Schnellenbach, Jan; Schubert, Christian

Working Paper
Behavioral Political Economy: A Survey

CESifo Working Paper, No. 4988

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Schnellenbach, Jan; Schubert, Christian (2014) : Behavioral Political Economy: A Survey, CESifo Working Paper, No. 4988, Center for Economic Studies and Ifo Institute (CESifo), Munich

This Version is available at:
http://hdl.handle.net/10419/103131

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.
If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Standard-Nutzungsbedingungen:
Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.
Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.
Sofern die Verfasser die Dokumente unter Open-Content-Lizensen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.
Behavioral Political Economy: A Survey

Jan Schnellenbach
Christian Schubert

CESifo Working Papers

CESifo Working Paper No. 4988
Category 2: Public Choice
September 2014

An electronic version of the paper may be downloaded
• from the SSRN website: www.SSRN.com
• from the RePEc website: www.RePEc.org
• from the CESifo website: www.CESifo-group.org/wp
Behavioral Political Economy: A Survey

Abstract

Explaining individual behavior in politics should rely on the same motivational assumptions as explaining behavior in the market: That’s what Political Economy, understood as the application of economics to the study of political processes, is all about. In its standard variant, those who played the game of politics should also be considered rational and self-interested, unlike the benevolent despot of earlier models. History repeats itself with the rise of behavioral economics: Assuming cognitive biases to be present in the market, but not in politics, behavioral economists often call for government to intervene in a “benevolent” way. Recently, however, political economists have started to apply behavioral economics insights to the study of political processes, thereby re-establishing a unified methodology. This paper surveys the current state of the emerging field of “Behavioral Political Economy” and considers the scope for further research.

JEL-Code: D780, D030, A120, D720.

Keywords: behavioral political economy, behavioral economics, rational irrationality, cognitive biases, social norms, voting, paternalism.

Jan Schnellenbach  
Walter Eucken Institute / Freiburg  
& University of Heidelberg  
Germany  
schnellenbach@eucken.de

Christian Schubert  
University of Kassel  
Institute of Economics  
Kassel / Germany  
dr.c.schubert@gmail.com
1. Introduction

Explanations of individual behavior in politics should rely on the same motivational assumptions that guide the economic analysis of market behavior. And the systemic processes that generate political outcomes should be understood as reflecting the interplay of supply and demand. That is what Political Economy – to be understood as the application of economic analysis to the study of political processes – is all about. In its standard (neoclassical) variant, aligning the motivational assumptions means, of course, to model agents as maximizing their expected utility not only in the marketplace, but also in the political arena – as voters, politicians, lobbyists or bureaucrats. As Brennan (2008) emphasizes, though, motivational symmetry does not necessarily translate into behavioral symmetry, given the weak incentives to invest in rational decision-making that are prevalent in the realm of collective decision-making.\footnote{This point seems to be supported by Kirchgässner (2008), referring to Adam Smith’s (1759) complex motivational theory.} This holds in particular for voters, whose perfectly rational individual behavior may lead to catastrophic collective outcomes, because in the political arena, individual feedback and learning mechanisms in politics are much weaker and more indirect than in the marketplace.

Thus, cognitive biases can be expected to play at least as important a role in politics as in the market. Political Economy is therefore one of the fields particularly likely to benefit from applying behavioral economics insights. Strikingly, though, behavioral approaches in Political Economy are by and large a relatively recent phenomenon. An explanation may be that the application of economic tools of analysis to the political sphere had been dominated by the Virginia School of Public Choice, whose members were busy exporting the rational choice approach to non-market areas of research, a project that seemed more urgent than improving economist’s toolbox itself (Wallerstein 2004). Things are more complex, though, when it comes to the leading figure in Public Choice, James Buchanan: While insisting on using a narrow version of homo economicus in constitutional economics, he was open to the use of non-standard behavioral assumptions to guide empirical research in Public Choice, and even believed non-standard assumptions to be more relevant in explaining decisions in politics than decisions in markets (Kirchgässner 2014). It is also noteworthy that parts of modern experimental economics can in fact be traced back to the Virginia School of Public Choice (Plott forthcoming). For obvious spatial restrictions, we will focus on Behavioral Political Economy, rather than on...
Experimental Political Economy. In general, Behavioral Economics and Experimental Economics are separate and only partly overlapping research fields (see, e.g., Heukelom 2014 on their historical background).

In applying behavioral approaches to Political Economy, stating that individuals are more prone to biases and other cognitive problems when they enter the political arena cannot be the end of the story. Rather, a systematic analysis of deviations from the baseline assumption of perfectly rational, optimizing actors is called for. Accordingly, many authors have encouraged scholars in the field to venture beyond the usual assumptions on rationality and utility-maximization (e.g. Simon 1995; Ostrom 1998; Kliemt 2005).

In this paper, we look at the essential contributions that have been made so far, and discuss promising avenues for future research. As to labels, we chose to follow Della Vigna (2009) and Berggren (2012) in speaking of “behavioral political economy” (henceforth BPE). We define BPE as the psychologically informed, economic analysis of behavior and its effects in the political arena. We distinguish between a “weak” and a “strong” variant of BPE. The former merely alters specific auxiliary assumptions on either agents’ cognitive capacities or the content of their utility functions, by arguing, for example, that voters not only care about political outcomes, but also about their “citizen duty” when going to the polls, or that they care about other aspects that seem irrelevant from an orthodox instrumental standpoint, like a candidate’s looks. In other words, weak BPE examines the implications of non-standard preferences, or cognitive biases – formerly known as “anomalies” –, as well as limits of attention and self-control, on the behavior of those active in the political arena. Useful as they may be, we

---

2 On which see, e.g., Fiorina & Plott (1978), Plott (2015), Palfrey (forthcoming) and the surveys on political behavior in Plott & Smith (2008).
3 From a methodological viewpoint, one may object, for instance, that the rationality assumption is essentially irrefutable: Under suitably redefined constraints, everyone can be described as acting rationally (e.g. Boland 1981). As Pelikan (2010) shows, however, this perspective breaks down as soon as we leave the one-person problem and examine settings with heterogeneous agents that differ in what is “best” for them. Thus, in the area of Public Choice we can safely treat rational choice as a refutable (and often refuted) empirical hypothesis.
4 See Wallerstein (2004). The definition suggested by Della Vigna (2009: 361) seems too narrow: According to him, BPE is about how “politicians change their behavior to respond to voter biases” such as limited attention (ibid.: 364); he contrasts it from “behavioral institutional design”, which is about the use of behavioral economics insights in designing policies (ibid.: 361, 364-365) and would then be closely related to “Behavioral Public Finance” (e.g. Mullainathan et al. 2012). “Behavioral politics”, suggested by Brennan (2008) seems to connote a normative bias. Berggren (2012: 199) defines BPE, also rather narrowly, as the application of the analytical tools of behavioral economics to “political decision-makers”.
5 Strictly speaking, merely introducing non-standard and non-procedural content into the utility function should not qualify as BPE at all, since the orthodox approach is agnostic as to what an agent chooses, as long as it’s outcomes. We suggest that future discussions of the field should take this into account.
therefore do not count models that merely incorporate information asymmetries or uncertainty into an otherwise standard setting as contributions to BPE. The “strong” variant of BPE goes beyond this and attempts to actually explain (rather than just postulate) motivational and other psychologically informed extensions to the standard model. For instance, it may try to examine the mental processes causing differences in agents’ susceptibility to certain biases. In the parlance of Simon (1978), while weak BPE joins orthodox economics in focusing on rational choice as a “product of thought”, strong BPE takes account of the processes underlying and preceding choice itself. We will discuss contributions from both types of BPE in this paper. We have decided to refrain from recapitulating the most important insights of behavioral economics in a separate section, referring the reader instead to the excellent surveys by Rabin (1998), Camerer et al. (2004), and of course to Kahneman (2011).\footnote{See also Kamenica (2012).}

There are a number of intellectual ancestors to BPE. Not surprisingly, Adam Smith already ventured into this territory when speculating, first, that a key rationale for government to exist is to protect private property from the transgressions of those who depart from “reasonable” behavior (which, for Smith, implied acting morally). Second, he argued that rational individuals will underinvest in the quality of political decisions: As George Stigler observed, Smith was a pessimist in this respect in the sense that he “gave a larger role to emotion, prejudice, and ignorance in political life than he ever allowed in ordinary economic affairs” (Stigler 1982: 140).\footnote{On Smith in this context see also Buchanan (1976). Smith’s argument, however, is not a harbinger of the modern concept of rational ignorance. It is rather an institutionalist argument: Individuals, from landowners to monarchs, are simply not trained for making good political decisions, or their habits constrain them from making them. Smith sees ignorance and inaptitude as unfortunate realities in political affairs, but they are not explained themselves as the outcomes of a rational choice between alternatives.} After Smith, the belief that individuals lose some of their capacity for well-informed, thorough reasoning upon entering the political sphere can be found, e.g., in Mill (1848), who in defense of laissez-faire warns of herd behavior in politics (Ch. XI, §3), incompetence due to a lack of specialization (§4), bad incentives due to small individual stakes in government decision-making (§5) and the danger that individuals lose their ability to voluntarily contribute to public goods if they get used to delegating more and more competencies to the government (§6). Mill therefore goes a step further than Smith, in offering explanations in terms of incentives for the supposedly low quality of collective action, such as governments crowding out the informal institutions underlying voluntary cooperation.
Taking earlier arguments such as those by Smith and Mill to the extreme, Joseph Schumpeter had a particularly grim perspective on the political process. For him, irrationality could be explained by the incentive structure prevalent there: “[The private citizen] expends less disciplined effort on mastering a political problem than he expends on a game of bridge ... Thus, the typical citizen drops down to a lower level of mental performance as soon as he enters the political field. He argues and analyzes in a way which he would readily recognize as infantile within the sphere of his private interests. He becomes a primitive again. His thinking becomes associative and affective.” (Schumpeter 1942: 261-62).8

Turning our attention to more recent developments in BPE, this paper is organized as follows. In the following sections, we discuss contributions that apply behavioral economics insights to the different sets of actors involved in the political game: voters in section 2, politicians in section 3, and bureaucrats, regulators and lobbyists in section 4. Section 5 briefly gives two examples of applications of BPE: the issue of “libertarian paternalism”, and the risk that the modern welfare state may foster irrational behavior. Section 6 offers an outlook.

2. Voter preferences and voter behavior

Voter behavior has received the bulk of attention by psychologically informed Political Economy. This is not surprising, since many empirical observations on voting behavior are difficult to reconcile with a basic rational choice model, unless it is enriched with additional assumptions that are, to a large extent, borrowed from psychology and also from political science. We will discuss, in turn, the way psychologically informed Political Economy has dealt with the voting paradox (by introducing a variety of direct benefits from the act of voting), followed by the phenomenon of expressive, rather than instrumental voting. Subsection 2.3 discusses Bryan Caplan’s model of “rational irrationality”, followed by the phenomenon of retrospective voting in subsection 2.4. A discussion of attempts to explain non-standard preferences for policies by, e.g., invoking framing effects concludes Section 2.

