

Chapter 1

Rawls, Buchanan, and the Search for a Better Social Contract

Brian Kogelmann

Because he was such an accomplished economist, it can easily be forgotten that James M. Buchanan was something of a great philosopher as well. In his important essay on the social contract tradition, Samuel Freeman places Buchanan in the company of contemporaries such as John Rawls, David Gauthier, and Thomas Scanlon, as well as giants from the history of thought such as Hobbes and Rousseau (1990 [2007], 17).¹ This is largely due to Buchanan's *The Limits of Liberty* (1975 [2000]), though Buchanan's foray into the social contract tradition really began in his coauthored masterpiece with Gordon Tullock, *The Calculus of Consent* (1959 [1999]). Though many interpret this work strictly as a piece of positive social science, Buchanan always maintained that *The Calculus of Consent* was primarily a normative project, part of the broader social contract tradition (Thrasher and Gaus, 2017).

Though often included among the greats, Buchanan's contributions to social contract theory are often either hastily dismissed or deeply misunderstood. The purpose of this paper is to show that such dismissals and misunderstandings are to the detriment of the social contract tradition as a whole. More precisely, many of the shortcomings found in contemporary approaches to the social contract—particularly the approach pursued by Rawls and his followers—can be remedied by adopting specific features of Buchanan's model. This does not mean that Buchanan's approach to the social contract is without flaw itself. As we shall see, Buchanan's model has its own unique problems. But in patching these holes while keeping those elements that circumnavigate the problems associated with other theories, we end with a vastly superior model that better achieves the philosophical goals of the social contract enterprise. We end up with a new and improved social contract.

The structure of this chapter is as follows. The first section examines Buchanan's forays into the social contract tradition, both highlighting a

tension in Buchanan's thought as well as pinning down the interpretation of Buchanan to be used throughout this paper. The second section examines the goals of the social contract tradition, introducing two success conditions any plausible social contract theory must satisfy. Section three turns to Rawls's famous contribution to the social contract, arguing that it fails to satisfy one of the two desiderata introduced in the prior section. Section four examines one version of Buchanan's theory of the social contract, arguing that it succeeds across the desideratum Rawls's theory failed across, yet fails across the desideratum Rawls's theory succeeds across. Section five pursues a synthesis of Rawls and Buchanan, showing how the strengths of the two theories can be combined to produce a better theory of the social contract that succeeds where the prior two theories failed. There is a concluding section.

BUCHANAN'S FORAYS INTO THE SOCIAL CONTRACT TRADITION

Buchanan's forays into the social contract tradition spanned the course of his career and are, at times, inconsistent. In focusing on Buchanan's theory of the social contract we must thus clearly specify *which* Buchanan's theory of the social contract we are talking about. As a philosophical project, the social contract is, at its core, a normative enterprise: it seeks to give us leverage to critique our current practices while also highlighting arrangements we can justifiably move toward if our current practices are in disrepute. But in being a social scientist, it is not obvious Buchanan is engaged in such normative theorizing. Indeed, this is something Buchanan himself struggled with, and is, I think, a point of tension between Buchanan's and Tullock's understandings of their shared work.

In the opening pages of the *The Calculus of Consent*, Buchanan and Tullock write,

Political theory has concerned itself with the question: What is the State? Political philosophy has extended this to: What ought the State be? Political "science" has asked: How is the State organized?

None of these questions will be answered here. We are not directly interested in what *the* State or *a* State actually is, but propose to define quite specifically, yet quite briefly, what *we think* a State ought to be. (Buchanan and Tullock [1962] 2004, 3; emphasis original)

Now detailing what they think the state *ought* to be certainly seems like a normative project, but it is not clear that Buchanan and Tullock agree on this point. In their separately written appendices to *The Calculus of Consent*, Buchanan and Tullock each describe how their jointly authored contribution

fits within broader schools of scholarly enquiry. Tullock primarily sees theoretical forerunners to their project in the then-developing game theory and social choice literature, citing the work of John von Neumann and Oscar Morgenstern, Duncan Black, and Kenneth Arrow (Buchanan and Tullock [1962] 2004, 310–31). On this understanding, *The Calculus of Consent* is decidedly a work of positive social science, not a contribution to normative social contract theory.

