by the cumulative distribution function to equal 0.344 of the entire area under the curve. The p associated with this probability would be 1, as indicated previously, except for points to the left of the lower discontinuity, where p is equal to 0. Assuming that the discontinuity point is at 0.4, the probability that X will receive 400 votes or less is calculated at 0.008, which is subtracted from 0.344 to yield an overall probability of being in the efficacious set—on the prospective losing (or tying) side—of about 0.336. Summing up the probabilities where X is expected to win is more complex, since the probability of being in the efficacious set changes at each point on the curve. Therefore, probability mass functions—each representing the probability of a discrete outcome—must be calculated for every possible winning outcome.²⁸ The associated p for each of these outcomes is

²⁸ For simplicity, this can be calculated only up to 3 standard deviations above the mean, here 670 votes, since the probability of X receiving more votes becomes infinitesimal. In any event, it would be calculated only up to the point of the upper discontinuity, here assumed to be 0.75, or 750 votes.

calculated by dividing the number of votes for candidate Y, plus 1, by the number of votes for X. The value of p at each outcome is then applied as a discount factor to the discrete probability associated with each outcome. Summing up the discounted probabilities for each outcome where X receives 501 votes or more (up to 3 standard deviations above the mean) yields an overall probability of 0.543. Adding this to the 0.336 calculated on the other side of the curve yields an overall probability of 0.879 that a vote for candidate X will be in the efficacious set.²⁹

The following formal equation represents the probability of being in the efficacious set (Pr_{ES}) with respect to a given distribution (μ), where X equals the expected number of votes for candidate X (in a two-candidate race), N equals the expected total electorate, while a and b are (respectively) the lower and upper discontinuity points.³⁰

$$\text{Pr}_{\text{ES}\mu} = \text{Pr}_{\mu} \left[X \leq \frac{N}{2} \right] - \text{Pr}_{\mu}[X \leq aN] + \sum_{X=\frac{N}{2}+1}^{bN} \frac{(N-X)+1}{X} * \text{Pr}_{\mu}[X]$$

Figure 1.3 applies this equation repeatedly as the expected vote share of X shifts between 0 and 1, illustrating a simulation of p calculated from a probability distribution, based again on a hypothetical sampling of 100 from an expected electorate of 1000, with discontinuity points set at 0.4 and 0.75. When the expected distribution of votes is evenly divided at a mean of 0.5, the overall probability of being in the efficacious set is

²⁹ If p drops to zero at 0.4, the probability of receiving 400 votes or less is small enough that it does not significantly alter the final calculation. However, locating the lower discontinuity point significantly above 0.4 would produce a greater effect on the final value. If it is instead located at 0.45, the overall probability of being in the efficacious set in this example drops to 0.807. For the upper discontinuity, the possibility of X receiving more than 0.75 of the votes is essentially 0, and the discontinuity point would have to move much further down from 0.75 to have any noticeable effect on the final value of p .

³⁰ This assumes an even number of voters. If N is odd, the summation term changes to: $\sum_{X=\frac{N+1}{2}}^{bN} \dots$

calculated at just above 0.9.³¹ The curve is again asymmetric, dropping more steeply where X appears more likely to lose since more of the distribution is being discounted to 0 under those expected outcomes. Correspondingly, p drops below 0.5 on the left-hand side of the curve just as X's mean expected vote share reaches 0.4, and p drops almost to 0 by the time that vote share reaches 0.3; however, on the right-hand side p drops below 0.5 when X's expected share is about 0.67, and does not settle to 0 until that share is above 0.8.

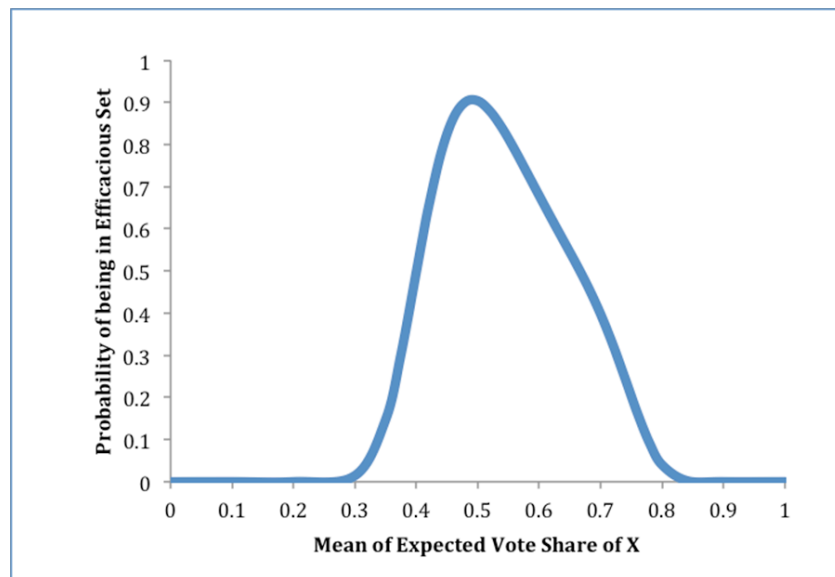


Figure 1.3: Value of p Calculated from Probability Distribution ($n=100$)