2.1. The paradox of voting

---

8 On Schumpeter’s general take on politics, see also Schubert (2013).
At the heart of Political Economy, as we define it, lies the “embarrassing predicament”\(^9\) of being unable to explain why people would bother to vote at all: Given the extremely low probability that a single vote will turn out to be decisive in large-scale majoritarian elections, there is no point in voting for instrumental reasons.\(^10\) To the extent that they were willing to explore voter motivation at all,\(^11\) early proponents of (weak) BPE avant la lettre tried to resolve this paradox by introducing non-standard arguments into peoples’ utility functions; this allowed them to stick to the utility maximization framework and to reconcile the act of voting with rational behavior. In other words, the “benefit” part in individuals’ cost-benefit calculus was modified.\(^12\) Most attempts to cope with the paradox of voting, still belong to what we refer to as weak BPE.\(^13\)

Already Downs (1957: 267) suggested the existence of an individual “sense of social responsibility” for the political system as a whole. People vote because they have been taught that voter participation is a necessary precondition for a well-functioning democracy, i.e. that high voter turnout is an end in itself.\(^14\) Individuals who harbor a preference for democratic institutions develop a sense of responsibility and vote in order to sustain these institutions. But in doing so, they voluntarily participate in the production of a pure public good (Samuelson 1954) which clearly should not be expected under strict rationality.

This immediately leads to the question why many individuals contribute voluntarily to this particular public good, and at the same time fail to contribute to others. In a much-cited contribution to the field, Riker & Ordeshook (1968: 27-28) suggested the extra argument \(D\) in the voter’s cost-benefit calculus, which can represent a voter’s concerns for the functioning of her polity, but also other intrinsic benefits provided by the act of voting itself, such as the warm glow of acting in a prosocial way, or the desire to avoid informal sanctions if peer pressure.

---

\(^9\) In the words of Ferejohn & Fiorina (1974: 525).

\(^10\) Brennan & Buchanan (1984) describe the relevant probability calculus.

\(^11\) See, e.g., Brennan & Buchanan (1984): “We cannot ... explain voter behavior in terms of preferences over outcomes ... People vote because they want to – period. And they vote how they want to – period.” (ibid.: 187).

\(^12\) Note that the “cost” part’s impact is rather uncontroversial: There is considerable empirical evidence that turnout declines as the costs of voting rise (Mueller 2003: 329). On the determinants of turnout see also Degan & Li (2015) who offer a model where voting is based on anticipated ex-post regret.

\(^13\) See, however, the end of this subsection for a notable exception of a strong BPE treatment of the paradox of voting, and Feddersen (2004) for non-behavioral approaches.

\(^14\) See also Tullock (2000). Complying with the social norms thereby imposed could be a source of utility in its own right, particularly if others comply as well (and can observe one’s own degree of compliance). See, e.g. Frey & Meier (2004).
demands participation (see also Tullock 1967). This explanatory strategy has been criticized as a sleight of hand, exemplifying “post hoc theorizing” (Green & Shapiro 1994: 50ff.) that effectively eliminates the predictive powers of rational choice theory (Mueller 2003: 306). In order to rectify this shortcoming, it is obviously necessary to find out which factors determine the postulated “sense of social responsibility”.

Carlsson & Johansson-Stenman (2010: 502-03), for instance, find that compliance with social norms as a motive for voting is more effective among women and the elderly. Empirical evidence of this kind is important because it sheds some light on the determinants of citizen duty, but it also calls for generalized theoretical explanations. It would thus be a task for (“strong”) BPE to understand how individuals learn to follow certain norms and behave accordingly, e.g., in an altruistic way. Norm-guided behavior would have to be understood as a conditional response to certain stimuli, and a response that is rooted in a history of past reinforcement experiences. This would require a relaxation of the rationality assumption, since the degree of norm-compliance would not be treated as a choice variable depending on expected rewards, but rather as a function of an individual’s conditioning history. On the other hand, the self-interest assumption would be retained.

The reverse, i.e., relaxing the assumption of self-interest while retaining the assumption of full rationality, is also possible. Kirchgässner (1992) argues that the costs of moral behavior matter: Since casting a vote is associated with very low costs in most established democracies, individuals can behave morally and contribute to the pure public good at a very low price. Under these circumstances, the argument that for many individuals is larger than the costs of voting gains plausibility. Hence, the focus shifts to the restrictions of decision-making. Individuals are – without further explanation – assumed to be inclined to cooperate, but their degree of actual cooperation depends on a rational assessment of the costs of doing so.

Among the non-standard sources of utility from voting discussed in the literature, we also find some that have less to do with social learning and interaction, and more with direct individual benefits. In particular, procedural utility and the utility gained by expressing one’s opinions are important. Consider the former: Already Buchanan (1954: 341) hinted at the

---

15 Della Vigna et al. (2013) present evidence from a field experiment that people vote “because others will ask”, i.e. for reasons of “social image” (esteem). A necessary assumption is, of course, that lying is costly.
16 Mueller (2003: 325-29) makes this case for what he refers to as “ethical” behavior, on which see also Morton (2015).
possibility that the “[v]oting choice does provide individuals with a greater sense of participation in social decision-making”. In modern behavioral economics, a sense of participation is seen as an important source of hedonic, process-based benefits. In recent years, interest in procedural utility has mushroomed. This type of utility is generally produced by personal feelings of being involved and respected. The psychological background is offered by the “self-determination theory” (Deci & Ryan 2000), which postulates that human beings are universally motivated by innate needs for competence, relatedness and autonomy. Jointly, these needs constitute the desire for self-determination. There is by now ample empirical evidence for the importance of these “pleasures of process” associated with the institutions that structure collective decision-making (e.g. Frey & Stutzer 2010; Hirschman 1989). For example, direct democracy appears to provide voters with procedural benefits by letting them participate in the decision-making process, independent of altering results in their favor (Frey & Stutzer 2005; Olken 2010).

Finally, an argument to be subsumed under the “strong” branch of BPE has been sketched by Riker & Ordeshook (1968), and elaborated upon by Quattrone & Tversky (1988). Riker & Ordeshook briefly speculate about the possibility that voters may overestimate the probability of their personal vote being decisive (ibid.: 38-9). Quattrone & Tversky lend some psychological substance to this idea by advancing a “voter’s illusion” hypothesis: Given the fact that people often fail to distinguish between causality and mere correlation, they argue that “some people may reason that if they decide to vote, that decision would imply that others with similar political attitudes would also decide to vote”, with the term “imply” meaning something like “serve as diagnostic” (ibid.: 733-34). If that were true (the authors in fact offer some supporting empirical evidence), this cognitive error would significantly inflate the perceived subjective probability of one’s vote being decisive. Recently, Dittmann et al. (2014) have produced experimental evidence for a positive “voting premium”, indicating either voters’ overconfidence in their own pivotal role, or even the neglect of the issue of pivotality itself. When that issue was made salient in the experiment, subjects displayed strong signs of overconfidence and overestimated the likelihood that others were biased as well. Evidence on the positive correlation between education and voting (e.g. Milligan et al. 2004) – and

---

17 See Frey et al. (2004) and Frey (2008: ch. 10) for surveys and Schubert (2012) for an application to “happiness politics”.
18 See also Urbatsch (2012) on the chances that less intelligent voters may overestimate the instrumental value of voting.
instrumental voting to boot —, appears to make a widespread voter’s illusion highly unlikely, though.\textsuperscript{19}

2.2. \textbf{Expressive Voting}

The specific source of utility referred to as \textit{expressive utility} is by now probably \textit{the} most widely accepted element of BPE.\textsuperscript{20} Hamlin and Jennings (2011) define the “expressive” aspects of voting behavior as reflecting benefits from the act of voting that neither derive from its instrumental nor from its consumptive value, but from its symbolic or representational aspect: not from the act, but from its meaning” (ibid.: 5). A much-cited analogy to illustrate that particular aspect of voting’s intrinsic benefits has been offered by Brennan and Buchanan (1984: 186, 196) who likened them to the utility gained by cheering at a sport event in order to “show preferences as such”. This source of intrinsic benefits is then positively reinforced by the victory of one’s preferred candidate (as it is by the victory of one’s preferred football team).\textsuperscript{21}

Expressive voting is not only able to explain why citizens vote, but also how they vote. Only individually inconsequential choices are “purely expressive” (Hamlin and Jennings 2011: 649), which is why the voting booth is a place most likely to encourage expressive action, at least in large-scale majoritarian elections.\textsuperscript{22} There, the individual voter is free to forego all instrumental considerations and to indulge in expressive benefits as she sees fit. In other words, voting offers her the opportunity to express her views about the public interest. The implications at the normative level are, however, unclear: It may well be that expressive voters act altruistically: As Mueller (2003: 321) argues, “since voting involves collective decisions that affect all members of the community, norms that govern conduct toward others might be expected to be particularly likely to come into play when individuals vote” (see also Morton 2015).

\textsuperscript{19} See also Carlsson & Stenman-Johansson (2010: 502).
\textsuperscript{20} There is by now a lot of empirical support for this hypothesis, see, e.g., Carter & Guerette (1992), Jones & Hudson (2000), Tyran (2004).
\textsuperscript{21} Riker & Ordeshook (1968: 28) arguably came close to that idea when counting among the determinants of their variable “D” the “satisfaction from affirming a partisan preference”.
\textsuperscript{22} It has been argued that the positive correlation between the closeness of an election and voter turnout suggests that instrumental reasons do outweigh expressive concerns (e.g. Geys 2006). As Rotemberg (2009) argues, though, the theory of expressive voting can very well explain why psychological benefits from voting are greater the closer the outcome is expected to be.
There are several arguments that can explain why individuals find it worthwhile to vote expressively rather than instrumentally. We may sort these arguments into three subsets. First, an act of voting may help the voter confirm her personal identity (either to herself or to others). The voter may achieve this, for instance, by identifying with some specific social group and gaining in esteem (or self-esteem) when succeeding: In both cases, voting may also help her confirm a personal identity she wishes to project, either on herself or on the group she identifies with. Political parties can then be modeled as clubs that offer attractive combinations of membership type and membership numbers. The effect of negative campaigning, which aims at strictly differentiating the identities of political groups, can be understood from such a perspective (Hamlin & Jennings 2011: 657).

Expressive behavior may be motivated by the wish to identify with political parties, positions or candidates and their perceived characteristics. Individuals, though, also appear to prefer candidates that already are similar to how they perceive themselves (McDermott 2009). Accordingly, Brennan and Hamlin (1998) offer a model that assumes that, if elections are dominated by expressive issues, people vote for positions that are sufficiently close to their own expressive bliss point. If positions are too far removed from that bliss point, they feel “alienated” and abstain. Candidates adjust their positions accordingly. Interestingly, the predictions differ markedly from the standard median voter model with endogenous participation: While there, moderate voters abstain because they are indifferent between the positions, in Brennan & Hamlin’s model, the extremists abstain because they feel alienated.

Expressive, identity-based voting can also lead to situations where money is left on the table. Hillman et al. (2015) demonstrate this for the case of clan-based voting, where clan members are pressured by their leadership to participate in local elections, where clans are large enough to exert a material influence on policy outcomes, but not in national elections.

---

23 Carlsson & Johansson-Stenman (2010) find that people generally believe that they themselves vote less selfishly (where selfish motivation is considered detrimental to one’s self-image) than do others. “Moreover, people tend to believe that others with similar political views as themselves vote less selfishly than do people with the opposite political views” (ibid.: 498, 504-07). On the economics of identity, see, of course, Akerlof & Kranton (2000). On the value of a positive self-image, see, e.g. Baumeister (1998). Hillman (2010) defines all expressive behavior as being motivated by the wish to confirm (the discretionary aspect of) identity.

