Explaining Self-Interested Behavior of Public-Spirited Policy Makers

Abstract: Public choice theory (PCT) has had a powerful influence on political science and, to a lesser extent, public administration. Based on the premise that public officials are rational maximizers of their own utility, PCT has a quite successful record of correctly predicting governmental decisions and policies. This success is puzzling in light of behavioral findings showing that officials do not necessarily seek to maximize their own utility. Drawing on recent advances in behavioral ethics, this article offers a new behavioral foundation for PCT’s predictions by delineating the psychological processes that lead well-intentioned people to violate moral and social norms. It reviews the relevant findings of behavioral ethics, analyzes their theoretical and policy implications for officials’ decision making, and sets an agenda for future research.

Evidence for Practice

• Ethical behavior is driven both by self-interest and by norms. Thus, officeholders tend to breach norms only to the extent that they can maintain their self-image as honest people.
• Public officeholders may act in self-interested ways because of automatic and unconscious motivations rather than deliberate and conscious calculations.
• The clearer it is to an officeholder that his or her interests diverge from those of the public at large, the less likely he or she is to give precedence to the former.
• Officeholders can recollect decisions they took that clearly promoted the public interest more easily than their self-interested decisions.

In recent decades, both sets of assumptions have been challenged by a growing body of experimental and observational studies by psychologists, economists, and political scientists that cast doubt on the premise of cognitive rationality by showing that people’s preferences and choices often fail to meet requirements such as dominance, transitivity, and invariance. The heuristics-and-biases literature has demonstrated that people’s choices deviate from these assumptions in systematic and predictable ways (Baron 2008; Keren and Wu 2015). Based on this literature, it has been argued that many suboptimal governmental decisions are attributable not to the economic rationality of public officials but to their bounded rationality (Lucas and Tasic 2015; Rachliski and Farina 2002; Simon 1972). Other studies have called into question the premise of motivational rationality by pointing to the role of other motivations, such as envy, altruism, and concerns about fairness and reciprocity (for an overview, see Gächter 2014). Accordingly, many suggestions have been made to qualify, supplement, and improve PCT by introducing more nuanced assumptions about policy makers’ motivation and cognitive ability (e.g., Jones 2001; Ostrom 1998).
Notwithstanding these and other compelling critiques, PCT has often provided quite accurate predictions of governmental policy choices. Many governmental decisions appear to systematically cater to the interests of officeholders, or to those of powerful interest groups, rather than the interests of the general public (e.g., Bartels 2008; Gilens and Page 2014; Jacobs and Page 2005; Schlozman, Verba, and Brady 2012). In line with PCT, such decisions advance the interests of those who can directly and indirectly benefit (or harm) public officials by providing (or withholding) campaign finance for future reelection, facilitate (or block) access to privately controlled media, offer (or deny) future jobs in the private sector, and the like.

In this article, we seek to bridge the gap between PCT’s predictive power and its questionable behavioral assumptions by integrating insights from the recent and growing study of behavioral ethics (BE) into the theory of political and administrative behavior. Broadly speaking, BE identifies the circumstantial, personal, and environmental conditions that lead ordinary people—through automatic, intuitive, and mostly unconscious psychological processes—to violate moral and social norms (for overviews, see Bazerman and Tenbrunsel 2011; Feldman 2014). These insights can explain when and how well-intentioned officials might make decisions that advance their own interests and those of rent seekers rather than the public interest. The fact that the meaning and implications of the norm that government authorities should pursue the public good are often ambiguous renders such violations particularly likely. BE studies lend support to the argument that public officials often advance their personal interests—but without necessarily portraying them as cynical. Thus, we argue that in many instances, the insights of BE provide a sounder behavioral footing for PCT’s predictions.

Our thesis is associated with social psychology and the emerging field of behavioral public administration (Grimmelikhuijsen et al. 2017). Yet, unlike studies that focus on citizens’ decision making (e.g., Caplan 2001; for an overview, see Schnellenbach and Schubert 2015, 397–404), here we discuss governmental decision making. Our analysis is part of behavioral economics broadly conceived. However, while most behavioral analyses of governmental decision making have hitherto focused on cognitive rationality (drawing on the heuristics-and-biases literature), we focus on BE—in particular, conflicts between self-interest and moral and social norms (we are unaware of previous integration of the BE literature into political psychology or behavioral public administration). Most significantly, whereas psychological insights are usually used to challenge and qualify PCT, our thesis aims to reinforce PCT’s predictions.

In addition to strengthening the behavioral foundation of PCT’s predictions, our theory offers testable hypotheses and policy implications that sometimes diverge from those offered by standard PCT or the heuristics-and-biases literature. For example, according to standard PCT, the more blatant a public official’s conflict of interest, the greater the danger that he or she will put his or her own interests before the general good. In contrast, BE predicts that when a conflict of interest is clear and unmistakable, officials are more likely to recognize and control their automatic tendency to advance their own interests. Thus, it is the less obvious cases of conflict of interest that pose a greater threat to a well-functioning public administration. Importantly, we do not argue that public officials are always well intended, that governmental failures are never attributable to cognitive biases, or that PCT’s predictions are always correct. Clearly, the picture is more complex: officials sometimes deliberately pursue their own interests at the expense of the public good, policy makers sometimes fall prey to cognitive biases, and governmental decisions sometimes do effectively promote the general good. Our point is that much of the success of PCT in predicting governmental decisions is attributable not to decision makers deliberately maximizing their own utility but rather to automatic and mostly unconscious psychological processes that increase the prevalence of self-interested behavior—often beyond the awareness and recollection of the acting public agents. This may account for the existing disjunction between findings that suggest that public officials possess a great deal of prosocial motivation and the considerable success of PCT in predicting policy outcomes.

The remainder of this article consists of four sections. The following section offers an overview of public choice theory and its critique and concludes with a puzzle: given the doubtful behavioral assumptions of PCT, why are its predictions often correct? To answer this question, the next section surveys the relevant findings and insights of BE and argues that they can provide a sounder behavioral basis for PCT’s predictions. The following sections then analyze the theoretical and policy implications of our thesis and suggest directions for further research.