Buchanan's appendix, however, places the piece squarely in the social contract tradition and explicitly as a normative project. Buchanan begins by noting that the science of politics contains both a positive as well as a normative component:

A positive science of politics should analyze the operation of an existing, or postulated, set of rules for collective decision-making quite independent of the efficacy of this set in furthering or in promoting certain "social goals." A normative theory of politics should, by contrast, array the alternative sets of rules in accordance with their predicted efficiency in producing certain ends or goals which should be, if possible, made quite explicit. (Buchanan and Tullock [1962] 2004, 292)

Importantly, Buchanan defines the core question of *The Calculus of Consent* as lying within the normative side of the science of politics: "What set of rules *should* the fully rational individual, motivated primarily by his own self-interest, seek to achieve if he recognizes that the approval of such rules must embody mutual agreement among his fellows?" (Buchanan and Tullock [1962] 2004, 296; emphasis original). This normative understanding of *The Calculus of Consent* is made even more prominent when Buchanan argues that the social contract is definitely not an empirical or descriptive project:

The relevance of the contract theory must lie, however, not in explanation of the origins of government, but in its potential aid in perfecting existing institutions of government . . . The origins of civil government and the major influences in its development may be almost wholly nonrational . . . if the purpose of investigation is solely that of explaining such growth, there is perhaps little purpose in inventing anything like the contractual apparatus. (Buchanan and Tullock [1962] 2004, 304)

This normative understanding of *The Calculus of Consent*, as well as the social contract tradition more generally, is later completely disavowed in Buchanan's *The Limits of Liberty* (1975 [2000]), published roughly ten years later. Here Buchanan argues—*contra* his earlier remarks—that "the origins of the state can be derived from an individualistic calculus . . . as we know from the writings of Thomas Hobbes as well as earlier contractarians"

(Buchanan [1975] 2000, 9). In reflecting on *The Calculus of Consent*, Buchanan notes that “the framework for analysis was necessarily contractarian, in that we tried to *explain* the emergence of observed institutions” (Buchanan [1975] 2000, 10; emphasis added). And in his description of the project pursued in *The Limits of Liberty*, Buchanan says, “In this book . . . existing and potential institutions as well as behavior within certain institutional constraints are *explained* in terms of failures of potentially viable contractual agreements to be made or, if made, to be respected and/or enforced” (Buchanan [1975] 2000, 11; emphasis added). As suggested in the introduction, much of the attention Buchanan receives as a social contract theorist in the philosophical literature comes from his work in *The Limits of Liberty*, not from his joint work in *The Calculus of Consent* (e.g., Gauthier 1986, ch. 7; Freeman [1990] 2007).

Even these remarks would not be Buchanan’s final thoughts on the *The Calculus of Consent* or the broader social contract tradition. In a piece written in the 1980s, Buchanan shifts back to a more normative understanding of this work: “*The Calculus* differed from the precursory works in one fundamental respect, namely, it embodied a *justificatory* argument” (Buchanan [1987] 2001, 70; emphasis original). So Buchanan’s understanding of social contract theory and his contributions to it as an essentially normative enterprise are not only in dispute, but in dispute within Buchanan’s own understanding of himself.² This chapter interprets Buchanan’s contributions to social contract theory as primarily a normative project, working on those questions social contract theorists are traditionally understood to be working on. Not only this, but this chapter takes Buchanan’s contributions in *The Calculus of Consent* as its core focus, contra most of the philosophical literature that primarily addresses *The Limits of Liberty*.

THE GOALS OF SOCIAL CONTRACT THEORY

As we navigate the social world we are confronted with rules and institutions that place restrictions on what we may do. Intuitively, some of these restrictions seem justified: though they may prevent us from doing certain things we would like to do at a particular point in time, they, overall, help us live better together. Also intuitively, some of these restrictions on what we may do are not justified: they are pernicious and harmful, making some worse off strictly for the benefit of others. As beings who can think about and act for moral reasons, it is natural for us to ask which rules and institutions are justified and which rules and institutions are not justified.

Social contract theory is but one way of addressing these difficult questions. Very roughly, rules and institutions are justified when they would be

the object of a hypothetical agreement of suitably modeled agents in a suitably modeled choice situation. Note that right from the get-go we stipulate that social contract theory is not concerned with actual consent by actual persons—what Gerald Gaus calls “justificatory populism” (Gaus 1996, 130–31). Rather, social contract theory—to borrow Rawls’s phrasing—is a philosophical tool we use that allows persons to ask “whether any of them has a legitimate complaint against their established institutions” (Rawls [1958] 1999, 53). If our current rules and institutions are the object of a suitably modeled agreement, then we have no legitimate complaint against them. If our current rules and institutions are not the object of a suitably modeled agreement, then there is something wrong with our current institutional order—we can justly complain about the restrictions placed on us, and now have some moral basis for demanding they change. In this way, social contract theory “tests” our current practices: if our practices are selected by contractors, then they are justified to us. If not, then not (Gaus 2011, 424–25).

Now in constructing a hypothetical agreement the question of how to precisely model the choice situation arises. To be sure, there are many ways of modeling the relevant parties, their preferences and information sets (i.e., beliefs), as well as the menu options they confront and much more. Furthermore, it is unlikely that there is one and only one best or correct way of fleshing out this model—there are many reasonable models of the initial choice problem, and thus many plausible social contracts we can use to appraise our current institutional order (Thrasher 2014). Yet, still, there are *some* constraints any suitable contract theory will place on the hypothetical choice situation. Such constraints follow from the very nature of the project we are engaged in—they are required so we do not, in David Schmidtz’s words, “set aside the problem” we are theorizing (2011, 777).