24 See, e.g., Schuessler (2000).

25 See also Hamlin & Jennings (2004).

26 Another important implication is that cycling due to the formation of multiple political equilibria is unlikely, because positions that are too far from voters’ expressive preferences will simply never be selected (Hamlin & Jennings 2011: 659). Alienation as a driver of abstention has been empirically confirmed by, e.g., Guttman et al. (1994).
Furthermore, elected representatives with a clan identity sometimes fail to agree on beneficial compromises, even if a compromise would be required to sustain future influence on local policy outcomes and access to rents. Hillman et al. (2015) explain this with expressive utility that can be secured by avoiding a compromise that would be materially beneficial, but would also imply an affront against the own clan-based identity.

Second, expressive behavior may be motivated by the benefits of *acting morally*. This argument also relates to the identity (“self-image”) considerations discussed above (see also Kliemt 1986), and it is not always easy to distinguish whether a person acts morally in order to cultivate a certain identity, or for other reasons such as receiving a warm glow from behaving morally. For instance, voters may wish to select candidates who display desirable moral qualities, which could *prima facie* serve as an argument for representative (as opposed to direct) democracy. On closer inspection, however, problems can arise when the perceived moral qualities of candidates become much more salient than their political positions or skills. Voters may then opt for candidates who, while apparently virtuous, hold irrational views on political issues. Moral concerns may also play a role when citizens vote for redistribution, even though instrumental considerations would point into a different direction. Voting for redistribution as a means to express one’s altruistic disposition – or to at least appear altruistic or “charitable” – is associated with zero expected costs in large elections. Experimental evidence indicates that voters are indeed more inclined to vote expressively when their probability of being decisive declines – in this sense, they act perfectly rationally (Shayo & Harel 2012).

Third, expressive voting may be motivated by a desire to express one’s gratitude in return for a political candidate that is perceived as “supporting” the voter. The idea to ground the expressive voter hypothesis on the theory of reciprocal giving was suggested by Fiorina (1976) and later by Brennan & Buchanan (1984: 198).

Further theoretical and empirical research is needed to identify the determinants of an individual’s expressive preference order, and also of possible conflicts between expressive and

---


28 As Buchanan (1954: 337) put it, though, “it is extremely difficult to determine whether the affirmative vote of a non-beneficiary individual for a public welfare project implies that he is either acting socially in accordance with a ‘nobler’ ordering of alternatives … or whether … he has failed to weigh adequately the opportunity costs of the project”.

29 See Tullock (1971).
instrumental preference orders. So far, the theory of expressive voting does not allow us to predict under which (institutional) conditions different kinds of expressive preferences emerge. Furthermore, even though the theoretical concept is clear, it may be difficult to empirically distinguish between expressive and instrumental preferences in actual policy-making processes (see above and FN 28). Another problem is that, even if it were possible to clearly distinguish expressive from instrumental preferences, the instrumental preferences could not per se be considered a superior representation of the citizen’s “true” preferences.

As Hamlin and Jennings (2011: 647) opine, “expressive motivations open up a new area of study which allow rational choice techniques to be employed in ways that more accurately reflect the meaning and symbolic significance of much political behavior”. Political outcomes should, in their view, be seen as resulting from the interplay between expressive and instrumental behavior. For instance, strong expressive voting on the part of citizens is likely to induce professional politicians to act instrumentally, in order to manufacture their positions so as to appeal to those voters. Hamlin and Jennings (2011: 653) use the example of the election speech to illustrate the interaction between instrumental and expressive motives in detail.

The effects of expressive voting are ambiguous. It may give rise to wasteful or even disastrous outcomes. For instance, it may result in inefficient election outcomes if every single voter free-rides on the non-expressive votes that she expects to be cast by others. For example, generally unsupported tax hikes may end up not being vetoed by a majority if they supposedly serve some “moral” ends, thus generating expressive utility (Kliemt 1986: 335-337). Hillman (2010) generalizes this approach and diagnoses “expressive-policy traps”, with majority-supported policies that the majority does not actually want (ibid.: 405). Inefficient foreign development aid exemplifies this. Hillman identifies institutional safeguards, though: At least when negative instrumental side-effects are large, the election winner may be expected not to actually implement expressive campaign promises, for fear of losing electoral support in the

---

31 Both sets of preferences can be held on the basis of full information and full consideration, even when in conflict: They may just be concerned with different domains. On the other hand, both sets can be subject to “framing” and manipulation; there is no “neutral” institutional setting that is able to reveal “true” preferences (Hamlin & Jennings 2011: 663). To illustrate, the expressive environment of political choice may be more appropriate in inducing a welfare-increasing supply of merit goods (assuming welfare to be defined in hedonic terms). On merit goods, see Brennan & Lomasky (1993).
32 See the “inefficient unanimity” discussed by Brennan & Lomasky (1984). A recent example is Jennings (2015), who shows how an expressive preference for upward social mobility leads to an inefficient outcome in the British education system.
subsequent election (ibid.: 414). This does, however, not seem to apply in all domains where the drive for expressive utility shapes policy outcomes: Lomasky (2008: 480-81) gives the example of the U.S. penal system as illustrating a case where “majoritarian enthusiasm” has led society into a negative-sum game. However the effects are evaluated, an interesting detail in these models is the assumption that politicians react in a perfectly rational way to the expressive behavior exhibited by voters.

What kind of institutional implications follow from the expressive voter hypothesis? As Brennan & Hamlin (2000) put it, the economic advisor should identify those institutional arrangements that are likely to prompt some particular (e.g., desirable) kinds of expressive behavior rather than others. For instance, the design of constitutions may be affected by expressive concerns – importing all sorts of political symbolism into it –, a fact that may limit the constitution’s operational efficacy. This applies chiefly to written constitutions, and in particular to those subject to popular referenda (Brennan & Hamlin 2002). The extent to which alternative institutions are able to elicit “responsible” (i.e., not inconsequential) behavior becomes an important issue for a yet to be developed theory of behavioral constitutional economics.

As an aside, note that those citizens who suffer from a significant voter’s illusion (see above) are less likely to behave expressively, since they believe in the material significance of their vote. In a second best world, where expressive behavior cannot be controlled directly, but responsible, as-if-consequential voting is desired, it may be efficient to amplify voter’s illusion in order to reduce, at least partially, the propensity for expressive political behavior.

2.3 “Rational irrationality” and biased beliefs

Downs (1957) already speculated that rational voters may invest very little in searching for information on political issues, thus staying “rationally ignorant”. An important step beyond this insight was to admit the possibility that voters may not only be uninformed, but that they perceive and process the little information that comes to their attention in a systematically biased way. As Akerlof (1989) showed, they do so in order to gain benefits from projecting an attractive

---

33 Well-informed electoral choice is a public good and will be insufficiently supplied – interestingly, in the case of voting also by the altruistically-minded (Lomasky 2008: 473).
self-image to themselves. Certain privately held “bliss beliefs” seem to be able to generate “good feelings” (see also Akerlof & Dickens 1982: 308-310).

Accordingly, a popular way to capture the apparent “irrationality” in voting behavior is to model beliefs as a normal good, the consumption of which is sensitive to the level and change of opportunity costs. Voters are assumed to prefer certain beliefs over others. There is also empirically well-supported evidence that people tend to stick to these beliefs even in the face of obviously falsifying evidence. The psychological mechanisms giving rise to this behavior remain unclear, though. As we will see, this extension of the standard utility function has important implications for a variety of issues in Political Economy, such as the question whether we can expect democracy to generate efficient outcomes.

In a seminal contribution, Wittman (1989) had argued that with rational expectations, voters will never accept democratic outcomes to be inefficient. In other words, the positive amount of state failure that can be expected will match the optimal amount, given costs and benefits of preventing inefficiencies. As Caplan (2001a) shows, however, the insignificance of an individual vote, combined with the (empirically well-supported) assumption that voters have preferences over beliefs implies that agents are “rationally irrational”, leading to inefficient political failures: “Collectively, people leave $20 bills on the sidewalk, but individually they do not” (ibid.: 312). Agents’ cognitive errors may even outweigh standard principal-agent issues in generating political failures (e.g. Caplan 2009).

Caplan’s model treats irrationality – understood as deviations from rational expectations – as a by-product of consuming a particular good: agents derive utility from holding irrational beliefs; they “want to believe what is pleasant” (ibid.: 319) and they trade-off holding irrational beliefs against wealth. Importantly, voters’ deviations from rational expectations are not randomly distributed, but rather systematically biased: They do not cancel each other out on average. Caplan assumes that individual preferences for irrationality can be represented by indifference curves for wealth and irrationality, i.e., irrationality is a consumption good that is traded-off against material wealth. However, the indifference curves are slightly non-standard, because they are C-shaped, and thus contain a range where their slope is positive. In the negatively sloped area, individuals would be willing to give up wealth for being able to consume

---

34 An important implication would be that the key assumption underlying principal-agent models, viz., that “when leaders stray, voters suffer” (ibid.: 161), may no longer apply. Interestingly, by making voters neglect the beneficial effects of incentive schemes for politicians, voter irrationality may be the source of principal-agent problems (ibid.).
a bit more irrationality. In the positively sloped area, they would need to receive more wealth to be willing to become even more irrational. In plain words, they do not desire to become as irrational as possible, but they have some “bliss belief” with a level of irrationality that they consider ideal. If the expected loss of wealth from holding that belief is zero – which is the case in large-scale elections – the bliss belief is always held.\textsuperscript{35} Holding one’s bliss belief means deliberately ignoring the negative side-effects of that belief; again, individuals do so because they act behind a veil of insignificance.

If however expected wealth would have to be sacrificed for being irrational (e.g., when committees are small), then the preference for irrationality yields a downward-sloping demand curve, with the price of irrational beliefs given by the private wealth losses generated by a given bias under given circumstances. On a smooth trade-off, individuals would then sustain some level of irrationality, somewhere between perfect rationality and the degree of irrationality implied by their personal bliss belief. Importantly, Caplan assumes agents to have rational expectations about the price of irrationality (ibid.: 312-313). He concludes that what he refers to as “near-neoclassical agents” (ibid.: 313) are not only ill-informed; rationally irrational voters also have no incentives to process information in a rational way, which makes their beliefs subject to systematic biases. What is more, Caplan’s model can also capture the overconfidence irrational voters tend to display with respect to the “truth” of their opinions, which they are often unwilling to correct in light of new evidence (resulting in a persistence that would be hard to explain when voters were merely rationally ignorant, ibid.: 315).

Specifically, Caplan identifies four biases that are widely cherished, namely, the “antimarke”, the “antiforeign”, the “make-work” and the “pessimistic” bias.\textsuperscript{36} All these cognitive errors may be subsumed under one key bias, viz., “a disposition to lend undue weight to what is readily observed at the expense of appreciating what is below the surface” (Lomasky 2008: 471). Caplan cites empirical evidence supporting this claim: According to the “Survey of Americans and Economists on the Economy” (1996), laypersons’ views on matters of economic policy differ markedly from those of professional economists, the more so, the less educated

\textsuperscript{35} Note, again, the stark contrast to private consumer choices in the marketplace.

\textsuperscript{36} In turn, they denote people’s tendency to underestimate the social value of market coordination and international trade, to overestimate the social value of job conservation (relative to productivity growth), and the belief that matters economic go from bad to worse over time.
laypersons are.\textsuperscript{37} Beyond the realm of opinions about fact, laypersons also differ from most economists with respect to the social values they would like policymakers to pursue: Fairness typically ranks far ahead of efficiency, whether static or dynamic (see also Haferkamp et al. 2009). Caplan’s “rational irrationality” approach has been applied to the economics of political reform by Thomas et al. (2013). They show how the repeal of the U.S. Prohibition laws became possible despite voters holding irrational beliefs on their effects.