Public Choice Theory and Its Critique

This section succinctly reviews the fundamental elements of PCT that are relevant to our thesis. This review is followed by a critique of PCT and concludes with the main puzzle that motivates this article: given its doubtful behavioral assumptions, why are the predictions of PCT often correct?

Public Choice Theory: General

In keeping with standard economic analysis, PCT assumes that political and administrative officials rationally maximize their own interests (Mueller 2003, 1). The premise of self-interest rules out direct concern for the welfare of others. This assumption allows that officials might consider the welfare of other people or concerns of fairness only to the extent that these affect their self-interest (Frohlich and Oppenheimer 2004). In the political domain, this assumption entails organized interest groups inducing self-interested government officials to pursue policies that advance the former's interests in return for benefits.

The pivotal role of self-interest in the development of this understanding is reflected in the revisionist theory put forward by Anthony Downs and its more recent iterations. The logic of collective action means that relatively small and homogenous groups are much more effective than the public at large at promoting their narrow interests. On the assumption that public officials, both elected and appointed, aim to maximize their own utility rather than promote the public interest, Downs further deduced that public officials cultivate coalitions of support of interest groups by adopting policies that favor these groups in return for their support (Downs 1957). These assumptions have been supported by subsequent studies of regulatory politics (Becker 1983; Peltzman 1976; Stigler 1971; Wilson 1980), leading to a broadly shared view
that relatively small groups use regulation to attain favorable results that they cannot attain through the market.

Admittedly, attending to the narrow interests of certain groups does not necessarily hinder the adequate functioning of democratic governance or undermine social welfare (Croley 2010). However, if public officials are predominantly self-interested, the disproportionate power of interest groups to affect the present and future utility of public officials may lead to policy choices that prioritize these narrow interests over the greater good (Levine 1981; McGinnis and Ostrom 2011). The resulting wealth shifts may be neither efficient nor fair (Shapiro and Tomain 2003, 81).

While the prospect of reelection is a key motivation for elected officials, bureaucrats are assumed to enhance their utility by maximizing discretionary budgets (Migué, Bélanger, and Niskanen 1974; Niskanen 1975) and their prospects of future employment in the private sector (Stigler 1971). These goals suggest that officials who wish to secure their position and capacity to intervene in the economy find it in their interest to be responsive to legislators on behalf of their constituents (Fiorina and Noll 1979). Several subsequent studies have provided empirical support for the responsiveness of bureaucrats to the preferences of their elected overseers (Weingast and Moran 1983; Wintrobe 1997; Wood and Waterman 1993).

One key entailment of the premise of self-interest in the administrative and political context is the agency problem. The agency problem was studied extensively in the social sciences during the twentieth century—first in the context of corporations (Berle and Means 1932) and gradually in other social, economic, and political contexts as well (e.g., Ferejohn 1986; McCubbins and Schwartz 1984; Miller 1992, 2005). The principal—agent theory posits that social and political interactions are founded on arrangements of delegation and accountability, whereby one party (the principal) delegates authority to another (the agent), and the welfare of the former is affected by the choices made by the latter (Arrow 1985; Eisenhardt 1989; Scott 1998). Consistent with economic rationality, both actors are assumed to be self-interested maximizers. While economic rationality is efficient in perfectly competitive markets, when coupled with asymmetric information, it leads to suboptimal equilibria. The agent’s self-interest typically results in a divergence of his or her interests from those of the principal. This conflict of interest is further exacerbated by the fact that (1) the agent typically possesses information that the principal does not (hidden information), (2) the agent’s behavior is usually difficult for the principal to monitor (hidden action), and (3) outcome-based monitoring is often hindered by uncertainties about the translation of the agent’s actions into results. For all these reasons, the agency problem is both a persistent feature of social life and inherently hard to solve (Miller and Whitford 2007).

To be sure, PCT may account for prosocial public choices, especially when circumstances limit one’s ability to act in a self-interested manner. For example, the theory of constitutional government (Buchanan and Tullock 1962) posits that when individuals are uncertain about their future positions, they are drawn out of their self-interest to choose rules that take into account the positions of all other individuals. Similar examples can be found in the theory of public goods, the prisoner’s dilemma, and insurable risks, which demonstrate how self-interested individuals can reach collective agreements (Mueller 2003, 603). In principle, economic analysis, including PCT, may also do away with the assumption of motivational rationality, taking all of people’s preferences, including other-regarding ones, as given (Francois 2000). However, more often than not, PCT predicts that officials’ decisions will further their own interests rather than the public good.

The behavioral assumptions underpinning PCT are indeed parsimonious and foster the development of models. However, numerous scientific works in psychology and behavioral economics have questioned their empirical validity. The primary critiques of PCT center on its assumptions that officials are perfectly rational and invariably self-interested. Let us briefly review these two critiques.

**Criticism of the Behavioral Assumptions of Public Choice Theory**

It has been argued that the behavioral assumptions of PCT—specifically, that public officials are cognitively rational and solely motivated by self-interest—are overly simplistic. Critics of the premise of cognitive rationality argue that suboptimal choices and policies are often the result of bounded rationality (Jones 2001), cognitive limitations, and heuristics and biases (Cooper and Kovacic 2012; Rachlinski and Farina 2002; see also Bendor 2015; Berggren 2012; Schnellenbach and Schubert 2015; Viscusi and Gayer 2015). Nothing in our analysis detracts from these cogent observations. However, our focus is on the critique of the second premise, namely, studies that demonstrate the prevalence of nonselfish motivations.