Since we are using social contract theory to gain some leverage on the reasons you and I have to endorse or reject our current set of institutions, it follows that the agreement between the parties in our hypothetical contract must, in some sense, actually speak to you and I. That is, it must be rational for actual persons in actual society to follow through and comply with whatever rules or institutions are ultimately selected. In his important work on the foundations of social contract theory, Gauthier notes that “the impartial perspective of the ideal actor must cohere with the perspectives of rational individuals actually engaged in strategic choice” (1986, 235). In their work on modeling hypothetical agreements, Gaus and John Thrasher call this the *identity test*, for actual persons in society must be able to *identify with* whatever our hypothetical parties choose, in that it is rational for you and I to follow through on whatever the output of the theory is (2015, 41). As an example, a contract situation that yielded the moral requirement that we all give to the point that we are paupers fails the identity test—it is hard to see

how people like us could see such a requirement as rational from our actual, non-normalized perspectives.

But when we ask whether our current institutional order is *justified*, and thus whether we have reason to continue supporting it, we are not asking whether we have any old kind of reason to do so—rather, we want to know whether we have *moral* reason to continue supporting it. Of course, the slaveholder certainly had *some* kind of reason to continue supporting the current institutional order in pre-Civil War America; yet, if such a set of rules were the object of agreement in our hypothetical choice situation then something has clearly gone wrong. Returning to Gauthier's work on the foundations of the social contract, he notes that contracts made in our hypothetical choice situations must be done from the Archimedean point of view, where any contractor's "concern is necessarily impartial, because it is based on the formal features of individual rational agency without the biasing content of a particular and determinant set of individual characteristics" (Gauthier 1986, 233). Gaus and Thrasher call this the *recognition test*, for you and I examining the agreement of our hypothetical parties must be able to *actually recognize* whatever is chosen as a bona fide moral proposal (Gaus and Thrasher 2015, 41). Note, insisting that the choice of our hypothetical parties be an impartial one does not commit us to there being one and only one impartial perspective. As Amartya Sen notes in his recent contribution to the social contract tradition, there may be many different and, indeed, incompatible impartial perspectives (Sen 2009, IX, 194–207). The recognition test merely requires our contractors be modeled using one such viewpoint.

Again, both the identity and recognition tests seem to intuitively follow from what it would mean to successfully complete the project: the recognition test follows from the fact that we seek *moral reasons* for why you and I have to comply with our current practices; and the identity test follows from the fact that we seek moral reasons for why *you and I have* to comply with our current practices. Yet balancing these tests is no easy task. Indeed, it is not obvious it can actually be done. As an example, though many may grant that Gauthier satisfied the recognition test with his theory of the social contract (although see Southwood 2010, ch. 2 on this point), the manner in which Gauthier tries to show his contract passes the identity test is deemed a failure by most. To show that it is rational to comply with his derived moral principle, Gauthier ends up changing the definition of rationality to a nonmodular conception at first (Gauthier 1986, ch. 6) and then finally to a conception that denies rational agents should behave strategically—they should simply choose what is Pareto optimal regardless what other players do (Gauthier 2013, 606–7). Note, this latter revision requires one to cooperate in prisoner's dilemmas. Revising the definition of rationality in such a manner in order to satisfy the identity test is a controversial move to say the least.

It is a fundamental conviction of the current project that no existing theory of the social contract satisfies both the identity and recognition tests. In repurposing Buchanan's contributions to the social contract, the goal is to develop a model of the social contract that satisfies *both* the identity test *and* the recognition test: we seek to model a contractual arrangement determining specific rules to live by that would be rational for us to follow but also strike us as bona fide moral proposals. If we succeed in doing this then not only do we offer a novel and more philosophically sophisticated take on Buchanan's social contract theory, but we also succeed where all others have failed.

THE FAILURES OF RAWLS—STRICT COMPLIANCE

We now turn to an analysis of Rawls's social contract theory in terms of the identity and recognition tests. Famously, Rawls held that the principles of justice for regulating the basic structure of society are those chosen by deliberators in the *original position*. When constructing the original position Rawls seems to have had the recognition test in mind: holding that "each according to his threat advantage is hardly a principle of fairness," Rawls heavily idealizes choosers so that morally arbitrary factors cannot influence the choice of principles, resulting in principles that are indeed recognizable as bona fide moral requirements (Rawls [1958] 1999, 58; 1971a, 141; [1971b] 1999, 207).