To conclude, none of the foregoing is meant as a claim that prospective voters can or should go through these types of calculations in practice, just as the statistical models of pivotal theory do not necessarily imply that voters can or should explicitly calculate their chances of being pivotal. Nevertheless, the simpler methods outlined above for

³¹ If the lower discontinuity is instead set at 0.45, the highest value of p —about 0.81—occurs near a vote share of 0.52, while for an evenly split electorate p is about 0.77. This is perhaps conceptually problematic, since one might expect the highest value of p to always occur when the race is closest.

calculating p could perhaps be understood by most prospective voters, in contrast to even the most basic calculations required to estimate p under pivotal theory. Yet the main point for present purposes is that the alternate interpretation described here leads to a starkly different understanding of the instrumental efficacy of voting. The initial theoretical claim is thus simply that when prospective voters consider the chance of their vote affecting the outcome of an election, they *could*—consistent with instrumental rationality—consider their chance of forming part of an efficacious set instead of considering their probability of being pivotal. However, there is also a stronger normative argument that the probability of having a causal effect on an election *should* be calculated in this manner, and that employing pivotal theory might actually represent a less rational—if not outright irrational—approach to the collective action problem of voting. This normative argument has vital implications for conceptions of the role of voting in democratic theory, and it also has important consequences for the legal and policy regimes associated with elections, as discussed in the following sections.

5) CRITIQUE AND NORMATIVE IMPLICATIONS OF EFFICACIOUS SET CAUSATION

In order to begin evaluating the normative implications of this new understanding of how individual votes contribute to electoral outcomes, it is helpful to review some of the critical response to Tuck's work. Notwithstanding the highly provocative nature of Tuck's conclusions—which pose a fundamental challenge to the conventional modeling of collective action problems—the response to *Free Riding* seems relatively muted. Reviews have been mixed, with some praising Tuck's innovative approach (Amadae 2008; Runciman 2008), while others respond skeptically to his argument against the Olsonian orthodoxy (Kuhn 2010; Thompson 2011; Congleton 2009). Jason Brennan

offers perhaps the most detailed criticism of Tuck's causal logic, directly addressing the theory of efficacious set causation in his book, *The Ethics of Voting* (Brennan 2011a; see also Brennan 2009).

Brennan advances several arguments against Tuck's position on the instrumental rationality of voting (Brennan 2011a, 28-34). Grounding his analysis, Brennan postulates the existence of two types of potential voters: Type-1 potential voters care not only about the electoral outcome, but they also care about being a causal agent of that outcome. Brennan's type-2 potential voters, on the other hand, care about the outcome but attribute "no special value to being the agent of causation" (29). His main argument is that abstention is the most rational choice for Type-2 individuals (who prefer candidate A), because: "Given what others are doing, voting for A and abstaining from voting for A are both sufficient for A to be elected" (33). However, Brennan never really defines the parameters of "what others are doing," and he does not mention close elections where the outcome may be uncertain. It may perhaps be understandable to miss the implications of Tuck's theory for uncertain elections, for as indicated, Tuck himself neglects to elucidate this matter. Brennan parenthetically states, "Recall that Tuck is not trying to argue that one should vote because there is some small chance one's vote will be decisive" (Brennan 2011a, 33). However, Tuck's theory does in fact imply that one should vote because—or more precisely, whenever—there is some *large* chance that one's vote will be necessary, if not "decisive," to the preferred outcome.³²

In cases of uncertainty, the logic of efficacious set causation suggests that participation could be instrumentally rational—subject to the *C* term of the calculus—

³² Brennan interprets efficacious set causation as the state of being "*minimally sufficient*" for the outcome (Brennan 2011, 29), but he does not contemplate it representing the probability that one's vote might be *necessary* to the outcome. Again, Tuck is unclear on this as well, since he focuses on the case where one's preferred candidate is fully expected to win (see *supra* n. 20).

even for Brennan's Type-2 individuals. For if these individuals truly do have a preferred outcome (the value of which is not outweighed by the costs of voting), they should want to do what they can to contribute—even in a very small way—to bringing about that outcome, regardless of how much “special value” they attach to their causal effect. The difference between Brennan's two types of potential voters might instead be exhibited in the location of the upper discontinuity in the probability of being in the efficacious set: Type-2 individuals could be expected to have more symmetry between their lower and upper discontinuities, as they would not care about having a share of causal credit for their preferred outcome in cases where that outcome is fairly certain. Yet assuming they do in fact care about the outcome, they should rationally want to participate in cases where it remains uncertain. Under Tuck's approach, caring about the outcome cannot be detached from attributing value to being an agent of causation, at least in cases of uncertainty, for if one truly cares about the outcome one by definition should want to influence it to the extent within one's power. In fact, Brennan's Type-2 individual appears to assume the counterfactual conception of causation employed by pivotal theory, taking for granted the virtual impossibility of affecting any election outcome, which makes Brennan's argument against Tuck seem circular.