Experimental studies have shown that people often deviate from strictly self-interested choices and are willing to forgo personal gain for the sake of fairness (Fehr and Fischbacher 2003; Gächter 2014). Indeed, unselfish motivations are no longer excluded from leading theories of economic behavior (Bolton and Ockenfels 2000; Fehr and Schmidt 1999). In the sphere of public administration, numerous studies have shown that policy makers are often motivated by a desire to promote public and social interests. Public service motivation (PSM) theory, for example, suggests that self-selection in the processes of joining the public sector results in organizations whose typical staff are disproportionately motivated by prosocial concerns (Georgellis, Iossa, and Tabvuma 2011; Perry 1996), which, in turn, mitigates some aspects of the agency problem (Gailmard 2010; Perry, Hondeghem, and Wise 2010; Perry and Wise 1990). These findings regarding the importance of intrinsic prosocial motivation rely on survey-reported values and attitudes as a proxy for such motives (Georgellis, Iossa, and Tabvuma 2011). These include a self-reported attraction to policy making, a commitment to serving the public interest, compassion and self-sacrifice (Perry 1996), and the desire to work “in a job that allows one to help other people” and “that is useful to society” (Crewson 1997, 502; see also Feeny 2008, 489; Georgellis, Iossa, and Tabvuma 2011, 480; Lewis and Frank 2002; Tschirhart et al. 2008). Other organizational theories suggest that a sense of common identity and commitment, as well as socialization, are instrumental in enhancing public servant’s prosocial behavior (Carpenter and Krause 2015; Dilulio 1994; Robertson, Wang, and Trivisvavet 2007). This is particularly true when social interaction and communication form part of the precontractual negotiation (Bottom et al. 2006; Charness and Dufwenberg 2006).
The Puzzle: Problematic Behavioral Assumptions and High Predictive Power

Given its debatable behavioral premises—particularly with regard to the centrality of self-interest—one might expect PCT to regularly fail in predicting governmental decisions and policies. In reality, however, PCT often appears to do this quite successfully. The empirical results of decades of studies indicate that a broad range of decisions by elected and nonelected public officials appear to be motivated by self-interest. Specifically, various policy decisions by elected officials have been found to be guided by reelection considerations. Notable examples include U.S. federal spending patterns across states (Wright 1974), the use of U.S. presidential veto power (Grier, McDonald, and Tollison 1995), decisions by British prime ministers with regard to elections timing (Smith 2003), and decisions by U.K. cabinet ministers on whether to appoint a public inquiry into a given crisis (Sulitzeanu-Kenan 2010). While reelection considerations may align with the general public interest, this is obviously not always the case.

Other successful predictions of PCT include patterns of regulation (Carroll 1983; Peltzman 1976; Stigler 1971), the relationship between campaign contributions and (lenient) regulatory policy (De Figueiredo and Edwards 2007), and findings that suggest that even the U.S. Food and Drug Administration—one of the most professional and reputable public organizations—is influenced by organized interests and the media in its drug approval decisions and enforcement activities (Carpenter 2002; Maor and Sulitzeanu-Kenan 2013). More broadly, several studies provide consistent evidence that both domestic and foreign policies reflect the preferences of elites and organized interests rather than the average voter or general public (Bartels 2008; Gilens and Page 2014; Jacobs and Page 2005; Schlozman, Verba, and Brady 2012).

To be sure, the predictions of PCT are not invariably correct, and some phenomena are better explained by PSM and other competing theories—or some combination thereof. Nevertheless, the intriguing disjunction between the predictive power of the self-interest premise and its most behavioral foundations calls for further investigation. The following section offers a possible resolution to this puzzle.

Public Choice Theory Meets Behavioral Ethics

Recent advances in the study of moral behavior delineate several consistent findings that enable us to propose a complementary behaviorally based theory for the choices made by public officials. In this section, we provide a bird’s-eye exposition of BE, followed by an explanation of how its insights complement PCT. Note that BE studies strive to account for people’s unethical behavior given their ethical convictions. These studies do not provide a normative assessment of those convictions or a normative critique of PCT or rational choice theory more generally. Much like PCT, the study of behavioral ethics is primarily descriptive and explanatory in nature rather than normative.

Behavioral Ethics

General. This subsection presents the main insights of BE, with particular emphasis on those with direct bearing on governmental decision making. (For overviews of BE, see Bazerman and Gino 2012; Bazerman and Tenbrunsel 2011; Bellé and Cantarelli 2017; Bereby-Meyer and Shalvi 2015; Feldman 2014; Kish-Gephart, Harrison, and Treviño 2010.) It should be stressed at the outset that BE does not offer a simple and coherent theory of human behavior of the sort put forward by standard economic analysis. As Daniel Kahneman points out, “life is more complex for behavioral economists than for true believers in human rationality” (2011, 412). There are, however, basic concepts that underpin much of BE—chief among which are dual reasoning and motivated reasoning.

Dual reasoning. As previously noted, recent decades have witnessed a wealth of experimental and observational studies finding that people’s choices and decisions routinely deviate from models of rational decision making. One major explanation for these deviations is that people use two modes of thinking—commonly known as System 1 and System 2 (Evans 2008; Kahneman 2011; Stanovich and West 2000). System 1 thinking is intuitive and holistic, context dependent, and implicit. It is based on heuristics that people acquire through personal experience, as well as innate ones, and it serves us well in most daily tasks. In contrast, System 2 is conscious, deliberative, and analytical. It employs rules that are explicitly learned. While System 1 is automatic, quick, and relatively undemanding of cognitive capacities, System 2 is controlled, slow, and exacting. The effortless, speedy, and autonomous nature of System 1 makes it dominant a priori—that is, it controls behavior by default, unless analytical reasoning intervenes (Evans 2006; Stanovich 2011, 19–22). While System 2 may intervene and regulate System 1 when the latter leads to suboptimal results, people often stick to the intuitive choices of System 1 and use System 2 mainly to provide justifications for those choices (Thompson 2009; Trouche et al. 2015). Dual reasoning has been adopted in political science (Lodge and Taber 2013), although mostly in the subfields of political psychology and public opinion rather than in studies of elite behavior and public administration.

The concept of dual reasoning plays a key role not only in studies of judgment and decision making but also in BE research. Many BE studies show how automatic processes facilitate the promotion of people’s interests and goals. Unlike standard economic analysis, which posits that people deliberately maximize their utility, BE focuses on automatic cognitive processes.

Motivated reasoning. The second basic mechanism that bears on people’s ethical behavior is motivated reasoning. In her review of a host of previous studies, Ziva Kunda (1990) argues that motivation affects reasoning through the cognitive processes by which people form impressions, determine their beliefs and attitudes, assess evidence, and make decisions. While people are sometimes motivated to reach an accurate conclusion, they often aim to reach a particular (“directional”) conclusion—usually the one that best serves their interests. When motivated by accuracy, people tend to use the most appropriate strategies to attain that goal. In contrast, directional motivations prompt people to use strategies that are likely to yield the desired conclusion. Interestingly, directional processing of information can be as detailed and thorough as accuracy-motivated processing. Thus, information processing can be simultaneously thorough and biased (Kunda 1990).