As in models of expressive voting, an important implicit assumption in Caplan’s reasoning is that politicians themselves do not suffer from perception biases. They are generally assumed to care about the odds of being reelected, rather than about the issues themselves. But they are also assumed to have more or less unbiased perceptions of the theories that are held by voters and that determine voting behavior.

Caplan’s approach has been criticized from many different angles. Lomasky (2008), for instance, questions whether the beliefs observed by Caplan qualify as “irrational”, since under the incentive conditions encountered in the political arena, voters are well advised to invest little in “belief cultivation”. In that case, Caplan’s would rather be a theory of “rational rationality”. But it seems that Lomasky overlooks the role of systematic biases people are subject to. Moreover, at least for the trained economist, the label “irrational” need not bear problematic normative connotations. After all, the biases identified by Caplan may very well result from people following heuristics that, in turn, are a perfectly sensible way to cope with the problems of, say, uncertainty in everyday life. In light of this, then, one may argue that a vote for a protectionist policy reflects rationally pursued expressive concerns rather than “irrationality” (Lomasky 2008).\textsuperscript{38}

\textbf{2.4 Retrospective voting}

\textsuperscript{37} Generally, more knowledge tends to make people more “conservative” (pro-market) economically and more liberal socially (Caplan 2009: 162). Some parts of the questionnaire used in that survey appear problematic, though. Consider, e.g., the following claim by Caplan (2006: 27): “[The survey] asks which of the following two positions is closer to the respondent’s views: One, we need a strong government to handle today’s complex economic problems; or two, the free market can handle these problems without government becoming involved. Fully informed opinion was more pro-market”. Further evidence for peculiarities of laypersons’ views is provided by Bastounis et al. (2004).\textsuperscript{38} In light of the 2008-2009 financial crisis, Lomasky (ibid.) also argues that the overall social costs generated by economic “irrationality” may well be small relative to the damage done by “economic virtuosi” (ibid.: 479).
Another objection has recently been advanced by Bischoff & Siemers (2013). They provide the first formal (game-theoretic) model of party competition that links the existence of persistently biased beliefs to the phenomenon of retrospective voting. As to the former, they argue that Caplan’s approach cannot explain why individuals are so reluctant to adapt mundane beliefs about, say, the impact of the minimum wage, to the empirical evidence. By focusing on the actual psychological process underlying voter choice, their model qualifies as a contribution to strong BPE. They suggest a micro-foundation that is based on the concept of mental models from cognitive psychology (e.g. Johnson-Laird 1983, Denzau & North 1994). A mental model is a highly simplified subjective representation of a complex real-world system, such as the economy. Due to their varied personal experiences and social background, agents typically entertain heterogeneous mental models; therefore, they reach divergent conclusions when reasoning, through a process of “mental simulation”, about the effects of some policy measure. Bischoff and Siemers argue that voters’ “irrational” assessments of the impact of policies do not reflect isolated beliefs, as implied by Caplan. Rather, they reflect conclusions from such “mental simulations”. That is why an agent cannot easily adapt these assessments without changing the whole mental model that gives rise to her conclusion. Hence the persistence of voters’ biased beliefs.

What happens, then, if reality contradicts these conclusions? Voters are assumed to reduce the ensuing cognitive dissonance, not through changing their mental models (that would be too costly), but rather by updating their beliefs about the “competence” of the political incumbent. That in turn gives rise to the phenomenon of retrospective voting: incumbents are held responsible for the state of the economy at election time; they are re-elected if the state is perceived as “good”, and voted out of office otherwise. This of course confronts policymakers with an intricate trade-off, on which they rationally maximize their expected utility: By pursuing popular, yet objectively counterproductive policies, they may maximize their short-term popularity, but risk worsening the economic outcome, thus losing perceived competence and decreasing their re-election probability. Thus, retrospective voting may serve as a mechanism that self-corrects politicians’ temptation to give in to voters’ biased beliefs.

In equilibrium, parties will choose a policy mix, combining good policies with bad, yet popular ones that ensure public acceptance; jointly, this generates mediocre outcomes. Bischoff & Siemers predict that the quality of policies depends negatively on the popularity of inadequate
mental models among voters and on the degree of polarization in the distribution of mental models; it depends positively on voters’ accounting of the parties’ past macroeconomic performance and on the degree to which policymakers in fact control this performance (as this influences their incentives to build up reputation for competence). There is an important caveat, though: “If ... the positive effects of good policies (or the negative effects of poor policies) do not become apparent in one term, the incentives to offer good policies are weak and outcomes are expected to be poor” (ibid.: 175). That is bad news for financial market regulation, public debt reduction and global warming.

Relatedly, Jennings (2011) suggests a model where voters can become emotionally attached to policies that do not lead to efficient outcomes. A combination of retrospective and expressive voting may then lead to an equilibrium where even good incumbents (i.e. those who are in fact motivated to implement efficient policies), who are aware of the fact that a given policy to which voters are emotionally attached is inefficient, nonetheless implement this policy in order to secure re-election. This outcome is particularly likely if those incumbents have a low time-preference and therefore a strong desire to stay in office. This model is yet another example where biased behavior by voters interacts with rational behavior by politicians.

Quattrone & Tversky (1988), in an earlier contribution to strong BPE, argue that retrospective voting – what they refer to as the “incumbency-oriented voting hypothesis” (ibid.: 723) – can also be explained by the insights of prospect theory into the asymmetries of risk preferences: If people are risk-averse for gains and risk-seeking for losses (relative to some pre-specified reference point), then “the less risky incumbent should fare better when conditions are good than when they are bad” (ibid.: 724). Moreover, the status quo bias should give the incumbent an extra advantage (ibid.: 725-26).

Frey & Eichenberger (1991) argue that the cognitive bias known as hindsight bias can help to explain why voters usually hold the incumbent accountable for bad economic outcomes, even if he is objectively not responsible for them. There is ample evidence that people, after learning about the true nature of events, overestimate the probability attached to that outcome; surprises and contingencies are neglected (e.g. Fischhoff 1980). As Frey & Eichenberger (1991: 39 On prospect theory, see Kahneman & Tversky (1979). Note that what counts is the fact that the challenger is here perceived as riskier, not better or worse, than the incumbent.
75) put it, “[i]f politics leads to unfavorable results, people wrongly believe that this was foreseeable. Therefore they blame government for having committed a grave mistake”.

Finally, Findley (2015) models retrospective voting with a behavioral twist and assumes that memories are subjected to hyperbolic discounting. This means that at first, memories fade away relatively quickly, while with more time, the rate of depreciation of memories decreases. He applies this model to the problem of political business cycles and shows how this kind of behavioral retrospective voting outperforms standard approaches in predicting different patterns of political business cycles.

2.5 Determinants of political preferences

Studies on the determinants of political preferences of voters are still rare, due perhaps to the complexity of factors that have to be examined. We discuss, in turn, contributions in the realm of weak BPE and in the realm of strong BPE.

As to the first, consider preferences for redistribution, one of the most important issues in the political debates of developed welfare states. It is still largely unclear what exactly determines the underlying individual preferences. While the existence of other-regarding preferences has been well established in laboratory experiments and in the field (e.g. Roth et al. 1991, Hoffman et al. 1996, Cameron 1999), it is an open question to what extent this translates into more realistic voting settings, i.e., whether the demand for redistribution is driven by other-regarding preferences or pure self-interest. Esarey et al. (2012) claim that the setting of previous studies, such as the ones by Scott et al. (2001) or Frohlich et al. (1987), is too far removed from real-world taxing and redistribution decisions, for those studies either involved extensive small-group discussions, or costless choices among alternative redistribution schemes. Esarey et al. (2012) design an experimental set-up where people’s choices on redistribution affect their own income. They find that “votes for a redistributive tax were almost entirely in accordance with self-interest” and that “a measure of subjects’ preferences for fairness or equality, their self-reported economic ideology, was unrelated to their voting behavior in this experiment” (ibid.: 621).

One may doubt, though, whether the Esarey et al-setting really represents “an environment similar to the one that people face in the outside world” (ibid.: 605). While the
possibility that subjects earned their economic position before voting certainly adds realism to the setting (compared to standard ultimatum or dictator games), the authors seem to neglect the “veil of insignificance” of votes in the real world: Given that voters have next to no influence on the actual political choice of a redistribution scheme, they might just as well act as if their choice was costless and use it to at least generate some expressive utility.

In contrast, strong BPE looks into the mental processes that give rise to (typically biased) voting behavior. A generally under-researched, yet recently burgeoning, field examines the origin of peoples’ political preferences. To illustrate, the question of which influence daughters have on their parents’ partisanship is hotly debated: While Oswald & Powdthavee (2006), relying on British household data, find that ceteris paribus, each additional daughter increases her parents’ inclination to vote for left-wing parties, Conley & Rauscher (2010) offer empirical evidence for the opposite effect. Washington (2007) argues that when it comes to legislators in the US, female offspring makes both US senators and members of the House more likely to support leftist positions. Obviously, the jury is still out here. A closely related literature examines the dynamics of children’s political preferences: Dinas (forthcoming), for instance, finds that those who are most likely to initially adopt their parents’ political views are also most likely to later abandon them as a result of their own political engagement.

Strong BPE shows, however, that even after socialization, political preferences often are far from settled. In a pioneering contribution to strong BPE, Quattrone & Tversky (1988) found that the decision whom to vote for is strongly influenced by framing effects, i.e., the way alternative options (policy programs, say) are framed. Framing effects violate the standard invariance principle in rational choice theory. The political supply side can easily use well-known biases, such as loss aversion or the status quo bias, in order to manipulate reference points that in turn shape the evaluation of alternatives. For instance, it seems to make a significant difference whether labor market policies are described as aiming at lower unemployment or higher employment (Chong & Druckman 2007). Interestingly, Kuenhanss et al. (2015) find that experts are as susceptible to framing effects in economic policy scenarios as the general public. In a recent study, Augenblick & Nicholson (2012) show that in direct-democratic decision-making, voters treat issues differently depending on their place on the ballot. This hints not only at framing effects, but also at a phenomenon called choice fatigue, where voters are constrained by their cognitive capacities when they have to deal with too many choices at once.
There are also contributions on the dynamics of political preferences over time. Kuran (1991) argues that the formation and subsequent evolution of political preferences is biased, due to individual susceptibility to the availability heuristic (Tversky & Kahneman 1974): Agents tend to pay serious attention only to those options that have previously been tried out, because information on those options is readily available in their minds.\textsuperscript{40} Kuran shows that the adjustment of private preferences in light of the biased and incomplete perception of past experiences may affect the evolution of publicly expressed preferences, which in turn may affect policy-making itself. As he concludes, “[t]his makes the whole evolutionary process inherently unpredictable, even in the absence of errors in the implementation of society’s collective choices” (Kuran 1991: 265).