Some of Brennan's other objections to Tuck's theory can be resolved by incorporating the other elements of the calculus. For example, he raises the existence of opportunity costs for the act of voting as an argument against Tuck's causal logic (Brennan 2011a, 31, 33; see also Mackie 2014, 45). However, all costs associated with voting—including any costs related to missed opportunities during time spent voting—are modeled in the C term of the calculus, and they should not directly affect estimation of p , whether under pivotal theory or efficacious set causation. Brennan also states that under Tuck's theory the individual utility of voting is equal to the probability that one's

vote will be in the efficacious set, multiplied by “the value of being in the efficacious set” (Brennan 2011a, 30). He then argues that “there is no obvious way” to determine the value of being in the efficacious set, citing Tuck’s rejection of the possibility of dividing up total utility among those with causal responsibility for bringing it about (Brennan 2011a, 184; citing Tuck 2008, 40-43). However, Tuck actually concludes that “each vote carries the full causal responsibility for bringing about the result” (Tuck 2008, 41). The utility associated with the probability of being in the efficacious set is therefore simply the differential utility derived from one’s preferred candidate winning—in other words, the value of B in the calculus. While it is true that estimating the value of B raises difficult issues, this has no bearing on the logic of efficacious set causation.

Brennan also argues that Tuck’s theory is fundamentally inconsistent with instrumental rational choice, which purportedly demands the conventional Olsonian logic. Brennan thus accuses Tuck of employing a theory of “rationality as effectiveness,” in which rational action is judged by *ex post* standards of whether success was achieved, regardless of the *ex ante* odds of achievement (Brennan 2011a, 32). Tuck is admittedly somewhat unclear regarding the probabilistic context of efficacious set causation, as indicated, but Brennan’s reading is nonetheless insupportable.³³ Tuck clearly maintains that the logic of efficacious set causation falls within the bounds of instrumental rationality, although it obviously modifies the conventional Olsonian understanding of what constitutes a rational choice in collective action situations (see Tuck 2008, 99-100). Furthermore, the extension of Tuck’s theory to elections with uncertain outcomes, as

³³ In accusing Tuck of employing rationality as effectiveness, Brennan cites Tuck’s statement that “the essence of instrumental action is, after all, that what we do is a *means* to an *end*, that is, *causes* it” (Brennan 2011, 32; citing Tuck 2008, 54). There is nothing in this quote, however, nor in its surrounding context (nor anywhere else in *Free Riding* for that matter), to indicate that Tuck endorses the concept of rationality as effectiveness, which as Brennan illustrates, leads to clearly irrational forms of decision-making.

developed here, reinforces the position of efficacious set causation within an instrumentally rational approach.

Brennan's critique helps reveal the normative implications of how Tuck's approach departs from the conventional assumptions of pivotal theory and the Olsonian view of collective action problems. Brennan submits his Type-1 and Type-2 potential voters as if they represented two equal attitudes, devoid of normative connotations. However, one might question the claim to greater rationality of the Type-2 position, where an individual desires to receive a benefit while attaching "no special value to being the agent of causation." This is of course the very definition of free riding: receiving benefits from the work of others while eschewing a contribution to the collective effort. The conventional Olsonian logic assumes that it is always most rational to have an overriding interest in trying to free ride whenever possible; hence the focus on not contributing unless one's participation is likely to be pivotal. However, it might be equally rational—or perhaps even more rational—to have an overriding interest in ensuring that a particular public good is provided, or that one's preferred candidate is elected. The theory of efficacious set causation provides a theoretical basis for questioning the conventional assumption that free riding is always the most rational response to a collective action situation, showing instead how it can sometimes be instrumentally rational to participate even in the largest of groups. In fact, the second half of Tuck's book reviews the extended intellectual history of collective action problems, demonstrating how prior to the mid-20th century the inclination toward free riding was viewed as a decidedly *irrational* response to these situations.³⁴

³⁴ As an example, Tuck characterizes David Hume's position on collective action problems as follows: "Psychological features, such as the propensity to think about short term outcomes, or a general ignorance of the instrumental point of collaboration, might induce people not to cooperate, but this (though understandable, and predictable on Hume's account of human character) was *an error in reasoning*" (Tuck 2008, 126, emphasis added).

In the individual turnout calculus, therefore, it may in fact make more sense—in effect be more instrumentally rational—not to discount the expected benefits associated with one’s preferred candidate (*B*) by the chance that one’s vote will be pivotal, but rather by the chance that one’s vote will be necessary to the result of an uncertain election, or might otherwise contribute to an efficacious set of votes. What then is truly the more rational approach, assuming that by rational one means optimally oriented for achieving a certain goal? It bears repeating, as Harsanyi indicates, that the designation of rationality is at root a normative judgment (Harsanyi 1986, 83; see also Zuckert 1995). Tuck likewise asserts that his argument “must be understood as a *normative* claim, and not at all (or at least not much) as a predictive claim about how human beings will as a matter of fact behave” (Tuck 2008, 111). Furthermore, pivotal theory sets a normative standard for rational behavior that has implications not just for voting, but also for collective action problems more generally. It is thus certainly worth questioning whether pivotal theory sets the best normative standard for how individuals should behave.