Motivated reasoning affects not only the decision process but also ex post recollection. People misremember both what they have done
and what they were told to do when that allows them to believe that they have acted ethically. In some experiments, participants who were given an opportunity to cheat tended to forget the precepts of an honor code they had previously read, far more than participants who were not given such an opportunity (Kouchaki and Gino 2016; Shu, Gino, and Bazerman 2011).

**Automaticity of self-interest.** Considerable evidence supports the claim that self-interest affects ethical behavior through System 1 thinking. Self-interest is “automatic, viscerally compelling, and often unconscious,” while compliance with professional obligations is “a more thoughtful process” (Moore and Loewenstein 2004, 189; see also Epley and Caruso 2004; Hughes and Zaki 2015). The automatic nature of self-interest makes it difficult for people to be aware of it—hence, they are unlikely to counteract its effect on their reasoning. For example, in one set of experiments, subjects were first asked to make estimates on behalf of one party (e.g., a prospective buyer or seller) and then incentivized to make estimates as objectively as possible (Moore, Tanlu, and Bazerman 2010). It was found that affiliation with one party not only biased the original estimates, as predicted by the notion of motivated reasoning, but that this bias carried over to the subsequent estimate—despite monetary incentives to being objective and accurate (see also Engel and Glöckner 2013). Furthermore, subjects tended to significantly underestimate their own bias. It appears that subjects actually believed their biased assessments. The notion that unethical, self-interested behavior is automatic (and only partly curtailed by self-control) is also supported by the finding that time pressure increases the incidence of self-serving unethical behavior, while ample time reduces it (unless one is able to come up with justifications for one’s actions) (Bereby-Meyer and Shalvi 2015; Shalvi, Eldar, and Bereby-Meyer 2012). Finally, a recent study showed that when subjects were experimentally manipulated into an intuitive/automatic mind-set, they tended to act in a more self-interested manner than when they were manipulated into an analytical/deliberative mind-set (Feldman and Halali 2017).

While the majority view in the literature emphasizes the role of System 1 in unethical behavior, the picture is more nuanced: System 1 thinking does not always lead to selfish behavior, and people do, sometimes, deliberately and consciously violate moral and social norms (Bereby-Meyer and Shalvi 2015; Feldman 2014; Hughes and Zaki 2015). In circumstances of explicit social exchange, such as economic games, cooperation and reciprocity—rather than self-interested deflection—appear to be the automatic response (Halali, Bereby-Meyer, and Meiran 2014; Rand, Greene, and Nowak 2012).

**Factors and mechanisms.** A common theme of BE studies is that ordinary “good people” sometimes do “bad things” (Bazerman, Loewenstein, and Moore 2002; Bersoff 1999; Mazar, Amir, and Ariely 2008). People tend to display moral hypocrisy—that is, they are motivated “to appear moral in one’s own and other’s eyes while, if possible, avoiding the cost of actually being moral” (Batson et al. 1999, 525). While rational choice theory might predict that people totally disregard ethical norms, in fact, they tend to breach those norms only to the extent that they can maintain their self-image as honest people (Bazerman and Gino 2012). As Nina Mazar and her colleagues put it, “people behave dishonestly enough to profit but honestly enough to delude themselves of their own integrity. A little bit of dishonesty gives a taste of profit without spoiling a positive self-view” (Mazar, Amir, and Ariely 2008, 633).

A telling demonstration of this observation is found in an experiment in which all subjects were “mistakenly” overpaid for their participation, and the conspicuousness of the unethicality of not correcting this mistake was treated by varying the identity of the “victim” of this behavior (an overseas firm that had financed the experiment or the experimenter) and the extent to which subjects were indirectly induced to deliberate on ethical issues (Bersoff 1999). The more difficult it was made for subjects to ignore the unethicality of keeping the excess payment and its adverse effect on a specific person, the more they tended to point out the overpayment. Increasing the saliency of dishonesty has been found to reduce cheating in other experiments as well (Gino, Ayal, and Ariely 2009). Similarly, it was found that specific rules reduce unethical decisions more than general ones because the former curtail people’s ability to rationalize their behavior (Mulder, Jordan, and Rink 2015). Finally, when people are faced with a self-benefiting choice that might harm someone else, they prefer not to know whether such harm would indeed ensue, so that they can make that choice in good conscience (Dana, Weber, and Kuang 2007; see also Mazar, Amir, and Ariely 2008; Shalvi et al. 2011).

People use several means of self-deception to avoid recognizing the unethicality of their behavior (Tenbrunsel and Messick 2004). These include the use of euphemisms (such as “creative accounting” or claiming to act “rationally” rather than “egoistically”), ethical numbing (when a morally dubious behavior is repeated), and putting the blame on others. Indeed, under certain circumstances, these and comparable mechanisms of moral disengagement may lead not only to lying and cheating but also to the perpetration of large-scale atrocities (Bandura 1999). Other mechanisms include moral justification (i.e., rationalizing the immoral behavior on the grounds that it serves an important social or moral purpose), advantageous comparison (contrast the behavior in question with even more reprehensible conduct), and distortion of consequences—in particular, minimizing the seriousness of the adverse effects of one’s behavior (Bandura 1999; Moore et al. 2012; see also Ayal and Gino 2012). Unethical behavior is also deemed more justified when it benefits not only oneself but others as well, thereby allowing one to present it as altruistic (Gino, Ayal, and Ariely 2013).

While some studies have shown that people care not only about fulfilling their interests but about fairness (Kahneman, Knetsch, and Thaler 1986; Zamir and Ritov 2011), it appears that fairness concerns often do not restrain unethical behavior, as fairness is a highly malleable concept. For example, people make self-serving judgments of fairness in the context of bargaining (Thompson and Loewenstein 1992). Therefore, while portraying people as solely self-interested is overly simplistic, people’s concern for fairness does not guarantee that they will act fairly.