Ridding us of these morally arbitrary factors in the choice situation is primarily done through use of the *veil of ignorance*, which is a just a way of specifying the information sets of parties in our hypothetical choice procedure. When behind the veil of ignorance deliberators do not know their place in society, their natural talents and abilities, their own conception of the good, the circumstances of their society, what generation they inhabit, or any similar kind of information that could lead to bias in decision-making. All deliberators know is that the circumstances of justice obtain, as well as general facts about economics, politics, and psychology. As Rawls notes, "If the original position is to yield agreements that are just, the parties must be fairly situated and treated equally as moral persons. The arbitrariness of the world must be corrected for by adjusting the circumstances of the original contract situation" (Rawls 1971a, 141). The main way we correct these arbitrary features is by placing deliberators behind the veil of ignorance; this assumption is key in ensuring that Rawls's theory passes the recognition test.

But how does Rawls's theory fare according to the identity test? Now there is some difficulty in answering this question, for two reasons. First, when we introduced the identity test, we said that the test requires that it in some way be rational for persons to comply with whatever principles are chosen by the parties in our hypothetical choice situation. But second, Rawls's principles

of justice are not normative principles regulating individual conduct; rather, they are normative principles for evaluating institutional performance. The difference principle, for instance, says that *institutions* are just only if they maximize the welfare of those least advantaged in society. How does an individual comply with such an evaluative standard? If they cannot, then it makes no sense to ask whether it is rational for individuals to comply with whatever is agreed upon in the original position, and thus no sense to ask whether the theory passes the identity test.

Though Rawls's theory first and foremost delivers us principles for evaluating institutions, it also chooses a set of institutions that best satisfy these principles. This is accomplished in a four-stage sequence: (1) first deliberators select principles of justice in the original position, (2) then a constitution, (3) then specific legislative policies, (4) then decisions on how to apply agreed-upon policies to particular cases that confront members of society (Rawls 1971a, §31). Rawls models these different choice problems by altering the original position. As deliberators go through each stage of the four-stage sequence, we lift the veil of ignorance more and more, which means we include more information in the deliberators' information sets. In the constitutional convention, for instance, the information sets of deliberators are altered so that general facts about society are allowed to be known, such as the size and economic level of advancement of society, society's institutional structure and natural environment, etc. In the constitutional convention this means we allow information of "the kind framers of a constitution would want to know" (Rawls [1993] 2005, 398).

Perhaps, then, Rawls's social contract theory passes the identity test if it is rational to follow the rules of those institutions and policies that are ultimately selected in the four-stage sequence. This interpretation of the identity test, though, does not go quite far enough; for merely following those rules laid down by our selected institutions will not guarantee that justice will actually be done, in that merely following those requirements of the selected institutions will not necessarily bring about a state of affairs satisfying the two principles of justice. Instead, we should interpret the identity test for Rawls as follows: Rawls's social contract theory passes the identity test if and only if it is rational for persons, acting within those institutions Rawls's theory prescribes to, bring about a state of affairs where Rawls's principles of justice are satisfied. If, given the institutions Rawls's theory prescribes, rational behavior leads to satisfaction of the two principles of justice or something close to it, then the theory passes the identity test; if, given the institutions Rawls's theory prescribes, rational behavior does *not* lead to satisfaction of the two principles of justice or something close to it, then the theory fails the identity test.

Does Rawls's theory pass the identity test given this way of fleshing the test out? It is highly unlikely, and the reason why has to do with a specific

assumption Rawls introduces into his model of the hypothetical choice situation. When modeling deliberators in the original position, Rawls stipulates that deliberators assume that “the parties can rely on each other to understand and act in accordance with whatever principles are finally agreed to” (Rawls 1971a, 145). This is known as the *strict compliance assumption* and has been the subject of intense debate very recently (e.g., Stemplowska and Swift 2012; Stemplowska and Swift 2014; Valentini 2012). A bit more precisely, “In assessing conceptions of justice the persons in the original position are to assume that the one they adopt will be strictly complied with. The consequences of their agreement are to be worked out on this basis” (Rawls 1971a, 145). That is, when assessing different principles of justice and their corresponding institutions, we model deliberators as assuming that persons in society will be moved by a sense of justice—they will be motivated to do as justice requires, not pursue their private interests.

Now this assumption of strict compliance, in my view, has little effect on the choice of principles of justice. However, I *do* think this assumption exerts an immense influence on the choice of institutions in the second and third stages of Rawls’s four-stage choice procedure: where deliberators choose the constitution best satisfying our already-chosen principles, as well as specific legislative policies that bring about the requirements of justice. The reason why there is such great influence here is because the strict compliance assumption essentially allows parties in the original position to assume away worries stemming from public choice theory: they do not have to consider the possibility of “politics without romance,” for by assumption, public actors are moved by their sense of justice (Buchanan [1979] 1999). In his discussion of political behavior among representatives in a legislature, Rawls says that such government officials “represent their constituents in the substantive sense: they must seek first to pass just and effective legislation, since this is a citizen’s first interest in government, and secondly, they must further their constituents’ other interests insofar as these are consistent with justice” (Rawls 1971a, 221).