The standard of pivotal theory eliminates the value of very small effects by framing them as negligible and hence inconsequential. In normative terms, however, even imperceptible effects can matter greatly to ultimate outcomes. As noted previously, Derek Parfit (1984, 70) refers to the devaluing of small effects as a “mistake in moral mathematics.” The implications of efficacious set causation for the ethics of imperceptible effects may thus apply in many areas beyond voting. For example, this discussion may be relevant to the question of whether individuals can rationally justify minor contributions to environmental degradation on the basis of negligibility. In contrast to pivotal theory, which would endorse what seems like antisocial behavior in such cases, the rational standard set by efficacious set causation is more socially oriented, taking into account the collective effect of many individual causal acts. As Tuck makes clear, this is

not a standard meant to apply regardless of what others are doing; it explicitly accounts for the anticipated activity of other individuals (Tuck 2008, 207). In a sense, this echoes Grafstein's (1991) "evidential" theory of turnout described earlier, in which individuals base their actions on expectations regarding the behavior of others. It might thus also help vindicate the rationality of "everyday Kantianism," which Elster (1989) criticizes as a form of magical thinking. Similarly, it might not be a completely irrational "illusion" to incorporate diagnostic contingencies into the causal logic of voting, as Quattrone and Tversky (1988) assert.

The movement away from pivotal theory can also be seen as part of a larger shift in recent scholarship away from conventional rational choice theory, and toward what Pildes and Anderson (1990, 2214) call, "a socially-situated understanding of individual and collective rationality" (see also Monroe 2001). Proponents of this shift have argued that theories of rationality have real-world implications, as normative standards for rational choice can have practical effects on how people behave. Pildes and Anderson thus suggest that academic adherence to conventional rational choice theory "might promote resignation, complacency, or at worst, a longing for antidemocratic politics" (2214). Similarly, Lars Udehn argues that the economic approach to collective action influences public attitudes, stating that "people become more egoistic by being told that this is what they are" (Udehn 1996, 194). Tuck himself reaches the identical conclusion, asserting that "the prevalence in modern economics and political science of the idea that it is not instrumentally rational to collaborate in large groups may well have led people to adjust their conduct accordingly" (Tuck 2008, 115).³⁵

³⁵ In addition, at least two experimental studies have found that exposure to pivotal voting theory leads to more negative attitudes toward participating in elections (Brunk 1980; Blais & Young 1999). Blais & Young find that such exposure reduces actual turnout among college students.

It may therefore be especially troubling when prominent economists or economically oriented political scientists cite pivotal theory as a compelling reason against participating in elections. For just a prominent example, the widely-read “Freakonomics” column by Stephen Dubner and Steven Levitt in *The New York Times* implies that economists should be embarrassed to be seen in a voting booth, and that the reason much of the general public does vote is possibly because “we are just not very bright and therefore wrongly believe that our votes will affect the outcome” (Dubner and Levitt 2005). Similarly, the following exchange took place just before the 2012 presidential election on Dubner and Levitt’s radio program:

Dubner: So Levitt, how can you...tell the difference between a smart person and a not so smart person?

Levitt: Well, one good indicator of a person who’s not so smart is if they vote in a presidential election because they think their vote might actually decide who wins (Dubner and Levitt 2012).³⁶

The theory of efficacious set causation provides an effective counterpoint to such encouragement of abstention, setting what is arguably a better normative standard of rationality, not only for participating in elections, but for contributing to collective action more generally. Moreover, beyond these broad normative implications, this new interpretation of how individual votes contribute to electoral outcomes has important consequences for how particular theories of democracy are translated into electoral institutions, as discussed next.

³⁶ See also the pre-election blog posts in support of Dubner and Levitt’s position by Phil Arena (2012), and Kindred Winecoff (2012); but see the response to Dubner and Levitt by Andrew Gelman (2012), citing his work with Edlin and Kaplan explaining that voting can be rational if one assumes an altruistic interpretation of B in the calculus (Edlin et al. 2008). The B -based explanation of the paradox is not inconsistent with the p -based resolution proposed here, but note that the logic of efficacious set causation could apply even with purely self-interested preferences.

6) IMPLICATIONS FOR ELECTION LAW AND POLICY

As noted in the introduction, design choices for electoral institutions are linked to core principles in democratic theory, as election laws and policies reify interpretations of democracy in particular legal and administrative contexts. Two areas in which this practical implementation of democratic ideals may take place are discussed in this section, the first relating to participatory democratic theory, and the second relating to competitive theory. Without deeply engaging the extensive literature in these wide-ranging schools of democratic thought, it is possible to outline how the interpretation of individual causation in voting reflects on “the law of democracy”³⁷ and its institutional manifestations from the broad perspectives of these two theoretical approaches.