**Social and organizational factors.** To fully understand unethical behavior and its implications for self-interested behavior of public officeholders, the role of social and organizational factors should be considered. Psychologists have long studied the conformity effect—namely, people’s tendency to adapt their behavior to suit
group norms (Forsyth 2012). This tendency pertains to unethical behavior as well as to other contexts. For example, it was found that viewing the unethical behavior of another person increased the unethical behavior of participants when that person was perceived as an in-group member but decreased it when the person was considered an out-group member (Gino, Ayal, and Ariely 2009; see also Ayal and Gino 2012). The closer people feel to someone who sets an example of unethical behavior, the less harshly they judge that behavior, and the more likely they are to engage in such behavior themselves (Gino and Galinsky 2012).

A meta-analysis in the context of organizations found a weak positive correlation between an egoistic work environment and unethical behavior and moderate negative correlations between benevolent and principled ethical climates and unethical choices (Kish-Gephart, Harrison, and Treviño 2010). In a recent meta-analysis, Bellé and Cantarelli (2017) found that social influences such as exposure to unethical behavior by in-group members or interdependence with other individuals benefiting from one’s unethical actions increase unethical behavior. A robust negative correlation was found between an ethical culture—including high ethical standards set by managers and the disciplinary enforcement of an existing code of conduct—and unethical conduct. Another experimental study further substantiated the effect of an unethical business culture in the banking system on the ethical behavior of its employees (Cohn, Fehr, and Maréchal 2014). More fundamentally, it has been demonstrated that cooperation significantly increases unethical behavior, to the extent that “dishonesty estimates observed in past work may have been conservative in comparison with settings in which people work in collaboration” (Weisel and Shalvi 2015, 10653).

The findings that unethical behavior is largely automatic and that many mechanisms exist that allow people to overlook the unethicality of their conduct pose serious challenges for any attempt to discourage such behavior. If unethical behavior is not the product of cost–benefit analysis, then raising its expected costs (for example, through legal sanctions) would not necessarily be effective (Feldman 2010; see also Pronin, Lin, and Ross 2002). These mechanisms allow people to deviate from consciously held values, such as those associated with being a public servant, while maintaining a conscious adherence to those values.

However, the application of BE insights in the public sphere may be challenged. As described in the previous subsection, most BE studies pertain to basic moral prohibitions on lying, cheating, and stealing rather than the duty of public officials to advance the public good. They look at individual and simple decision making by private individuals during a single laboratory session rather than collective, procedurally structured, and repeated decision making by public officials over an extended period of time. All of these differences call for caution when applying insights from BE to governmental decision making and for additional inquiry into its special characteristics—issues we elaborate in the next two sections. We nevertheless believe that such application is warranted—in some respects even a fortiori.

Let us begin with the pertinent norms—the prohibitions on cheating and stealing versus the duty to pursue the public good rather than that of powerful interest groups or one’s own interests. Arguably, promoting one’s self-interest (or those of small interest groups) instead of the public good is different from lying and stealing, while promoting the public good is not an entrenched, basic moral norm, and it is more ambiguous than the prohibition on dishonest behavior, as it is often unclear what constitutes the “public good” and which course of action best serves it.

In response, there is an important common denominator between BE studies and the current context—namely, the role of self-interest vis-à-vis competing norms. This is true whether the competing norm is “Do not lie,” “Do not steal,” or “Do not promote your interests at the expense of the public good.” When a public official is entrusted with promoting the public good, he or her own interests (or those of a particular interest group) instead, such conduct is akin to unfaithfulness and cheating. Just as the prohibition on cheating is entrenched in commonsense morality, the proscription of conflicts of interest is a common, fundamental, and well-known tenet among public officials.

More importantly, the differences between the private and public spheres make the insights of BE all the more relevant in
the latter sphere, as the psychological mechanisms that facilitate unfaithfulness and cheating are especially suited to enabling the self-serving decisions predicted by PCT. Among other things, the fact that the duty to promote the public good and to avoid conflicts of interest is less obvious and less clear than the moral prohibitions on lying and cheating makes violations of the former less conspicuous. In this context, reflecting on the politics of interest groups elicits new insights. Powerful interest groups may be influential because of their ability to gather, analyze, and provide officials with convincing evidence and arguments—and they may also hold the power to harm or benefit those same officials. While standard PCT concentrates on the latter aspect, behavioral ethics draws attention to interest groups’ ability to obscure the boundaries of the “public interest,” thereby enabling officeholders to concede to interest groups’ demands while maintaining a moral self-perception. This insight is in line with recent advances in political science that construe lobbying as a legislative subsidy (Hall and Deardorff 2006). However, it extends the strictly informational benefits that this model entails by showing how such information makes it easier to rationalize questionable policy choices. Consequently, the fact that governmental decision making is often a prolonged process does not necessarily make self-interested decisions less likely, as more time fosters greater rationalization.

As for the argument that most BE studies refer to individual decisions rather than collective ones and do not examine repeated decisions over extended periods of time, it should first be noted that many public officials do make decisions individually. More importantly, the available data indicate that infringements of moral and social norms may be more—rather than less—prevalent in collaborative environments (Weisel and Shalvi 2015) and in repeated settings (because of ethical numbing). Two other characteristics of public officials’ decision making that make norm infringement more likely are the organizational setting of the public sphere, which makes it much easier to cast blame on others (Thompson 1980), and the fact that in the public sphere, the “victims” of biased policies are typically unidentified, faceless people (Bersoff 1999).

Another difference between the private and public spheres, which makes BE potentially more important in the latter, is that in the private sphere—particularly in market settings—there is nothing wrong with pursuing one’s own interest. On the contrary, the beauty of the competitive market is that it facilitates the maximization of social utility when each person pursues his or her own interest (with the exceptions of agency problems, as in the case of corporations or in attorney-client relationships).