Kuran & Sunstein (1999) present some detailed case studies where widespread preferences for – often very costly – policy interventions have evolved in response to risks that are very small from a scientific perspective. Applying a conventional cost-benefit analysis, these policy interventions would be judged to be inefficient in the sense that their explicit monetary costs outweigh expected benefits. But more importantly, they are inefficient in terms of opportunity costs: There would be other, more relevant risks that could be reduced instead through other policy interventions. Kuran & Sunstein explain this irrational policy-making with the interaction of availability cascades and reputation cascades. The former are based on the availability heuristic. In social communication, availability effects are reinforced in cascades, and public discussions of minor risks can lead voters to significantly overestimate them. Representatives and regulators respond to such unfounded concerns, despite knowing better, out of reputational concerns. These make it individually rational to be politically responsive even to clearly unreasonable risk assessments. Again, inefficient policies become a plausible outcome of biased citizens interacting with perfectly rational politicians.

This framework also encompasses the phenomenon of preference falsification, which has been introduced and discussed at length by Kuran (1995): Agents may find it rational to act publicly on preferences that differ from their true, privately held preferences. This is due to reputational concerns that may play out differently in different institutional settings. In autocratic regimes and dictatorships, citizens often have a straightforward incentive to conceal their true preferences.\footnote{\textsuperscript{40} Or certain past experiences may be particularly salient, such as hyperinflation in the 1920s for Germans (Frey & Eichenberger 1991: 75). For a general approach to the study of cognitively biased inference-making, see Gennaioli & Shleifer (2010).}
preferences. But as Noelle-Neumann (1993) has shown, a “spiral of silence”, where individuals do not voice minority opinions in fear of social ostracism, can be observed in democracies as well.

Biases can also arise from the exact opposite of the availability effect, namely from (individually rational) ignorance with regard to complex policy issues that are not conveyed in an easily digestible form in social communication. A prominent example is fiscal illusion, which comes in different forms (see Oates 1988 for a survey of the theoretical issues). Consider debt illusion, where voters systematically underestimate the true costs of public goods if they are financed through public debt, rather than through taxes (Buchanan & Wagner 1977). Another source of fiscal illusion is the complexity of the tax system, which can help to disguise the true overall tax burden (Buchanan 1967), and also make it difficult for individuals to ascertain the incidence of particular taxes. Empirical evidence generally supports the relationship between the complexity of revenue systems and the degree of fiscal illusion (Dollery & Worthington 1996). Recent empirical evidence that fails to corroborate Ricardian equivalence by showing that voters prefer debt over tax finance also lends some support to the existence of debt illusion (Banzhaf & Oates 2012; Dell’Anno & Dollery 2012).

In an interesting experiment, Tyran & Sausgruber (2011) confirm that individual voters are often unable to correctly account for the true distribution of tax burdens, which leads them to prefer inefficiently high tax rates. They also show that by playing multiple rounds, their subjects learn and de-bias with experience. However, it is doubtful whether such learning processes will also work outside of the laboratory, where opportunities to learn from policy changes and their results are few and far between, relative to the orderly playing of subsequent multiple rounds in an experiment. The second result may, unfortunately, be more relevant in practice: Communication between the subjects in the laboratory increases polarization in opinions between groups of individuals, and thereby also increases the likelihood that agents fail to account for the true costs of taxation. This nicely fits with our discussion above: If communication reinforces prior beliefs, and also reinforces group identities, then preferences may emerge that have no foundation in the objective conditions and restrictions of a political decision.

3 Policy-makers
In this section, we start with a brief diversion into normative territory in Subsection 3.1 and outline the discussion (so far only brief and sketchy) on what politicians should do under non-standard assumptions regarding individual preferences. Section 3.2 outlines the mostly empirical discussion on the efficacy of incentives and rewards, which has uncovered some surprising patterns of behavior. In Section 3.3, contributions are discussed which take a more evolutionary, process-oriented perspective on policy-making than standard models. Finally, Section 3.4 discusses the role of political entrepreneurs, the lesser-known Schumpeterian relatives of business entrepreneurs who, for example, disturb political equilibria through the introduction of new and innovative policies.

3.1 What should politicians do?

When it comes to the study of politicians, a key question is normative in nature: What should they actually do? Following John Neville Keynes’ (1917) influential classification, answering this question either refers to the “instrumental” branch of economics (“the art of political economy”, as Keynes puts it), or to the normative part of economics (“applied ethics”). While the former examines which policy measure is best suited to achieve a given goal, the latter examines alternative goals and criteria themselves. It seems to us that both branches are still dramatically under-researched at this point.

As to the instrumental branch, there are as yet very few attempts to derive policy implications from an explicitly “behavioral” model of how citizens respond to policies. Bolton and Ockenfels (2012) suggest an approach they call “behavioral economic engineering” that tries to integrate such ideas. Harstad and Selten (2015) also model citizens as responding to policies in a non-optimizing way: They are guided instead by “lower-dimensional” rules of thumb that are much easier to apply. The authors argue, quite convincingly, that policymakers should take these observations into account, as many important policy recommendations go awry when inadequately modeling citizens as homines oeconomici.41

41 Also to be included here is psychologically informed advice as to whether constitutions should be designed so as to contain “knavish” policy-makers (see subsection 3.2, below).
As to Keynes’ normative branch, it seems that a lot of it needs to be rebuilt from a behavioral perspective. For if we take the insights from behavioral economics seriously, then many established concepts and tools of standard normative (or “welfare”) economics can no longer be applied: If individuals can no longer be assumed to reveal coherent utility functions through their observable choices (e.g. due to framing effects or dynamic inconsistencies), a well-behaved social welfare function may no longer exist. More fundamentally, the standard preference-satisfaction account of welfare no longer qualifies as an appropriate normative benchmark (Hausman & McPherson 2009). That insight makes Hargreaves Heap (2013) conclude that even if behavioral economics were not to survive the methodological attacks from mainstream economists (such as Gul & Pesendorfer 2008), it would still have a lasting impact on normative economics, at least to the extent that normative economics is individualistic.

To illustrate the challenges welfare economists face here, consider the recent development of a behavioral welfare economics (e.g. Bernheim & Rangel 2009, Rubinstein & Salant 2012), which, when applied to policy-making, translates into “behavioral public economics” (Bernheim & Rangel 2007). Behavioral welfare economics aims at keeping the traditional definition of welfare – the satisfaction of coherent preferences – as far as possible under “behavioral” conditions. This means that empirically revealed preferences must be “reconstructed” (or rather “cleansed”) in a way that makes them coherent again. Happiness researchers have suggested subjective well-being as an alternative policy goal or, more ambitiously, as an empirical proxy for social welfare (e.g., Kahneman et al. 2007, Layard 2006).

Apart from a catalogue of normative worries (such as paternalistic tendencies), the well-known traditional Public Choice critique applies to both proposals: Given the intricacies of real-world political processes, we cannot expect welfare-maximization (in either its behavioral or hedonic variant) to be implemented. To illustrate, a policy goal such as subjective well-being can hardly be introduced without distorting its own empirical basis: rational agents would mend their answers to happiness questionnaires in a way that maximizes their personal interests (Frey & Stutzer 2012).

---

42 One might, of course, object that normative economics is not a legitimate part of economics (Gul & Pesendorfer 2007).
43 The use of happiness as a policy goal has been criticized from many different angles, see, e.g. Schubert (2012), Fumagalli (2013), Van der Rijt (2013) and Saint-Paul (2011).
Instead of hoping to maximize some measure of welfare, it appears more reasonable to search for constitutional rules and other institutions that frame the political process in an acceptable way. But again, what benchmark can be used to discern “acceptable” from unacceptable policy-making from a behavioral perspective? One option, advanced by Robert Sugden in an explicitly contractarian framework, could be to use the agents’ preferences, whether they be coherent or not (Sugden 2004, 2008, 2010), as the normative yardstick and to assess institutions as to the “opportunities” they provide to satisfy whatever preference one may happen to have. It turns out, though, that Sugden’s approach is based on questionable empirical assumptions. Anyway, there is a lot of exciting work to do in this area.

Long before the advent of behavioral economics, Public Choice had of course criticized traditional welfare economics and its “public finance” applications on different grounds, namely for its neglect of the actual political process. Welfare-maximization did not appear to be an empirically plausible outcome of real-world policymaking. But in practice, Public Choice has itself introduced normative criteria that do not stand the test of behavioral economics. Consider the median voter theorem (Black 1948; Downs 1951). When used as a predictive theorem, it requires strict assumptions on individual preferences and on the nature of the political problem to hold (Arrow 1951). But the median voter result has – often more implicitly than explicitly – also become a normative benchmark in Political Economy. Although the median voter outcome is not generally welfare-maximizing – since standard voting cannot transmit the intensity of preferences in the same way as the price mechanism can (Besley 2006: 72) – it is an attractive normative benchmark, because it is understood as a measure for the quality of a competitive, democratic political process. In this sense, it is prima facie an appealing normative criterion: A setting where politicians are perfectly disciplined – not by the voters, but at least by one of them, viz. the decisive voter. The median voter outcome thus plays a twofold role in Political Economy (see e.g. Persson and Tabellini 2000): On the one hand, it is a yardstick for the quality of political processes in terms of disciplining agents (politicians, bureaucrats); on the other hand, it also serves to analyze and predict political failure. The latter can result even with perfectly disciplined political agents: Collective decision-making may lead to economically inefficient results, depending for instance on the distribution of voter preferences, but also on institutional

---

44 See Schubert (forthcoming).
45 An overview on the general issues is given by McQuillin & Sugden (2012). See also Cordes & Schubert (2013).
or even cultural influences that shape voter incentives. Clearly, this normative benchmark relies on some standard neoclassical assumptions, in particular on individual voters having stable and technically well-behaved preferences. It seems necessary to redirect the focus of normative analysis and to look for policies that can actually be reconstructed as the result of voluntary political exchange (Buchanan 1987). While this is standard Public Choice reasoning, the crucial difference caused by the shift to BPE appears to be that along with the new emphasis on collective “irrationality”, political failure and preference variability also comes a new emphasis on political learning and political change. The adaptability of policy-making may become a normative criterion, as does the diversity of policies across jurisdictions, which gives sovereign citizens an opportunity to choose and policy-makers an opportunity to learn (Schnellenbach 2005). There are, thus, some first steps towards normative theorizing under behavioral assumptions, but it is still one of the clearly under-researched sub-fields of BPE.

3.2. Incentives and rewards

How to control self-interested representatives has always been one of the central issues in Political Economy. The relationship between voters and politicians is modeled as a principal-agent situation. Not only electoral competition, but also many other sophisticated incentive schemes have been proposed to induce the agents to act in the best interest of the principals, or to make sure that good candidates self-select into becoming candidates for public office.46 This line of research is clearly rooted in the neoclassical tradition, as evidenced in particular in the constitutional approach suggested by Laffont (2001), where the design of a constitution is understood as an exercise in mechanism design, aiming to produce an elegant incentive scheme for perfectly rational politicians. In most cases in traditional Public Choice, the incentives under examination are “hard”, i.e. they have some effect on the politicians’ monetary budget constraints.

It is a well-known empirical insight in behavioral economics that hard incentives of this type may have unintended consequences (Kamenica 2012). A popular example is the observation by Gneezy and Rustichini (2000) that a fine intended to discourage immoral behavior may be

46 See e.g. Besley (2007) for a contribution that provides a thorough introduction to modern rational-choice political economics.
interpreted as a fee to be paid for immoral behavior, and that individuals who abided by the rules before the introduction of the fine break the rules afterwards, gladly paying the fine as a price for doing so. Apparently, providing a monetary price for unwanted behavior crowds out a pre-existing intrinsic disposition to play by the rules. A similar problem may occur with positive bonus payments; in this case, the so-called hidden costs of reward (Lepper & Greene 1978) may become relevant: Paying individuals for something that they have done without external incentives before may crowd out their intrinsic motivation. They will henceforth reduce voluntary contributions, and their effort will depend on further external rewards. An explanation for this effect may be that the reward also contains a signal to the agent (Bénabou & Tirole 2003). For example, the agent may interpret the reward as telling him that the principal actually expects less of him, and adjust his effort accordingly. In a different context, Frey (1997b) argues that a feeling of being externally controlled may permanently alter the preferences of the agent and make him less cooperative.