The paradox of turnout that emerges from pivotal voting theory represents a serious challenge to participatory theories of democracy, with potentially important consequences for how elections are structured. To begin, McCaffery, Crigler, and Just, in discussing the prospects for electoral reform in the United States, seem largely dismissive of the possibility for major reforms based in participatory theory, concluding that “Downs’s paradox lives” (McCaffery et al. 2004b, 232). When individual votes are viewed as only infinitesimally likely to have a causal effect on the outcome, it is presumably more difficult to endorse a position in democratic theory that places a high value on participation. Consequently, the new interpretation of p offers needed support to participatory theory, and it provides a stronger foundation for electoral reforms based in this school of democracy.

³⁷ The term is borrowed from the title of the noted election law casebook by Issacharoff, Pildes, Karlan, and Persily (2016).

Strengthening the theoretical foundations of participatory democracy could thus provide guidance in analyzing election law cases of the type Daniel Tokaji calls the “new vote denial,” which assess the constitutional validity of various administrative burdens on casting a vote (Tokaji 2006).³⁸ The issue of voter identification has become a flashpoint for debate in this area, though many other administrative procedures are also implicated, including registration requirements, provisions for absentee or early voting, and various other voting regulations and polling practices. As mentioned in this chapter’s introduction, while there is a general consensus that courts need better theoretical guidance for analyzing and deciding these cases, there remains disagreement as to what exactly that guidance should entail. The perhaps dominant approach traces to the work of Issacharoff and Pildes (1998) advocating a turn to “structuralism” in election law. Structuralists argue that burdens on the right to vote should not be analyzed through the conventional individual rights and equal protection modes of analysis, which involve balancing the rights of individuals against state interests served by burdensome election regulations. Instead, structuralists argue that courts should assess the aggregate effects of these regulations across population groups (Elmendorf 2008; Overton 2007), shifting the analytical focus to the “structural mechanisms for monitoring the proper operation of a constitutionalized system of representative democracy” (Gardner 2010, 457; see also Stephanopoulos 2014).³⁹

³⁸ For Tokaji, who builds upon on a voting rights typology originally proposed by Pamela Karlan (1993), the “new” vote denial cases signal a return to the jurisprudence of participation that characterized the early voting rights cases, before a shift toward aggregative interests emerged in later cases alleging collective vote dilution rather than individual vote denial (Tokaji 2006, 692, n. 15). More recent cases, like those concerning voter identification, are thus seen as returning the focus to the individual participatory interest in casting a ballot, as opposed to the emphasis on group interests implicated in earlier challenges to methods of aggregation in electoral districts (718).

³⁹ Stephanopoulos’ theory of representational “alignment” between the preferences of the median voter and the election outcome can be associated with participatory theory to the extent that his approach is concerned with electoral regulations causing a divergence between the actual and the eligible median voter (Stephanopoulos 2014, 325). Likewise, one might worry that interpreting *p* in terms of pivotal theory

Other scholars, however, continue to endorse the traditional individual rights and equal protection analysis (Hasen 2003, 139), and particularly for deciding cases of vote denial (Fishkin 2011). In his extensive treatment of the theoretical foundations of voting rights, Joseph Fishkin sets forth requirements for deciding these cases:

This enterprise demands a more carefully specified account of why, and in what way, the individual right to vote matters.... To distinguish more severe burdens from less severe ones, and to develop precedents for deciding when such burdens outweigh state interests, courts will need to make at least implicit use of a theory of vote denial: a theory that tells us in what way(s) disenfranchisement harms individuals (Fishkin 2011, 1332).

Fishkin associates rational choice theory generally with the structuralist approach, pointing out the challenge pivotal theory poses to an individual rights analysis (1333). The practical significance of this point is well illustrated by the appellate court decision of Judge Richard Posner in the voter ID case ultimately decided by the Supreme Court, *Crawford v. Marion County Election Board*.⁴⁰ In finding that the State of Indiana's photo-identification requirement did not impose an impermissible burden on the constitutional right to vote, Judge Posner writes, "The benefits of voting to an individual voter are elusive," parenthetically adding that "a vote in a political election rarely has any *instrumental* value, since elections for political office at the state or federal level are

affects turnout decisions in a way that causes such divergence. Stephanopoulos indicates that participation is a structural value (297); however, it can also be relevant to an individual rights approach, for as Farber (2004) indicates, structural and individual rights concerns are often implicated simultaneously. Note that the pivotal situation associated with the vote choice of the median voter, which emerges from the arrangement of individual preferences in relation to candidate positions, is different from the pivotal situation in the turnout decision. In the turnout decision, as indicated, no one individual can ever be solely pivotal, but in the context of vote choice, the one pivotal voter with median preferences can be precisely identified, at least in theory.