Additionally, it has been shown that winning a contest produces a sense of entitlement that induces people to act less ethically (Schurr and Ritov 2016). If this is the case, it stands to reason that elected officials, and other functionaries who owe their positions to winning in an election or in a competitive environment, would be particularly prone to violating norms for their own benefit. Finally, although public officials are typically experts who use professional procedures when making decisions over extended periods of time, several findings suggest that this practice cannot, in and of itself, guarantee that their choices will be ethically sound. First, studies have shown that the degree of myside bias is similar across intelligence and cognitive ability levels (for a review, see Stanovich, West, and Toplak 2013). Second, the use of explicit reasoning (System 2) and enhanced expertise has been found to be associated with higher rates of motivated reasoning (see, respectively, Kahan 2011, 2013; Taber and Lodge 2013).

Before proceeding to examine the theoretical and policy implications of our thesis, it is important to reiterate that our argument is not based on the naive assumption that all public officials are “good people” who fail to do the right thing solely because of automatic and unconscious psychological processes. Some public officials undoubtedly deliberately and consciously breach legal, organizational, and moral norms. The crux of our argument is that even if most public officials are well-intentioned and honest people, they are liable to make socially undesirable, self-serving decisions. Hence, the predictions of PCT need not hinge on the disputable assumption that public officials are strictly rational maximizers of their own utility, who do not care about the general good.

**Theoretical and Policy Implications**

Our proposed thesis entails a number of theoretical and policy implications. First, it provides a stronger, empirically based foundation for PCT’s predictions without relying on simplistic, questionable assumptions about human motivation. Rather than using behavioral insights to challenge the predictions of PCT—as commentators have long done—our thesis strengthens the behavioral basis of PCT’s predictions.

Concomitantly, some of the predictions and policy implications of our thesis clearly diverge from those of standard PCT. The following propositions highlight these divergences:

**Proposition 1:** Holding the potential gains and risks involved in a self-interested decision constant, the more salient a decision maker’s conflict of interest is, the less likely he or she is to make a self-serving decision.

**Proposition 1a:** A conflict of interest is more salient when the personal gain from a decision is larger, it is more tangible and concrete, and the divergence between the private and public interests is greater.

PCT advocates measures that aim to align the interests of public officials with those of the public at large. A perfect alignment would indeed eliminate the problem of conflicts of interest—but this is rarely, if ever, attainable. Partial alignment of interests mitigates the agency problem, but at the expense of rendering the conflict of interest less salient. Alas, diminished saliency renders unethical behavior more likely, as it makes it easier for people to believe that their decisions serve the overall good. Therefore, it is a mixed blessing.

To put it another way: according to standard PCT, the clearer it is to the decision maker that his or her interests diverge from those of the public at large, the more likely he or she is to give precedence to the former (at least insofar as the decision maker believes that he or she could get away with it). BE, however, predicts the opposite: the starker this divergence, the harder it is for the decision maker to use
From a heuristics-and-biases perspective, professional expertise is by no means a panacea against decision failures (e.g., Zamir 2015, 33–36). Nonetheless, it has been argued that professional civil servants are less likely to fall prey to the cognitive biases that affect elected politicians, thanks to the decision strategies that they develop to overcome the perils of intuitive decision making. Hence, the optimal arrangement may be to divide decision making between political, professional, and judicial bodies (Rachlinski and Farina 2002).

The picture becomes more complex once BE insights are taken into account. On the one hand, conflicts of interest may be more prevalent among elected officials than appointed ones. On the other hand, professional decision makers may be more susceptible than laypeople to some of the automatic psychological mechanisms that facilitate norm violations. These include slippery slopes—the psychological numbing that comes from repetition—and the prevalent use of euphemisms (Bazerman and Tenbrunsel 2011, 91–94, 123–24; Tenbrunsel and Messick 2004). Indeed, self-serving biases have been found among professionals such as experienced negotiators (Babcock and Lowenstein 1997) and academic coauthors (Caruso, Epley, and Bazerman 2006). Moreover, it has been demonstrated that professionalism—or, more specifically, seeing oneself as a professional, that is, as one who is technically or morally superior in some way—may actually increase unethical behavior because it licenses one to act unethically (Kouchaki 2012).

From a policy perspective, if simple incentives are less likely to counteract automatic and possibly unconscious process, what else can be done to improve governmental decisions? One measure—which is already in place in many legal systems—is to disqualify public officials from making decisions in cases in which they may be biased (Wade and Forsyth 2014, 384–404). This rule not only invalidates biased decisions but also seeks to eliminate the potential for such decisions. A common justification for the rule is the appearance of propriety: justice must not only be done but also be seen to be done. Indeed, courts and scholars have noted that a major (if not the major) purpose of this rule is not to prevent deliberate abuse of power but rather to prevent nondeliberate abuse (Farina 1993, 296–97; see also Regina v. Inner West London Coroner ex p Dallaglio [1994] 4 All ER 139, 152).

However, the concern that automatic and largely unconscious biases may hinder the promotion of the public good extends far beyond the instances covered by the legal rule; hence additional measures seem necessary. One option that comes to mind is to use mild, ncoercive means—similar to the much-discussed “nudges” in the behavioral economics literature (Thaler and Sunstein 2009). The following proposition may guide such interventions:

**Proposition 3:** Holding potential gains and risks from self-interested behavior constant, self-serving behavior may be reduced by greater awareness of the relevant norms—which, in turn, may be achieved by drawing officials’ attention to them before they make decisions, emphasizing personal responsibility, and inculcating an ethical work environment.

It has been shown that moral disengagement and motivated forgetting of ethical rules can be significantly reduced by simple devices, such as having participants read or sign an honor code.

---

**Figure 1** The Hypothetical Likelihood of Acting Self-Interestedly under PCT and BE Assumptions, Holding the Likelihood of Detection Constant

A host of mechanisms that would shield him or her from recognizing the immorality of her decisions—thus making him or her less likely to act immorally. Suppose, for example, that the term for a contractor’s agreement is about to end, and a public official faces a choice between issuing a new competitive tender and extending the agreement with the current contractor. While the government may get better performance and possibly save money through a competitive tender, the new contractor may not be as good as the previous one, and the tasks involved in handling the tender and changing to a new contractor are likely to entail considerable extra work for the official. Unlike situations in which a public official stands to gain directly and substantially from her decision, here the subtle divergence between public and personal interests may easily be overlooked, and the official may well make a self-interested decision without being aware of her self-servicing bias. In the same vein, from a BE perspective, the risk that public officials might make biased decisions as a result of a monetary bribe is smaller than the risk that they might do so in return for immaterial benefits (either actual or expected), such as political support or peer recognition. This is because immaterial benefits are much less likely to trigger a critical reflection by System 2 thinking and therefore more likely to compromise the moral conduct of “good people” (Feldman 2014).