The empirical literature on motivational crowding in political contexts focuses to a large extent on the incentives of citizens set by politicians or bureaucrats. For example, Frey (1997a) empirically shows that taxpayers tend to have higher tax morale in polities where they are treated with respect by the tax authorities, and where they have more direct-democratic control over policy. Rules that are based on the assumption that every taxpayer is a potential criminal in need of tight control are associated with lower tax morale. Experimental evidence from the laboratory points in a similar direction (e.g. Tyran & Feld 2002). Thus, any policy proposal needs to take into account that setting hard incentives may backfire through behavioral adaptations of citizens.

The perspective can also be turned around, though: Do incentives provided for politicians and bureaucrats lead to motivational crowding effects? Frey (1997a) believes that drafting a constitution for “knaves” will crowd out intrinsic motivation of politicians, but presents no direct evidence in this respect. There is some evidence showing that positive incentives are associated with the desired positive effects on performance. Kotakorpi and Poutvaara (2011) show that a salary increase in Finland drew more skilled representatives into parliament, but have no evidence on actual performance. For Italy, on the other hand, Gagliarducci & Nannicini (2013)

47 See also the recent survey by Alm (2012) for a broad overview on the empirical literature on tax evasion and its determinants, which generally supports the theoretical relationships discussed here.
find evidence that higher paid politicians are both more educated and more efficient. They attribute this to a selection effect, rather than an on-the-job incentive effect.

Selecting not only more skilled, but also more cooperative-minded, less “knavish” individuals into public office can be seen as generally preferable to the tedious task of controlling knaves once they made it into office (Brennan 1996). Mansbridge (2009) points out that the selection approach has three components: (i) self-motivated agents, who (ii) pursue the same objectives as the decisive voter, and (iii) a selection mechanism that carries these individuals, and only them, into public office. Galasso and Nannicini (2015) discuss the efficacy of formal electoral systems in performing the task of selecting skilled politicians. But beyond political systems, there is also evidence that tentatively points into the direction that positive incentives are useful in selecting good politicians, where positive incentives do not necessarily need to be monetary, but can also be of a pure reputational value, such as awards (Frey & Neckermann 2008). However, a constitution that disregards knavish behavior is still not preferable: Minorities, who are not part of the winning coalition, need protection from an agent whose objectives are closely aligned with those of the winning coalition.

In a different context, Choi et al. (2009) show empirically that the quality of work provided by American judges does not increase with salary. This is interesting insofar as judges undergo a careful selection process, but judicial independence leaves very limited mechanisms to control or discipline judges once they are in office. In our broader context, this result may hint at the necessity to look in greater detail at the interaction of selection and control, and the possibility that selection through positive incentives is less effective if it is not complemented by any meaningful sanctioning mechanisms. Recent theoretical work indeed shows that the interaction of rewards and sanctions is complex and its effects depend on further auxiliary conditions (Sasaki et al. 2012).

3.3. The process of policy-making with beliefs and ideologies

---

48 Corazzini et al. (2013) however show in an experiment that the breaking of campaign promises appears to be associated with psychological costs for the politician. In this sense, a psychological contract between voters and politicians could under some circumstances lead to the implementation of desired policies, without an external sanctioning mechanisms.
We have already discussed, in Section 2, how the specific assumptions of BPE on voter behavior lead to a different understanding of policy-making, relative to conventional models that assume rationally ignorant voters casting purely instrumental ballots. In particular, we have seen that problems may arise if voters decide behind a veil of insignificance, and representatives passively adjust their policies to expressive voting. In the resulting expressive policy traps, policies may be implemented with the backing of an electoral majority that eventually will have adverse effects on the welfare of the very people who voted for them (Hillman 2010). This is, however, only a first step towards a behavioral economic analysis of policy-making.

There are BPE approaches that adopt a more process-oriented approach and take a closer look at the origins of policy preferences, in order to see how they are translated into actual policies. Dixit (1996), with his transaction-cost approach, can be seen as one of the pioneers in this area.49 Positive transaction costs in politics imply that constitutional rules, understood as the result of a constitutional contract, will always be incomplete. They always leave room for interpretation, and even for change of the perceived meaning of the constitution without changing the text (Voigt 1999). The framing and also the re-interpretation, and in some cases the explicit re-negotiation of a constitution all depend on the relative bargaining power of groups in society, but they also depend on the beliefs concerning the properties of alternative sets of rules that prevail in a society at any given time (North 1990). On the one hand, this leads to a certain hysteresis in policy-making. Beliefs and bargaining powers are often quite persistent, and therefore inefficient policies can remain in place for a long time, if they still have the support of constituencies who benefit from them, and if side-payments to compensate these constituencies cannot be organized due to political transaction costs. On the other hand, today’s policies exert an influence on tomorrow’s beliefs and bargaining powers. Policy-making becomes path-dependent and real (i.e., irreversible) time starts to matter. History enters the economic analysis of politics, and so do many beliefs, goals and values that people hold apart from economic efficiency, and that influence policy (Dixit 1996: 144-156; North 2010; Boettke et al. 2013).

Bednar et al. (2015) introduce behavioral spillovers as a novel micro-foundation for path-dependence in the evolution of institutions. They model these spillovers between individuals, within communities, and show that the specific way in which individuals in communities learn

---

49 See also North (1993) for an earlier, programmatic paper suggesting the development of a transaction-cost oriented Public Choice.
how to play games influences the degree of efficiency of different institutions. There are complementarities between specific institutions and specific learning dynamics, and the latter could be broadly interpreted as the different cultures that establish themselves in different communities.50

Other approaches extend the process-oriented perspective. For example, Meier & Slembeck (1997) argue that a theory of economic policy-making should take into account how politicians and citizens interpret political problems. According to their model, social communication constructs the beliefs in which concrete policy-measures are rooted.51 It is then rational for some agents, such as representatives and lobbyists, to manipulate processes of communication and interpretation in their favor, so that rationally ignorant voters are influenced accordingly. A theory of economic policy would then need to be open to explanatory approaches from social psychology and sociology, which help to understand how common beliefs are formed and dispersed (hence, moving the theory from weak to strong BPE). There is a similarity between this approach and that of expressive voting, in the sense that social communication matters. But it goes a step further in emphasizing a possible active role of politicians. The politician not only faces a new constraint that restricts his scope of action, but he can also actively attempt to influence collective beliefs and thereby manage this new constraint.52 We will return to this issue when discussing the phenomenon of political entrepreneurship below. Interestingly, the collective beliefs that constrain politicians differ widely across countries (Bjørnskov & Paldam 2012). The influence of beliefs on policy has also entered more traditional political economy analyses. An example is Alesina & Angeletos (2005), who argue that beliefs about the fairness of a market order influence redistributive policies.53 Recent empirical research also indicates that the ideological affiliation of governments does systematically predict policy outcomes, at least in some areas (e.g. Bjørnskov and Potrafke 2013; Belke and Potrafke 2012).

50 See also Zhang & Wang (2015), who apply the idea of institutional path-dependence to China and identify the roots of path-dependence in the interactions of local governments and businesses.
51 See also Slembeck (1997).
52 A distantly related approach, although far deeper rooted in a conventional rational choice paradigm, is the impact of issue salience on policy-making. For example, Roemer et al. (2006) demonstrate how individuals vote against their own material interests in distributive politics because their strong preference for xenophobe policies dominates their distributive concerns. Constellations of this type also give policy-makers an opportunity to strategically work with voters’ preferences on different policy issues in implementing their own preferred policies.
53 As Leong (2015) demonstrates, fairness perceptions seem to depend, among other things, on the thickness of the veil of ignorance behind which agents may be placed.
Empirical evidence already suggests that private media serve as an arena where preferences and beliefs of individuals are influenced.\textsuperscript{54} It should be noted at this point that some proponents of a behavioral welfare economics have even suggested that individual preferences become a lever for active policy-making. For example, Layard (2007: 22-23) argues in favor of a purposeful sculpting of individual preferences. Public education, he argues, ought to make use of the results of happiness research and manipulate beliefs of individuals such that it will subsequently be easier for them to lead happy lives. From a Political Economy perspective, advocating such re-education policies, even if done in good faith, appears to be a dangerous approach. Gentzkow and Shapiro (2004) give an example for the severe biases and misinformation that may result if individual perceptions are purposefully engineered by politicians and state-controlled media through propaganda policies.

On the level of the individual politician, Ostrom (1998) argues that we cannot expect fully rational actions by policy-makers themselves. Instead of carefully evaluating all possible alternatives, they will typically use heuristics and follow rules of thumb. In turn, they may themselves be susceptible to the influence of “behavioral” influences (Kuehnhanss et al. 2015). While in general, the use of heuristics appears to be a quite efficient strategy and not, as one might think, associated with bad decisions due to information left unused (Gigerenzer & Gaissmaier 2011), matters may again be different in politics, where rules of thumb tend to be related to stable ideologies, which may not offer very precise guidance to solving policy problems. Baron & McCaffery (2008) give an example: Among American conservatives, “starve the beast” is a popular rule of thumb, i.e. the belief that tax cuts will automatically be followed by spending cuts, thereby reducing government activity. However, even those who want to “starve the beast” often find it difficult to identify actual spending positions that they would approve to be cut. Experience shows that tax cuts without immediate spending cuts tend to result in increased public debt rather than reduced public spending, yet the “starve the beast” metaphor still guides the decisions of many politicians and voters. It is rational for politicians to use it as an ideological signal towards their constituencies.\textsuperscript{55}

Another issue on the level of the individual politician is the susceptibility for emotional and affective behavior. Van Winden (2015) gives a broad overview over the effects of such


\textsuperscript{55} Osterloh & Heinemann (2013) also present survey evidence for politicians, which indicates that they individually have relatively coherent belief systems, which however differ between ideological camps.
behavior in different games that are played. He shows, for example, how expecting a feeling of anger in a loser of a contest for rents will reduce the value of winning this contest, and thereby reduce the problem of rent over-dissipation. On the other hand, fear of humiliation will have the opposite effect and lead to an inefficient increase of investments into the contest. Emotions in politicians can therefore have both beneficial and problematic effects from a social point of view.

Psychological research has also proposed explanations for the attachment of individuals to certain ideologies, broadly defined as sets of beliefs that guide their political decisions. For example, Kahneman and Renshon (2009) pick seven from the large smorgasbord of uncovered cognitive biases and argue that these seven biases are responsible for hawkish, trigger-happy behavior, in particular (but not only) in foreign policy. Recently, Stern et al. (2014) have empirically associated political ideologies with beliefs of uniqueness. While conservatives overestimate how people with similar political preferences as themselves also share roughly the same non-political preferences, left-wingers overestimate how unique they are in their personal, non-political preferences. This suggests that conservatives are relatively conformist and suffer from a false consensus bias, while leftists suffer from a strong overestimation of their own uniqueness. There is also a branch of the literature that relates political ideology to other personal characteristics, such as confidence in one’s own abilities. While much of the earlier literature has argued that conservatives tend to be more overconfident in their abilities than leftists, Snowberg & Ortoleva (2015), based on a formal model and the analysis of variation in time of survey data, find that there is no systematic relationship between overconfidence and conservatism.