⁴⁰ 553 U.S. 181 (2008).

never decided by just one vote.” Judge Posner then proceeds with a structural analysis of the case, reasoning that “the fewer the people harmed by a law, the less total harm there is to balance against whatever benefits the law might confer.”⁴¹ Fishkin criticizes this structuralist reasoning, maintaining that the right to vote should be enforced more strictly at the individual level in cases alleging vote denial. He thus articulates a detailed normative argument—based in the historical development of conceptions of equal citizenship and a resulting “dignitary harm” in disenfranchisement—that a voting regulation could be unconstitutional even if it unduly burdens only *one* potential voter (Fishkin 2011, 1296, 1357).

Fishkin makes a forceful argument, but the logic of efficacious set causation provides an additional—and arguably more fundamental—basis for valuing and protecting the individual interest in casting a vote. Moreover, this new interpretation of *p* shows that the rational choice approach need not lead inexorably to a purely structuralist theory of voting rights.⁴² The strengthening of participatory theory that emerges from the rejection of pivotal theory provides a straightforward justification for stricter scrutiny of regulations that potentially infringe on the right to vote at the individual level. Individual votes can and do have instrumental effects on election outcomes, and they should be deserving of protection on this basis alone. This does not mean that instrumental

⁴¹ *Crawford v. Marion County Election Bd.* 472 F. 3d 949, 951-952 (7th Cir. 2007). In affirming Judge Posner’s decision, the Supreme Court in *Crawford* (2008) utilized mainly the traditional individual rights versus state interests balancing test, although the majority opinion also invoked elements of structural analysis by questioning the overall number of individuals who would be affected by Indiana’s voter ID law (see Fishkin 2011, 1330). Judge Posner has since appeared to repudiate the result of his 2007 opinion upholding Indiana’s ID law (see Schwartz 2013), but there is nothing indicating his disavowal of the structuralist approach or the principles of pivotal theory (see Bauer 2013; also see Judge Posner’s dissent from denial of *en banc* rehearing in *Frank v. Walker* WL 5326463 (7th Cir. 2014)).

⁴² Fishkin cites Judge Posner’s dismissive attitude toward voting costs—holding that potential voters who fail to clear administrative hurdles in the voting process simply “disenfranchise themselves”—as typical of the rational choice voting model (Fishkin 2011, 1337). However, it may in fact be unnecessary to assume that a rational choice approach inevitably supports what Fishkin calls the “formal” conception of the right to vote expressed by Judge Posner, as opposed to the more “substantive” conception advocated by Fishkin.

motivations are sufficient to fully embody the individual interest in voting, as expressive motivations might conceivably predominate, even in highly competitive elections. However, there is no reason to assume that instrumental rationality is wholly antagonistic to the participatory approach of individual rights protection, so it seems less necessary to resort to Fishkin's conception of dignitary harm to justify more rigorous protection of the right to vote at the individual level. Voters need no longer be viewed as "little more than worker bees" (Gardner 2010, 462), but can instead claim their fair share of democratic sovereignty as bearers of authentic agency and influence over the outcomes of elections. It accordingly becomes more difficult to maintain the view that voting rights do not merit robust protection at the individual level.

The second area of law and policy relevance relates to the competitive school of democracy. Competitive theory is often identified with the minimalist approach of Joseph Schumpeter, who famously held that elections are merely officially sanctioned competitions for power and authority, and democracy is essentially just a non-violent mechanism for resolving the practical problem of who should govern (Schumpeter [1942] 2003). However, Schumpeter's minimalism is given added substance in the work of more recent competitive theorists like Ian Shapiro, who holds up Schumpeterian competition as a model for structuring political institutions in a way that limits forms of domination—hardly a minimalist enterprise (see Shapiro 2003, 51).⁴³

The new interpretation of *p* is naturally in harmony with competitive theory insofar as individual votes become more likely to form part of an efficacious set as

⁴³ It is often assumed that competitive theory is fundamentally inconsistent with participatory theory. This assumption is based on the work of early competitive theorists like Schumpeter, who clearly seemed quite cynical about the value of mass participation (see Fishkin 2009, 69; see also McCafferey et al. 2004a, 9–11). However, under an updated interpretation of competitive theory—like Shapiro's—these two approaches to democracy are not necessarily inconsistent, and a joining of the participatory and competitive strands of theory may in fact be perfectly coherent.

elections become more competitive. When potential voters perceive a race as clearly uncompetitive, although many voters on the expected winning side could still have instrumental reasons to participate, those on the prospective losing side might see themselves as effectively disenfranchised, with their votes again becoming as instrumentally useless as under pivotal theory.⁴⁴ This, to an extent, is perhaps unavoidable in a majoritarian electoral system; nevertheless, the new interpretation of *p* draws attention to reforms designed to ensure that more elections feature true competition, so greater numbers of prospective voters face a high *ex ante* likelihood of having an instrumental effect on the outcome. Empirical evidence is largely consistent with the prediction that individuals are more likely to vote when elections are perceived as competitive (Blais 2000, 60; Blais 2006, 119).⁴⁵ The value of competition, however, involves more than just voter turnout at the individual or aggregate levels; it entails a *structural* concern with democracy. This is the central point of Issacharoff and Pildes (1998), who argue that representational democracy requires robust competition, and courts should intervene when necessary to “destabilize partisan lockups” of the