This claim is graphically depicted in figure 1. The x-axis denotes the perceived divergence of interests between the official and the public at large (whereby 0 = perfect alignment, and 10 = maximum misalignment of interests), and the y-axis depicts the likelihood of the official advancing his or her own interest. While standard PCT predicts a monotonic increasing function, BE predicts a concave function, whereby the likelihood of a self-serving decision initially increases but begins to decrease at a certain point as the conflict of interest becomes more salient.

**Proposition 2:** While expertise and experience may improve governmental decision making because professional civil servants are less likely to rely on simple heuristics, professional civil servants are not necessarily less susceptible to self-serving biases than elected officials.
before engaging in an activity (Shu, Gino, and Bazerman 2011), or by other explicit messages regarding the relevant norm (Feldman and Halali 2017). It has also been experimentally demonstrated that people who might be tempted to be dishonest in a written report are significantly less tempted if they are asked to sign the report before (rather than after) completing it (Shu et al. 2012). It may thus be advisable to structure administrative processes such that explicit declarations of fidelity (e.g., signature, oath, etc.) would precede and be temporally proximate to the provision of information or the making of decisions to which these declarations refer. That said, measures taken to increase the saliency of ethical norms—such as asking people to disclose their conflicts of interest—may also backfire because of moral licensing: people who disclose their conflicts of interest may feel morally licensed to act in a more biased manner following such disclosure (Cain, Loewenstein, and Moore 2005; on disambiguation and accountability as antidotes to unethical behavior, see also Feldman 2014).

These mixed results call for caution when deriving policy recommendations from the available findings. Some researchers have conjectured that nudge-like measures may encounter greater resistance in the context of unethical behavior than in the heuristics-and-biases context (and thus be less successful) because in the latter case, nudges are often used to align people’s decisions with what is demonstrably their true self-interest, whereas in the former context, nudges are meant to suppress the actor’s self-interest (Feldman 2014). However, if unethical behavior is often the result of automatic biases, nudge-like measures should help ordinary people who view themselves as honest to overcome those biases (as demonstrated by Shu et al. 2012).

BE studies also show that implementing mild measures of detecting and sanctioning improper behavior may sometimes be counterproductive because they reframe the situation as one of risky costs (probability of detection and sanction levels) and benefits (from unethical behavior) to the actor rather than as an ethical dilemma. Such measures may therefore crowd out intrinsic motivation to act properly (Tenbrunsel and Messick 1999; see also Langevoort 2005).

Other means may prove to be more effective. For example, to counteract displacement of responsibility (Bandura 1999), it may be desirable to emphasize the decision maker’s personal responsibility for the outcomes of his or her decisions (rather than that of superiors or the entire organization). Such means are congruent with some of the measures adopted in New Public Management reforms aimed at clarifying responsibility (Hood 1998; Schick 1996). However, those measures focused on responsibility for performance rather than ethical conduct (Hood 1991; Schwartz and Sulitzeanu-Kenan 2004). Insofar as the ethical culture in an organization affects the behavior of individuals (Cohn, Fehr, and Maréchal 2014; Kish-Gephart, Harrison, and Treviño 2010), instilling a culture of high moral standards, enforcing those standards, and associating them with people’s professional status may prove beneficial. However, the effectiveness of formal systems aimed at encouraging ethical behavior depends on their interaction with the informal systems within any organization—and because the latter vary considerably, it is difficult to come up with uniform solutions for different organizations (Smith-Crowe et al. 2015).

As further detailed later, more studies are necessary in this regard. While considerable progress has already been made in identifying the factors affecting unethical behavior, much less is known about how self-serving biases can be debiased and how unethical behavior can be counteracted (Feldman 2014). Once it is understood that improper decisions are the product of situational factors, social norms, and individual differences, simple recipes are no longer appropriate. The fact that much unethical behavior is the product of automatic processes makes simple pricing of unethical behavior—which would have affected people’s cost–benefit analysis—less effective. Furthermore, the realization that people vary in their conscious pursuit of self-interest and in their susceptibility to unconscious self-serving biases implies that different people may react differently—or even in opposite directions—to the same measures. Designing measures that are beneficial to all people, or that send different messages to different people (the notion of acoustic separation), is very challenging (Feldman and Smith 2014).

From another perspective, the perception of governmental decision makers as being self-interested and captured by rent seekers—and public officials’ resentment of this perception—results in mutual distrust between the law courts and the general public on one side and public administration on the other. This, in turn, gives rise to calls to curtail the discretion held by public officials and aggressive judicial and public oversight of the government. However, if it transpires that suboptimal governmental decisions and policies are not necessarily the product of ulterior motives, but very often of human cognitive limitations and susceptibility to automatic and largely unconscious biases, mutual suspicion may be replaced by greater trust and respect, and there may be fewer calls to limit administrative agencies’ discretion (compare Rachlinski and Farina 2002, 610–15).

Finally, zooming out from the political context and PCT, BE may be able to contribute to a better understanding of principal–agent relationships in other areas, in both the public and the private domain. These include the relationships between managers and corporations, between employees and employers, and between professionals such as attorneys and physicians and their clients. Even more fundamentally, BE can arguably account for the remarkable ability of standard economic analysis to predict human behavior despite its debatable premises about human motivation: even if people are not egoistic maximizers of their own utility, they often behave as though they are because of automatic and partly unconscious self-serving biases.

**Future Research**

While great progress has been made in recent years in behavioral studies of people’s ethicality, BE is a relatively young field of research. There is still much to be learned about the situational, personal, and social-environmental variables that affect ethical behavior and their interrelationships. In particular, most BE studies have been conducted in laboratory settings, involved simple moral dilemmas about relatively minor dishonest behavior, and used untrained and inexperienced people as subjects. While the few studies that have attempted to test BE hypotheses in field settings provide preliminary indications that the theory is not limited to isolated individual decisions of subjects in laboratory experiments (e.g., Shu et al. 2012), there is an urgent need to address the challenges of external
validity and generalizability of the (predominantly) laboratory-based findings in more complex and realistic scenarios—particularly with regard to governmental decision making.