To sum up, the BPE approach to policy-making is significantly broader than that of traditional Public Choice analysis, especially due to its interpretation of policy-making as a path-dependent process that tends not to be based on rational expectations. This openness does, however, come at a cost: ambiguity. It is difficult to come to definite predictions if the underlying model is open in the sense that it allows for changes in preferences or beliefs. So far, a convenient workhorse model such as the median voter model of traditional Public Choice is lacking in BPE, as is, a fortiori, a unified theory of policy-making.

3.4. Political Entrepreneurs
The issue of political entrepreneurship, i.e. the introduction of new policies by innovative politicians who overcome social dilemmas, or political and institutional barriers to reform, can help to further illustrate the BPE approach. Simplifying somewhat, in traditional Public Choice, the implementation of a given policy is either the result of the decisive voter wanting it, or of self-interested politicians implementing it against the decisive voter’s will. It all depends on what is assumed on the efficacy of electoral control. Changes of policies may come about due to exogenous events, such as economic shocks that have an impact on the relative bargaining power of groups of voters, but there is little scope for endogenous change. With regard to reactions to changed conditions, the main puzzle is the fact that economically efficient reforms are often delayed, or even not occurring at all (e.g. Alesina & Drazen 1991). From a behavioral perspective, acknowledging cognitive biases, faulty information-processing, and other aspects of bounded rationality, it is not surprising that opportunities for policy reform lay idle and need to be discovered, that new coalitions need to be forged, which involves political transaction costs, and so on.

In such a framework, entrepreneurship does have a role to play also in politics. Accordingly, Wohlgemuth (2002) argues in favor of a “Schumpeterian” approach to politics, where the personal characteristics and skills of politicians matter. The focus is, to some degree, shifted from institutions and other objective restrictions to the capacity of specific politicians to propose, popularize and implement novel policies (Schumpeter 1942: 290). This new perspective, however, also gives rise to the question under which conditions representatives endowed which such skills self-select into candidacy for public office, and become elected (Bernecker & Gathmann 2013). It has also been discussed whether incentives such as those engendered by territorial political competition associated with fiscal federalism increase the propensity to introduce political innovations (e.g. Rose-Ackermann 1980; Strumpf 2002). Regarding the ability of politicians to innovate, Schumpeter (1942) himself, maybe under the impression of contemporary politics, believed that it would be relatively easy for skilled political entrepreneurs to manipulate public opinion and gain support for the introduction of political innovations. In contrast to this view, more recent contributions such as Kuran (1995), who describes “collective conservatism”, and Schnellenbach (2007) emphasize the self-stabilizing properties of publicly voiced opinions. In times of routine politics, when no severe crisis
invalidates status quo beliefs and values, large-scale political entrepreneurship becomes unlikely and small, incremental policy changes within a given political paradigm are predicted.\footnote{See also Kingdon (1995), who argues that politicians normally act not as entrepreneurs, but as gatekeepers, who allow only those issues onto the agenda that promise a relatively secure gain in popularity, without the risks associated with public entrepreneurship.}

There is some tentative evidence suggesting that the \textit{framing} of reform proposals matters. Traut-Mattausch et al. (2008) argue, based upon experimental evidence, that attitudes towards reform will be significantly different, depending on whether a reform proposal has been justified with expected improvements, or in a more negative form as the necessary introduction of new limitations. Again, conflicting representatives and interest groups will rationally engage in attempts to frame issues such that voters perceive them in a way that is compatible with the interests of representatives and interest groups. Analyzing international survey data, Heinemann \& Tanz (2008) find that general trust – i.e., the general belief that other people will by and large act cooperatively rather than opportunistically – is positively associated with the quality of economic institutions, which they interpret as evidence for a positive impact of trust on the willingness to perform efficiency-enhancing reforms. This argument is corroborated by Bjørnskov \& Méon (2013). Thus, “soft” or behavioral conditions, such as the entrepreneurial skills of politicians and a general sentiment of trust on the side of the voters appear to influence the success of political innovations.

At this stage, this research on the political effects of trust by and large belongs to “weak” BPE: It is assumed that trust shapes individuals’ perceptions of cost and benefits of political measures, and that individuals then rationally base their choices upon these perceptions. The step towards “strong” BPE would in this case require a causal explanation on the sources of the specific type of trust that is relevant for the success of reform policies.

It should also be noted that traditional Political Economy caveats still apply. In the literature reviewed here, political entrepreneurship and policy reform are usually positively connotated; Heinemann \& Tanz (2008) even measure willingness to reform indirectly with the quality of institutions. This is an optimistically biased view. The ability to convince citizens that new policies are preferable clearly is a capacity that can be used in a purely self-interested fashion, and against the enlightened interests of citizens. It would therefore be premature to argue, for example, for a dismantling of formal checks and balances to facilitate political entrepreneurship (Schnellenbach 2007). On the marketplace, private consumers can easily opt
out and refuse to buy innovative products that they do not prefer. Entrepreneurial policies, by contrast, are usually collectively enforced through government power, which should require a more cautious evaluation.\footnote{For example, it may be tempting for politicians to evoke affective reactions by voters (Van Winden 2007), and even to evoke feelings of hatred to specific groups (Glaeser 2005).}

There are, however, exceptions. Elinor Ostrom (1990, 2005, 2012) has throughout her career analyzed the possibility to overcome social dilemmas through self-organization, and through the spontaneous establishment of institutions that help to stabilize cooperative behavior. This stabilization through institutions is necessary because even under conditions that favor initial cooperation, a collapse of cooperative behavior due to increasing free-riding in time can be theoretically predicted and empirically observed (Fischbacher and Gächter 2010). This decentralized negotiating of cooperative institutions is political entrepreneurship from the bottom up, and therefore of a different type than the Schumpeterian political entrepreneurship with its emphasis on leadership from above.

4. **Bureaucrats, regulators, and lobbyists**

The choices of bureaucrats, regulators and lobbyists are to a large extent still uncharted territory. While there is a very broad literature on how bureaucrats should efficiently regulate the actions of individuals suffering from choice imperfections, so far there is very little research on the biases that regulators themselves may be subject to (see however Kuenhanss et al. 2015 and Tasic 2011 for a first survey). Our discussion on the availability bias in Section 3 (Kuran & Sunstein 1999) can offer a first hint: Politicians, as the principals of bureaucrats, cannot be counted on to give their agents efficient mandates for regulation. Guided by re-election concerns, they will choose to regulate those risks that are perceived to be particularly salient by the general public at a given point in time, which may direct regulatory resources away from other, objectively more pressing but less salient risks (Jolls et al. 1998).

Cooper & Kovacic (2012) argue that under these conditions, regulators face a trade-off between, on the one hand, maximizing social welfare, and, on the other hand, serving their career concerns by following the politicians’ demands. Cooper & Kovacic then examine whether the well-known biases uncovered by behavioral economics are likely to be found in bureaucratic...
decision-makers, and they generally believe this to be the case. The argument is that unlike firms in markets, bureaucracies lack regular and noticeable external feedback mechanisms to learn from mistakes.\textsuperscript{58} The authors suggest some institutional remedies that could alleviate the problems of bias in bureaucracies. An example is “de-biasing” by assigning a team within the bureaucracy to play the role of \textit{advocatus diaboli}, and thereby make sure to get all arguments on the table.\textsuperscript{59} Another proposal is to incentivize bureaucrats by making parts of their rewards dependent on long-term outcomes. This is however probably very difficult to implement in practice.

Furthermore, there is once again the risk that intrinsic motivation is crowded out by monetary rewards conditioned on long-term outcomes. Some empirical evidence suggests that intrinsically motivated individuals select themselves into jobs in the public sector, which are often paid less than comparable positions in the private sector (Georgellis et al. 2011). Charness et al. (2013) offer experimental evidence suggesting that if individuals in an organization receive information on relative performance, but are paid flat wages, they will invest in destructive effort and sabotage their internal competitors’ work. Thus, even non-monetary incentives may have adverse effects in bureaucratic organizations, which may also pose a problem to using awards as incentives, as proposed by Frey & Neckermann (2008). In general, the empirical literature on the effects of performance-related rewards in bureaucracies is still inconclusive (Hasnain et al. 2012), and a golden rule for proper incentives for bureaucrats is not available at this point.

A behavioral theory of bureaucracy, regulation and lobbying casts a much wider net than the neoclassical view, which has long been represented by the Stigler (1971) and Peltzman (1976) approaches associated with the Chicago School. There, special interest politics is merely a mechanism to organize wealth transfers, and competition between politicians ensures both that these transfers are organized efficiently, and that any idiosyncrasies of individual politicians or regulators, such as ideologies or beliefs, cancel each other out in the process of electoral competition. BPE, though, has no reason to assume that competition and feedback mechanisms in politics are sufficiently strong – i.e., even stronger than they are on markets – to weed out cognitive biases in politics and principal-agent problems within a bureaucracy (Glaeser 2006).

\textsuperscript{58} See also Glaeser (2006) for a similar argument with regard to government in general.
\textsuperscript{59} In a similar vein, Choi & Pritchard (2003) are skeptical about the capacities of a monopolistic regulator and would favor a competition of regulators.
As regards a behavioral approach to lobbying, Jolls et al. (1998) observe that, perhaps unexpectedly, many laws seem to implement general fairness perceptions, rather than cater to the preferences of special interest groups. This may come as a surprise because the traditional theory of interest groups (Olson 1965) would predict that small groups with specific interests are easier to organize, and exert more influence. However, the discussion in Section 2 suggests that voting behavior behind a veil of insignificance may lead people to abstract from their own material interests. General-interest policy-making on broad issues that concern all or most citizens may therefore well be oriented at general norms, rather than particular interests of small, decisive coalitions. Nevertheless, a lot of scope for special-interest policy-making will remain, using laws targeted at small groups.

One new facet of lobbying recognized by behavioral economists is the so-called “availability entrepreneur” (Jolls et al. 1998), who exploits the aforementioned availability bias. He identifies suitable current events, publicizes them, and connects them with the kind of policies he prefers. For example, the lobby of domestic car producers might publicize a production error in a batch of foreign cars in order to argue that these are inherently unsafe, and that foreign-made cars must undergo costly additional safety checks, which of course increases their relative price. There is some similarity here to the traditional approach of informational lobbying, where provision of information to the legislator can be used along with campaign contributions to influence the politician (e.g. Bennedsen & Feldmann 2006). But a crucial difference is that the behavioral approach explains why lobbying is so often not only an activity behind closed doors between legislator and lobbyist, but rather aimed at overtly influencing public opinion in general. Again, the behavioral approach turns out to be more process-oriented, with beliefs and preferences being endogenous and subject to change.

An important issue debated in political science is the fact that a large fraction of studies fail to detect significant success in lobbying activities. For example, Gerber (1999), in an analysis of referenda in the United States, fails to find a significant influence of resources invested by special interest groups to sway public opinion on referendum results. In general, surprisingly many empirical studies fail to detect a significant direct influence of lobbying on policy outcomes (see Lowery 2013 and the extensive literature cited therein). Hall & Deardorff (2006) survey a literature of empirically reported anomalies in lobbying behavior. Among them are the facts that lobbyists invest a lot of resources on legislators who already share the lobbyists’
policy preferences, and that legislators receive relatively low monetary contributions in exchange for their votes. The authors also argue that it is unlikely that lobbyists actually convince legislators on a regular basis by providing information, because representatives have access to many other sources of information.