Looking more to the tangible impact of this assumption, a cursory glance suggests that the institutions Rawls ends up endorsing are quite possibly the worst possible institutions we could select from a public choice perspective. For instance, Rawls holds that the legislative rule best satisfying the two principles of justice is bare majoritarianism, with either no or very minimal constraints placed on what simple majorities may do (Rawls 1971a, 224). Yet the public choice tradition has shown us that such a rule has particularly undesirable equilibria (Buchanan and Tullock [1962] 2004, chs. 10–11; Tullock [1959] 2004). As a further example, those economic institutions Rawls endorses—property-owning democracy and market socialism—likely face huge incentive problems when we consider rent-seeking, regulatory capture,

and the massive bureaucracies that will be required for these economic systems to function as they should (Vallier 2015, 294–96). But, when deliberators assume that persons in society have a sense of justice, such problems simply cannot be considered.

The upshot here is that precisely because Rawls introduces the strict compliance assumption into his initial choice situation, his theory fails the identity test. When persons are acting within those institutions that Rawls ends up prescribing—property-owning democracy and market socialism with no or very minimal restrictions on bare majoritarian rule—it is unlikely that rational action within such an institutional environment will bring about a state of affairs where the two principles of justice are satisfied. That is, it is unlikely that rational behavior within such institutions actually protects the basic liberties and secures fair equality of opportunity, all the while maximizing the welfare of those least advantaged. The opportunities for discriminatory legislation, rent-seeking, regulatory capture, and the like will be far too great. As a result, the outcome that is ultimately realized from Rawls’s preferred set of institutions will likely look quiet foreign from the perspective of justice as fairness. So though Rawls’s theory can get us recognition, it cannot get us identity. The principles are without a doubt, moral, but it is not clear in what sense they are rational for people like you and I to follow. We now turn to see if Buchanan can get us identity, while preserving Rawls’s recognition.

THE SUCCESSES AND FAILURES OF BUCHANAN— HOMO ECONOMICUS AND THE VEIL OF UNCERTAINTY

In at least one of his models of the social contract, Buchanan presents a group of individuals choosing constitutional rules under which normal politics can then proceed. As Buchanan and Viktor Vanberg note, it is not only conflicting interests that lead to diverging preferences over constitutional rules. Individuals can also differ in terms of how they predict different constitutional rules might operate—that is, they can differ in terms of the predicted outcomes of different constitutional rules (Vanberg and Buchanan 1989). Even if preferences are identical, such social scientific disagreements can still result in conflicts over which constitutional rules are best.

In dealing with this problem of prediction, Buchanan (in conjunction with Tullock) quite revolutionarily proposed we apply the economist’s rational choice behavioral model to the study of political institutions ([1962] 2004, ch. 3). Even more specifically than the basic rational choice model—which merely says that individuals have complete and transitive preference orderings, and choose their most choice-worthy option—Buchanan further insisted that we apply a very restricted version of the rational choice model, where

individual utility functions are defined narrowly so that persons prefer more wealth to less (Kirchgässner 2014, 4–7). We shall call this the *homo economicus* behavioral postulate. In Buchanan’s words, “We must place certain restrictions on individual utility functions, restrictions which are precisely analogous to those introduced in economic theory . . . we must assume that individuals will, on the average, choose ‘more’ rather than ‘less’ when confronted with the opportunity for choice in a political process, with ‘more’ and ‘less’ being defined in terms of measurable economic position” (Buchanan and Tullock [1962] 2004, 28).

Why use the *homo economicus* behavioral postulate specifically? Buchanan (often when coauthoring with Geoffrey Brennan) offers several philosophical defenses of this behavioral assumption (Brennan and Buchanan 1981; 1983; [1985] 2000, ch. 4), all of which have recently been subject to critical scrutiny (Kogelmann 2015). These arguments aside, our interest is in whether the inclusion of *homo economicus* allows Buchanan’s model of the social contract to satisfy the identity test. Now recall why Rawls’s social contract failed the identity test: when persons act rationally within the institutions Rawls proposes, it is unlikely that such behavior achieves the desired result—it is unlikely that such behavior brings about a state of affairs where justice as fairness is satisfied. This was because of Rawls’s use of the strict compliance assumption in the derivation of his proposal.