⁴⁴ Justin Buchler contests the value of competition in democratic elections, and one of his primary arguments is that closer elections yield less “representative” government, since they result in larger numbers of individuals in the minority who voted on the losing side of the election (Buchler 2011, 26). However, the situation of voters who end up on the losing side of a fairly close election does not seem akin to disenfranchisement, since these voters had a chance at winning, but simply failed to reach a majority (or plurality). On the other hand, when *ex ante* it’s clear that one’s preference has no hope at all of winning, it arguably is more problematic from a perspective of representational theory, because it appears to be—if not an actual disenfranchisement—at least a clear case of disempowerment. The theoretical value of competition under majoritarian (and primarily two-party) democracy thus seems tied to how it ensures a realistic chance that one’s preference might prevail, which of course provides the chance to gain power at some point over the longer term, and not remain a permanent minority.

⁴⁵ Blais’s 2006 study is based on a meta-analysis of 32 previous studies, 27 of which found a causal relationship between closeness and turnout. Blais comments on the relationship as follows: “This is the most firmly established result in the literature. I cannot see how this finding could be wrong” (Blais 2006, 119). But see Cann and Cole (2011), finding that the increase in turnout in competitive states during U.S. presidential elections is not a direct result of competition, but rather reflects mobilization effects. Cann and Cole attribute the apparent lack of a direct turnout effect to the workings of pivotal theory (351), but it might instead point to the predominance of non-instrumental motivations for voting.

democratic process. The structuralist approach of competitive theory is thus particularly appropriate for analyzing cases that allege unfair partisan gerrymanders of legislative districts.⁴⁶

This is another area where scholars see a need for greater clarity on foundational democratic theory in order to better guide law and policy (see e.g. Ortiz 2004). The Supreme Court has held that allegations of partisan gerrymanders are justiciable (*Davis v. Bandemer*),⁴⁷ and that excessive partisanship in drawing congressional districts can be unconstitutional (*Vieth v. Jubelirer*),⁴⁸ but the Court has been unable, or unwilling, to formulate a standard for when a particular districting scheme should be struck down (see Berman 2005).⁴⁹ Issacharoff has taken the lead in employing competitive theory to argue that partisan gerrymandering should be prohibited, and he advocates a bright line rule that would mandate institutionalized non-partisanship in the drawing of districts (Issacharoff 2002). In his words, the harm of gerrymandering is that it results in a “constriction of the competitive processes by which voters can express choice” (600). Similarly, Pildes writes of how “the constitutional violation [of partisan gerrymandering] lies in the structural harm to representative self-government” (Pildes 2006a, 271).⁵⁰ The essential principle

⁴⁶ Note that even Joseph Fishkin endorses the use of structural analysis for deciding cases where the aggregate effects of electoral law and policy transcend individual interests, like cases of gerrymandering and vote dilution (Fishkin 2011, 1305; Fishkin 2012, 1893; see also Charles 2007, 651).

⁴⁷ 478 U.S. 109 (1986).

⁴⁸ 541 U.S. 267 (2004).

⁴⁹ Although the Court has criticized partisan gerrymandering for violating principles of fairness and equality, it does not appear to have specifically embraced a democratic value of competition (see Stephanopoulos 2013, 677).

⁵⁰ Nathaniel Persily argues that a concern for competition seems more fitting to an individual rights than a structuralist approach, since lack of competition reduces the chance that an individual might cast a “tie-breaking vote that will decide the election” (Persily 2002, 678 n. 98). This is conceptually problematic, since the chance of casting a pivotal vote in a very large election is essentially zero even with perfect competition. In any event, the theory of efficacious set causation arguably reflects a more structural perspective on competition, since individuals look to the chance that their one vote will be aggregated into a larger set of winning votes, as opposed to calculating whether their vote might prove individually decisive.

states that if citizens are to fully exercise their democratic authority as collective sovereigns, they should have some realistic chance at electing their preferred representatives, which (at minimum) implies the existence of competitive elections. Interpreting p in terms of efficacious set causation, which allows levels of competition to produce meaningful variation in the instrumental motivation for voting, thus helps reinforce opposition to partisan gerrymanders from a structuralist perspective.

This is not to argue that increasing competition should be the only—or even necessarily the overriding—factor in drawing districts, but it strongly supports the notion that competitive theory is worthy of increased doctrinal emphasis. The problem of gerrymandering is complex and contentious, both normatively and empirically. Nevertheless, the theory of efficacious set causation provides encouragement to arguments for judicial intervention to ensure that districts are not intentionally and systematically drawn to be clearly uncompetitive. More substantively perhaps, this new understanding of how individual votes contribute to electoral outcomes might support arguments for laws requiring districting schemes that aim to intentionally increase competition, at least to the extent consistent with other representational values. Moreover, the proposed model for interpreting p implies that districts need not be drawn that highly competitively in order to provide an instrumental motivation for voting, but rather just enough to ensure that prospective voters on the more likely losing side perceive a “realistic” possibility of prevailing.⁵¹

⁵¹ The competitive aspect of this new interpretation of p could also have implications for thinking about U.S. presidential elections. While the Electoral College leads to competition being concentrated in only a few “swing” states, a national popular vote would distribute competition more widely, at least when the race is relatively close at the national level. Like partisan gerrymandering, however, the Electoral College raises complex issues beyond the current scope, and while it may be problematic from the perspective of competition, it may (or may not) serve other important purposes.