What is the relationship between legislators and lobbyists then about? Hall & Deardorff (2006) argue that lobbying is a subsidy to selected legislators: Lobbyists provide manpower to legislators in order to “help” them with their work in parliament and in accumulating influence there. But the anomalies reported would also fit a different explanation. Van Winden (1999) argues that the real goal of lobbyists could be the maintenance of social ties with legislators, and to prevent a network of like-minded individuals from eroding. The stabilization of social networks in the political arena over longer periods of time, rather than direct persuasion of legislators, would then be the true objective of lobbying. This branch of the lobbying literature therefore qualifies as weak BPE. It works with standard neoclassical assumptions on individual behavior, but introduces new objectives into their utility functions that have been disregarded before.

5. Applying BPE: Two examples

5.1. Libertarian Paternalism

One result of recent attempts to draw policy implications from behavioral economics insights has become particularly influential both in academia and politics: “Libertarian Paternalism”, as popularized by Thaler & Sunstein (2008), has also had an impact on the new behavioral welfare economics of, e.g., Bernheim & Rangel (2009). The starting point is the assumption that cognitive biases, some of which we have discussed in this paper, are symptoms of defective decision-making by individuals. Strict neoclassical “rationality” is maintained as a normative benchmark, which may seem surprising: While behavioral economists have invested a lot of effort into a broad empirical critique of homo economicus, the proponents of libertarian paternalism still believe that individuals should ideally act as if they were perfectly rational in

---

60 See also Sunstein & Thaler (2003a, 2003b) and Camerer et al. (2003) for early theoretical foundations that ignited the debate on soft paternalism.
the traditional sense. Furthermore, they believe that “choice architects” should guide individuals into this direction, in most cases by deliberately designing choice situations such that an individual is “nudged” into acting in accordance with the full rationality approach.

An example for libertarian paternalism is the incentive scheme devised by Thaler & Benartzi (2004). Individuals are believed to suffer from a tendency to procrastinate – formally modeled as hyperbolic discounting (Laibson 1997) – and often regret that they have saved too little once they reach pension age. The suggested “soft” incentive scheme involves allowing employees to commit not to immediate payments, but to payments in the not too distant future (to circumvent the procrastination), and also to have savings automatically increase with wages. While this by itself may sound like traditional paternalism, it is in fact supposed to be libertarian because freedom of choice is preserved: the employee can opt out of this plan at any time. He is nudged, rather than coerced. Other proposed policies of this type follow different goals; purposeful choice architectures are conceivable with regard to many behavioral anomalies.

Clearly, the approach belongs to behavioral welfare economics, rather than positive BPE. Moreover, it does not address the individuals who are believed to suffer from choice defects, but a fictitious benevolent planner, who supposedly has no preferences or objectives for himself (Sugden 2011). One can doubt whether a paternalist planner will be able to gather all the information necessary to intervene efficiently. If an individual has incoherent preferences, e.g. conflicting long-term and short-term preferences, then it is far from obvious how the individual’s conflicting preferences are to be weighed and aggregated. This would require a cost-benefit analysis, but one using subjective valuations that are not observable to any paternalist planner (Sugden 2008). And even if the information were readily available, one could still doubt that it is legitimate to nudge her without her own explicit consent. Note that the possibility to influence an agent’s choices while preserving his freedom of choice critically depends on the ability to exploit her cognitive biases (as, e.g., myopia in the scheme sketched above). Many critics have argued that this amounts to objectionable paternalism. In other words, it can be argued that – despite people being formally permitted to opt out of paternalist choice architectures – there is a danger that nudging, rather than working with transparent regulations, involves the manipulation of individual choices. Addressing the issue to the individual herself instead of a paternalist planner would alleviate this problem: The individual’s own capacity to deal with choice inconsistencies

---

would be strengthened, which may be preferable to being the object of inconspicuous manipulation. (Sugden 2011; Dalton & Ghosal 2011).

From a BPE perspective, though, the most important problem is this: Libertarian paternalists do not ask whether we can expect real politicians acting in real political processes to actually implement the policies that are deemed efficient. One issue here is that politicians also suffer from making biased choices. Glaeser (2006) argues with regard to libertarian paternalism that there is no reason to expect politicians or bureaucrats to be able or willing to nudge individuals into the direction of the *homo economicus* benchmark, even if they had all the information needed. In particular, paternalist planners themselves may suffer from their own availability bias, and focus on issues that are particularly popular at a given point in time (Schnellenbach 2012).

The case of Libertarian Paternalism therefore illustrates the difference between BPE and behavioral welfare economics, which is very akin to the old conflict between Political Economy and traditional welfare economics. Behavioral welfare economists are currently at risk of repeating the mistake of neglecting the real-world political process with its many intricacies. Under these conditions, policy advice addressed to an imaginary social planner may not only be useless, but even dangerous, if it helps to promote policies that have unintended, negative consequences under real-world conditions.

5.2. The Long-Term Effects of the Welfare State

The welfare state, like all private and public insurance mechanisms, suffers from the problem of inducing moral hazard. Knowing that transfer mechanisms are installed to absorb negative income shocks, individuals may react by adjusting their behavior. They might be more willing to take risks in their occupational choices, they may be less motivated on the job if they are less concerned about the danger of losing it, or they might extend the period of job search and voluntarily live off transfers. Recently, a literature has surfaced that analyzes the problem of behavioral adjustments to the welfare state. Already Lindbeck (1995) and Lindbeck et al. (1999) argued that while traditionally social norms have often repelled individuals from applying for transfers (even if they were legally entitled to them), this social norm of living off one’s own means as long as possible may be instable. The authors argue that a reason for this development
may be the way the welfare state is framed. Increasingly, politicians have come to interpret the reception of transfers as basic citizen’s rights, rather than as a stigma. With a changed framing of policies, the social norm evaporates. Lindbeck & Nyberg (2006) extend the approach by arguing that the existence of a welfare state will also negatively affect the propensity of parents to teach their children norms that reflect a strong work ethic.

If the true reasons for an individual becoming a transfer recipient are unobservable, then social norms that discourage people from claiming benefits play an important role. Such norms may even have been a precondition for large-scale transfer schemes to have been considered feasible in the first place (Algan & Cahuc 2009). Heinemann (2008), using data from the World Values Survey, finds empirical evidence in support of a deterioration of social norms by the welfare state: Both higher transfers and waves of high unemployment in macroeconomic downturns result in an increased willingness to claim transfers, even if one is legally not eligible to do so. While the first effect could be interpreted as a direct price effect, the second indicates the deterioration of the social norm when people get more used to claiming benefits.62

If the negative effect of the welfare state on work ethics will be corroborated in future empirical studies, then this would be an interesting relationship from the perspective of BPE. On the one hand, it may serve as evidence for the relevance of changing individual preferences, in response to changing social norms. It also indicates that endogenous social norms are an important object of analysis for a behavioral approach to political economy. If the original assertion from Lindbeck (1995) is correct, then this is also an interesting case for the study of framing effects in politics: The degree of the deterioration of social norms may depend on the way access to the welfare state is rhetorically framed by politicians. And finally, there may also be the question whether politicians have an incentive to strategically influence social norms, if they are able to do so.

6. Conclusion and outlook

62 Halla et al. (2010) also find a negative long-term effect of the generosity of the welfare state on benefit morale. Corneo (2012), on the other hand, looks at actual work norms rather than benefit morale, and does not find a robust relationship. However, Arnold (2013) shows that there is a negative link between self-reported benefit morale and the frequency of taking sick leave. Thus, using benefit morale in empirical studies is informative for actual behavior.
Our discussion of the emerging field of Behavioral Political Economy has shown that behavioral approaches become more and more commonplace in the economic analysis of politics. This trend is to be distinguished from the inspiration the early Public Choice literature provided for experimental economics (a subfield we did not cover here). The observation that Political Economy is, relative to other fields in economics, a late bloomer in this respect could be explained by the fact that the original impetus of economists analyzing political behavior was to export rational choice analysis into areas of non-market interactions. The initial skepticism towards behavioral approaches may, thus, have been more pronounced in this research field than elsewhere in economics.

We have also seen that what makes political behavior a particularly suitable candidate for applying insights from psychology is the fact that it exhibits incentive structures that differ markedly from those prevailing in the marketplace: Behind the veil of insignificance, people are essentially free to pursue any kind of non-standard goals. This, of course, also complicates the construction of general models that are able to explain empirically observed behavioral patterns.

In this respect, it is also interesting to note that BPE models still often display an asymmetry with regard to their basic assumptions, when, for example, “behavioral” voters are modeled as interacting with perfectly rational policy-makers or lobbyists. There is a danger here to introduce a new dichotomy in behavioral assumptions without much concern for the empirical evidence.

Our paper has also briefly addressed the issue of policy implications. Generally, with the emergence of “political economics” (Persson & Tabellini 2000), policy advice seems to play a less salient role than it used to (Jennings & McLean 2008). Having said that, the current trend in Behavioral Welfare Economics (or Behavioral Public Economics) is to use empirical evidence on individual choice imperfections as a starting point for designing policies that can correct such imperfections. It seems tempting to argue that the behavioral perspective adds a third kind of “failure” to the set: Besides market failure (a focus of old-fashioned public economics) and government failure (the specialty of Public Choice), there may now be “voters’ failure” when voters fail to act according to their interests (Jennings & McLean 2008: 73). That may be misleading, though. BPE relates to the behavioral welfare economics just like orthodox Political Economy relates to standard welfare economics: BPE emphasizes problems in the process of
policy-making that cast serious doubt on the capacity of government to accomplish systematic improvements of seemingly imperfect individual decisions.

As a final remark, some methodological worries should be mentioned that concern behavioral economics more generally. For many of the cognitive biases discussed in this paper, alternative interpretations are possible. Cason & Plott (2014), for instance, argue that a failure of subjects to recognize the correct game form leads to behavioral patterns that could also be interpreted as evidence for non-standard preferences. They also criticize framing theory for its lack of falsifiability: Standard choice theory allows us an unambiguous prediction of rational behavior, and systematic deviations can be used to reject either the hypothesis that individuals are able to correctly perceive the situation they are in, or the hypothesis that they have standard preferences. The predictions of framing theory, on the other hand, are not robust to changes in the experimental framework. If, for example, subjects are trained to correctly recognize the game they are playing, then this changes the choice framework, and changes in behavior can be attributed to the new framing of a given choice. But a clear distinction between mistaken individuals and those that simply have non-standard (frame-dependent) preferences cannot be made.

This example indicates that the interpretation of at least some of the empirical evidence that motivates Behavioral Political Economy may still be subject to change. However, we have also seen that the emerging field does not rest only on a few, specific biases or on one concept of non-standard preferences. The indisputable fact that the political arena features its own peculiar set of incentive structures makes sure that Behavioral Political Economy is here to stay and can safely be predicted to diversify further (preferably leaning toward strong BPE) in the foreseeable future.

**Acknowledgments**

We are indebted to three anonymous referees for their very careful reading of the paper and their constructive comments. We also thank the participants of the 2014 CESifo Venice Summer Institute Workshop on Behavioral Political Economy for their discussion and comments, in particular T. Scott Findley, Ron Harstad, Arye Hillman, Colin Jennings, Rebecca Morton, Charles Plott, Erik Snowberg, Heiner Ursprung and Frans van Winden.
References