Will persons acting rationally within the rules and institutions chosen by Buchanan’s social contract satisfy the underlying normative criteria? (And here, Buchanan leaves open with what normative criteria deliberators appraise different constitutional arrangements.) That all depends. Namely, it depends on how close persons in the real world approximate the *homo economicus* behavioral postulate. If deliberators in the hypothetical contract choose constitution *C* because *C* best satisfies normative criteria *N* under the assumption that people behave as *homo economicus* does, but persons in the real world do *not* behave as *homo economicus* does, then actually implementing *C* might not satisfy *N*. That is, rational action within the chosen institutional arrangement might not produce a state of affairs satisfying our underlying normative desiderata. It thus seems that Buchanan’s model of the social contract also fails the identity test, as did Rawls’s model.

In response, we note that the *homo economicus* behavioral assumption is not as integral a component of Buchanan’s model of the social contract as it might initially seem. As he notes, “The self-interest assumption, for our construction, serves an empirical function. As such, it may or may not be ‘realistic’: this can only be determined by a comparison of some of the positive analytical implications with observable real-world facts” (Buchanan and Tullock [1962] 2004, 297). But, what *is* an integral component of Buchanan’s social contract is a commitment to politics without romance; or, in other

words, a commitment to “take men as they are, not as [we] would like them to be” (Buchanan and Tullock [1962] 2004, 296). “Men as they are” *could* mean *homo economicus*, or it *could* mean something quite different—as Buchanan notes, this can only be determined empirically. But if it *is* something quite different, then it is perfectly in line with—indeed, perhaps required by—Buchanan’s social contract theory to model persons in such a manner. If one does this, then the model *will* pass the identity test: if deliberators in the hypothetical contract choose constitution *C* because *C* best satisfies normative criteria *N* under the assumption that people behave in manner *B*, then when *C* is implemented it *will* satisfy *N* because, by hypothesis, people do in fact behave in manner *B*. Buchanan’s behavioral postulate—if we understand it not as insisting on *homo economicus* strictly but rather as taking men as they are, be this *homo economicus* or not—thus allows his model to succeed where Rawls’s model failed: it allows the model to pass the identity test.

Though Buchanan’s social contract excels at identity, it falters at recognition. The reason why has to do with the epistemological assumptions Buchanan places on contractors—what we shall call the *veil of uncertainty*.³ In modeling the social contract, Buchanan (in conjunction with Tullock) describes the information sets of the parties as follows:

Recall that we try only to analyze the calculus of the utility-maximizing individual who is confronted with the constitutional problem. Essential to the analysis is the presumption that the individual is *uncertain* as to what his own precise role will be in any one of the whole chain of later collective choices that will actually have to be made. For this reason he is considered not to have a particular and distinguishable interest separate and apart from his fellows. This is not to suggest that he will act contrary to his own interest; but the individual will not find it advantageous to vote for rules that may promote sectional, class, or group interests because, by presupposition, he is unable to predict the role that he will be playing in the actual collective decision-making process at any particular time in the future. He cannot predict with any degree of certainty whether he is more likely to be in a winning or losing coalition on any specific issue. Therefore he will assume occasionally he will be in one group and occasionally in the other. His own self-interest will lead him to choose rules that maximize the utility of an individual in a series of collective decisions with his own preferences on separate issues being more or less randomly distributed. (Buchanan and Tullock [1962] 2004, 74)

By my reading, Buchanan here intends to place *no* constraints on the information sets of the contracting parties at all. Rather, Buchanan thinks that the *mere exercise* of constitutional choice places those involved in the choice problem in a state of relative ignorance—no restrictions imposed by *the*

theorist are required. This is in contrast to Rawls, who does specifically model contractors as behind the highly artificial veil of ignorance.

Now there are two things we need to note about Buchanan's veil of uncertainty. First, in saying that the information sets of deliberators are determined by the actual exercise of drafting a constitution, Buchanan seems to suggest that the social contract is an *actual* agreement *actual* persons undergo, not a hypothetical choice model as explicated in §3. This, I think, adds another tension to Buchanan's theorizing about the social contract to the one already outlined in §2 above. In his great essay, "Positive Economics, Welfare Economics, and Political Economy," Buchanan critiques welfare economists for making welfare judgments based on assumptions about people's preferences: walking a radically behavioralist line, Buchanan asserts that "a specific change may be judged to be Pareto-optimal or 'efficient' only after it has, in fact, been proposed and the individual preferences for or against the change revealed," where preferences are revealed "by the overt action of the individual in choosing" (Buchanan 1959 [1999], 194–95). That is, we can only make judgments as to whether persons are better off by looking at actual, not hypothetical, choices.