7) CONCLUSION

This chapter has described and developed an arguably more plausible and normatively superior alternative to the conventional interpretation of p in the rational choice calculus of voter turnout. The theory of efficacious set causation, based on the pioneering work of Richard Tuck, provides an instrumentally rational justification for individuals to participate in large elections, credibly resolving the so-called paradox of voter turnout. Rejecting pivotal voting theory and restoring the potential for instrumental value in voting has been shown to have major normative implications for the rationality of participating in elections and collective action efforts more generally. The new interpretation of p also provides support to arguments from participatory democratic theory calling for stricter judicial scrutiny of procedural burdens on the individual right to vote, and it can contribute as well to a proposed jurisprudence of partisan gerrymandering based in competitive democratic theory.

Nonetheless, this chapter potentially just scratches the surface of the deeper significance of this transformational approach to voting and other collective action situations. This work is intended to form part of the broader ongoing critique of rational choice theory, even though this work employs that theory's own terms in the turnout calculus, and it aims at redefining the meaning of rational choice rather than supplanting the theory entirely. Hopefully, this might contribute to progress in the wider debate on the nature of rationality while at the same time helping to define a better conceptual framework for the motivating factors of voter turnout. As stated previously, the argument here is not meant to imply that turnout is ever fully explained by instrumental motivations. Individuals may also derive significant expressive, or otherwise non-instrumental benefits from participating in elections, as represented by the D term of the calculus, and furthermore, p may often still be equal to zero even under this new

interpretation. The objective here has been merely to rescue instrumental motivations from the theoretical abyss created by the counterfactual causal logic of pivotal voting theory. Much more analysis is needed to parse the distinctions between instrumental and expressive motivations in theory and in practice, as discussed in the following chapters.

The argument of this chapter could provide a provocation to empirical research, as the theoretical model described above leads to specific hypotheses about voting behavior that could be tested to confirm or deny the theory's validity. First is the basic question of whether individuals ordinarily (or ever) think in terms of efficacious set causation when (or if) they consider the chance that their vote will affect an election outcome, or whether they more commonly think in terms of pivotal theory. Furthermore, if individuals do think in terms of efficacious set causation, they may or may not assign greater instrumental value to being on the winning side of an election, as predicted by Tuck's theory of bandwagon incentives.⁵² In addition, pivotal theory predicts an "underdog effect" due to the strategic context of a majoritarian election, with higher turnout among the prospective minority than those expecting to be in the majority (see e.g. Levine and Palfrey 2007; Morton and Tyran 2012).⁵³ In contrast, the theory of efficacious set causation predicts that voters on the prospective winning side will overall have more motivation to participate, although in close elections the incentive could be

⁵² For a possible confirmation of Tuck's theory, see Howell and Justwan (2013), who unexpectedly find that among those on the losing side after an election, satisfaction with the democratic system is unaffected by the electoral margin, while satisfaction among those on the winning side decreases for wider margins. This is perhaps consistent with the logic of efficacious set causation and bandwagon effects, since on the losing side, where the *ex post* probability of having been in the efficacious set is always zero, the margin of victory should be irrelevant, while on the winning side, where that probability is almost 1 for very close elections and then decreases gradually, margins of victory should matter more.

⁵³ Given costly voting, in equilibrium prospective voters on the expected winning side reduce their likelihood of participation, while those on the expected losing side increase their likelihood (see Morton and Tyran 2012, 10). Mackie indicates that his theory of the mandate value of voting also predicts that voters on the prospective losing side will be more likely to participate (Mackie 2014, 66); however, it is not clear why voters on the prospective winning side might not be equally motivated to increase the perceived mandate of their party or candidate.

somewhat higher for those deemed (slightly) more likely to lose, given that p equals one for those prospective voters. This also raises questions about the existence, and typical locations, of the hypothesized discontinuities in the proposed model for calculating the value of p , as well as more basic questions about how different values of p might be perceived and acted upon in practice.

Beyond these empirical issues, however, the normative and policy implications of this chapter are important on their own. Moreover, as noted in the introduction to this chapter, the interpretation of p affects how the other elements of the voting calculus are interpreted, and this new understanding of how individual votes contribute to election outcomes thus paves the way for a wider research agenda that better accounts for the varying individual and institutional factors that influence voter turnout. Fundamentally, the conception of the role of an individual voter in a large election is an essential element of democratic meaning under modern conditions of mass representation. This work aims to defend the much-maligned value of electoral participation under these conditions, and to thereby assist in development of the critical links between the theories and practices of democratic elections.