One key difficulty facing future research in the context of governmental policy making and decision making is that the predictions of BE and standard PCT often overlap, and it is difficult to tell apart conscious self-interested calculation from automatic and partly unconscious self-serving biases. For example, under standard PCT, elected public officials who represent certain constituencies are more likely to make decisions that benefit their respective constituencies at the expense of the overall social good than officials elected by the entire population (Calabresi 1995; Weingast, Shepsle, and Johnsen 1981). This, according to PCT, is how a rational maximizer who wishes to maximize his or her chances of reelection would act. Without sharing this cynical perception of elected officials, BE would plausibly make the same prediction based on automatic psychological process. Similarly, the finding that people tend to behave more unethically when their dishonesty is less observable and verifiable is consistent both with the findings of BE studies and with deliberate calculation of the risks involved in unethical behavior.

However, as demonstrated earlier (see figure 1), the two theories also offer diverging predictions. Theoretically specifying and empirically testing these diverging predictions is a promising challenge of future studies. In addition to the propositions laid out here, contrary to standard PCT, BE predicts that for a given likelihood of being sanctioned, decision makers would be more inclined to violate a norm if the uncertainty is rooted in its ambiguity rather than imperfect detection and enforcement—as empirically found in other contexts (Feldman and Teichman 2009). The same is true of BE’s prediction that surveillance may be counterproductive because it crowds out the intrinsic motivation to act ethically.

Another set of new research questions pertains to the possible differences between elected and unelected officials in terms of their susceptibility to automatic self-serving biases, as envisaged by BE. As previously noted, the basic predictions of both PCT and BE apply to elected and unelected officials alike. However, the two groups typically differ in terms of their professional expertise, long-term goals, self-perception, and public accountability. Ideally, empirical examination of how these differences affect the self-serving behavior of these two groups and their susceptibility to various biases should involve real officials, as emulating their characteristics and roles in a laboratory setting appears to be quite difficult.

Importantly, there is considerable room for studying the efficacy of various antidotes to self-serving decisions by public officials—from simple incentives to nudge-like measures. Field experiments in which different governmental decision makers are subjected to various treatments in terms of the surveillance systems, level of enforcement, or the inculcation of ethical standards, while all other variables are kept as similar as possible, may be particularly fruitful in this regard.

In addition to academic researchers, a uniquely important role in this context may be played by behavioral task forces such as the U.S. Social and Behavioral Sciences Team and the British Behavioral Insights Team, which have been established in recent years to examine the incorporation of behavioral insights into governmental policy (Choueiki 2016; Lunn 2012). These task forces have the methodological competence and direct access to governmental bodies, which may facilitate field experiments in this sphere; they also can cooperate with academics specializing in BE. Such cooperation may also yield cross-fertilization between BE and public administration studies (compare Grimmelikhuijsen et al. 2017). Extending BE studies from the laboratory to the field, from laypeople’s behavior to that of professionals, and from the private to the public sphere, would be hugely beneficial not only to public administration and political psychology but to BE as well.

Finally, the reality of political and administrative decision making is exceedingly complex, hence it cannot be fully explained by any single theory. To use Guido Calabresi’s famous image, each theory—be it PCT, PSM, BE, or any other—offers but one view of the cathedral (Calabresi and Melamed 1972). A major challenge of future research is thus to delineate the contexts and circumstances under which each theory (or any combination thereof) provides a more fruitful explanation of reality.

Conclusion
Economic analysis is a powerful analytical tool. It facilitates the understanding of complex social phenomena through rigorous, simple models; it questions accepted truths; and yields thought-provoking (and often testable) predictions. At the same time, its simplified assumptions about human nature often cast doubt on its empirical validity. While we share the view that the behavioral assumptions of economic analysis and PCT are flawed, we point to recent psychological findings that actually reinforce most of PCT’s predictions. Although people are often not rational maximizers of their self-interest, but rather mindful of the welfare of others and attentive to moral and social norms, various automatic psychological mechanisms skew their ethical judgments and decisions in a self-serving manner. Consequently, the decisions made by public officials—be they corrupt and cynical or well intentioned—often serve their own interests and those of powerful interest groups, as though they were all rational maximizers of their own utility.

Just as one must be careful not to unquestioningly embrace the tenets of standard PCT or to assume that public decisions invariably promote the public good, one must take care when applying the insights of BE to governmental processes and decision making. We still know too little about the mechanisms underlying people’s unethical behavior, and even less about how to induce people to behave more ethically. However, if we wish to truly understand, predict, and influence governmental decisions, we must consider the findings of recent studies of ethical behavior and further study their implications for politics and public administration.

Notes
1. We leave out of the discussion individual and demographic characteristics (Kish-Gephart, Harrison, and Treviño 2010; Moore, Tanlu, and Bazerman 2012), as they are less relevant to the implications of BE for self-interest behavior of public officials.
2. Leading scholars of PSM have observed that an open question remains regarding “the relative importance of PSM compared to other motives” (Perry, Hondeghem, and Wise 2010, 688).
3. As noted earlier, situational, organizational, and social conditions may act to alter the perceived location of a choice on the horizontal axis.
Acknowledgments
We would like to thank Yoav Dotan, Asaf Eckstein, David Enoch, Yuval Feldman, Joshua Guetzkow, Daphna Lewinsohn-Zamir, Adi Libson, Ofer Malcai, Moshe Maor, and Barak Medina, as well as participants in the Faculty of Law of the Hebrew University seminar, the 2016 conference of the Israeli Association of Law and Economics (Bar-Ilan University), the 2016 meeting of the European Public Choice Society (Freiburg), and the 2016 conference of the European Group for Public Administration (Utrecht) for valuable comments on earlier drafts of this article. Shmuel Baron, Tal Mendelson, Eldad Spiegelman, and Roi Yair provided excellent research assistance. This research was supported by the I-CORE Program of the Planning and Budgeting Committee and the Israel Science Foundation (Grant no. 1821/12).

References