Yet, in the very same essay, Buchanan goes on to say that the unique approach to political economy he is putting forward *must* assume that "the social group is composed of reasonable men, capable of recognizing what they want, of acting on this recognition, and of being convinced of their own advantage after reasonable discussion" (Buchanan [1959] 1999, 294). But, because Buchanan realizes that "societies in the real world are not made up exclusively of reasonable men," the political economist must make some *as if* assumptions about preferences, and thus some judgments concerning when a contract would indeed be the subject of unanimous agreement, even if it is not in fact unanimously agreed upon (Buchanan [1959] 1999, 294). So it seems Buchanan tells us both to *only* rely on *actual* agreement as normatively relevant, *and* that *hypothetical* agreements should be used as a simplifying assumption in dealing with the unreasonable. In a similar vein, Buchanan's discussion of the veil of uncertainty seems to describe the information sets of participants in an actual constitutional convention. Yet, Buchanan's explicit engagement with the normative social contract tradition (highlighted in §2 above) suggests that he is working on a hypothetical model of an idealized choice situation. Since the veil of uncertainty seems to describe the actual information sets of actual persons, it remains a bit unclear just how Buchanan intends to specify the beliefs of deliberators in our hypothetical choice procedure.

This leads into my second point. If we are to understand the veil of uncertainty as describing the information sets of actual persons engaged in

a constitutional contract, then many of Buchanan's claims become operationalizable hypotheses subject to testing—does the very act of drafting a constitution put persons in the kind of epistemic state Buchanan seems to think it does? And, indeed, several have pursued this very line of inquiry. Overall, the results are rather mixed, if not slanting against what Buchanan says—several case studies show that the mere exercise of constitutional construction and choice is *not* sufficient to induce impartiality among drafters. As one study notes after examining the dynamics of constitution drafting in Greece,

the hypothesis that constitutional writing in Greece was the result of disinterested framers designing institutions behind a veil of ignorance to cope with uncertainty receives less than solid support, while there seems to be support for the conjecture that constitutional writing responded to the experiences and reflected the interests of whoever held political advantage at the time of writing. (Imbeau and Jacob 2015, 185)

Even the American Revolution and subsequent drafting of the Constitution by the founding fathers—what has been called a “best case scenario so far as aiming to probe the constitution-making process under conditions of deep and extensive uncertainty” (Imbeau and Jacob 2015, 99)—has been charged as biased and partial. Famously, Charles Beard argued that the drafting of the Constitution reflected the personal financial interest of the founding fathers, not impartial reflections on the common good (1913).

What does all of this mean? It means that *if* we take the veil of uncertainty to be determined by plausible information sets of actual constitutional drafters, *then* it is unlikely that the information sets will achieve the kind of impartiality we seek for our theory of the social contract. But if deliberators in our choice situation are not choosing impartially, then our contract fails the recognition test—the resulting constitution will not strike us as a bona fide moral proposal if deliberators game the choice to benefit themselves. For the end result is just what Rawls feared: a social contract where each is given according to their threat advantage. And this, Rawls notes, hardly seems fair.

A BETTER SOCIAL CONTRACT

Here is where we stand: Rawls's social contract passes the recognition test but fails the identity test. Buchanan's social contract passes the identity test but fails the recognition test. In this final section, I sketch an outline of a new model of the social contract that gets us both. Naturally, such a model will

be inspired by those successful elements of Rawls's social contract and those successful elements of Buchanan's social contract.

Return first to Buchanan's inability to satisfy the recognition test: the problem here was that the veil of uncertainty allows too much knowledge of personal circumstances into deliberators' information sets. As theorists, this strikes us as inappropriate from the moral point of view. In remedying this failure, the simplest thing to do is alter the information sets so that they more closely resemble Rawls's veil of ignorance. On this model, deliberators in Buchanan's constitutional convention choose a constitution to govern postconstitutional politics with no knowledge of their place in society, their natural talents and abilities, their own conception of the good, the circumstances of their society, as well as what generation they inhabit. But upon constructing such a model, a question immediately arises: *How* are deliberators to choose a constitution without any knowledge of this kind? How do they select a constitution when they literally know nothing about themselves or the society they inhabit?

Here we turn back to Rawls. In dealing with this problem, Rawls assumes that deliberators in the original position have preferences over what he calls *primary goods*, which are defined as "things that every rational man is presumed to want" (Rawls 1971a, 62). Included here are rights and liberties, powers and opportunities, income and wealth, as well as self-respect. More specifically, in Rawls's original position it is assumed that deliberators simply prefer more primary goods to less (Rawls 1971a, 93). But how can we make such an assumption about deliberators' preferences if they literally have no knowledge of the persons they are supposed to represent? Rawls reasons that "though men's rational plans do have different final ends, they nevertheless all require for their execution certain primary goods, natural and social . . . whatever one's system of ends, primary goods are necessary means" (Rawls 1971a, 93). That is, *regardless* what kind of person deliberators end up being, in order to advance whatever conception of the good they end up holding, more primary goods will better facilitate this advancement when compared to less primary goods. In other words, it is rational to want more primary goods regardless who one is.

Running with this assumption that deliberators prefer more primary goods to less, we return to Buchanan. Suppose we now assume that deliberators in Buchanan's constitutional convention prefer more primary goods to less, and that they seek the constitution that maximizes their bundle of primary goods. But here, it is not clear in what sense deliberators can make such judgments about the effects different constitutions have on distributions of primary goods. The constitutional political economy research program can tell us things more specific: that without constitutional constraints on bare majority rule we can run afoul of things like inflation traps, debt traps, and high-tax

traps (Brennan and Buchanan [1985] 2000, ch. 6). But it remains unclear to what extent the public choice research program can tell us which sorts of legislative rules and constitutional protections maximize things as abstractly defined as primary goods.

Hitting a snag, we turn back to Rawls. Rawls's assumption about preferences over primary goods works quite well in the original position when deliberators consider competing principles of justice, not specific constitutions and legislative policies. This is because the principles of justice just are cashed out in terms of primary goods: utilitarianism is understood as the theory requiring that basic institutions maximize total primary goods; the difference principles as the theory requiring that the primary goods of those least off be maximized, and so forth. Under these assumptions, Rawls holds that the unique solution to the rational choice problem posed in the original position is his two principles of justice, or justice as fairness. The assumption of deliberators preferring more primary goods to less, then, does not make the choice situation impossible to navigate in the original position as it does in Buchanan's constitutional convention. A well-defined choice problem is presented, and one set of principles of justice selected.

But if we simply follow Rawls here, will we not run afoul of the identity test? This is where we bring Buchanan back in. After deliberators select principles of justice in the original position, they move on to the constitutional convention. But now, when deliberators consider potential constitutions they (1) look for the constitution that best satisfies the principles of justice chosen in the original position, but (2) do so under the assumptions Buchanan employs in his constitutional political economy research program. That is, deliberators examine constitutions in terms of how well they will perform at doing what justice requires, under the assumption that persons in society will *not* want to do as justice requires. Buchanan's self-interested behavioral postulate thus replaces Rawls's strict compliance assumption. Deliberators examine constitutions in terms of how well they satisfy the principles of justice, under the assumption that men are as they are—however it is they in fact are—not as we would like them to be.

Note the constitutional political economy research program is better-suited to serve this role because the requirements of justice are more concrete than the abstract nature of primary goods.⁴ As an example, in their recent study, Torsten Persson and Guido Tabellini ask,

If the United Kingdom were to switch its electoral rule from majoritarian to proportional, how would this affect the size of its welfare state or of its budget deficits? If Argentina were to abandon its presidential regime in favor of a parliamentary form of government, would this facilitate the adoption of sound policy towards economic development? (2005, 7)

Similarly, deliberators in our constitutional convention must ask this: Which sorts of voting rules in the legislature will most likely result in legislation satisfying the difference principle? Would presidential or parliamentary forms of government better protect the basic liberties? What election schemes will best protect the fair value of the political liberties? To take an example more relevant to Buchanan's research in particular, if we want to maximize transfers to the poor, then we need to avoid hitting a suboptimal point on the Laffer curve (Rawls [1967] 1999, 145), yet high-tax traps by definition do just that. Buchanan's work on tax traps and political equilibria can thus help us devise a system that avoids these pitfalls so that we are better able to actually maximize the welfare of those least advantaged, as justice requires (Buchanan and Lee 1982a; 1982b).

It should be noted that our new social contract will involve a significant invisible hand element to it. In his analysis of the market, Adam Smith notes that participants in the marketplace act as if they are "led by an invisible hand to promote an end which was no part of his intention" ([1776] 1981, 456). According to Smith, if market institutions are structured properly, then consumers maximizing consumption bundles and producers maximizing profits will induce a state of affairs that benefits all, even though no participant in the market actually intended to do so. Similarly, in using Buchanan's constitutional analysis in conjunction with Rawls's normative principles, we seek to find the constitution that most likely brings about justice as fairness under the assumption that persons in society do not care about justice—they only, we assume, pursue their private interests. Just as market actors in Smith's market place produce a Pareto efficient state of affairs with no intention to, our hope is that with the right constitutional rules, political actors in our society will produce justice with no intention to. On our new social contract theory, justice is an emergent property; it is not the result of persons with a sense of justice trying to bring it about as it is in Rawls's model, but rather the result of human action, not human design.

So here, in summary, is the proposal. Our social contract consists of a two-stage choice sequence. In the first stage deliberators choose principles of justice as they do in Rawls's model of the original position. In the second stage, deliberators choose the constitution with the tools and assumptions of Buchanan's constitutional political economy research program, yet with an eye to justice: assuming men are as they are and not as we would like them to be, they try to pick the constitution that best brings about what justice requires anyways. This, I think, is our best hope of developing a social contract theory that passes both the identity and recognition tests. We pass the recognition test because the principles of justice chosen, just like with Rawls's social contract, intuitively strike us as bona fide moral principles given the impartiality modeled into the designed choice procedure. We pass the identity test

because, just like with Buchanan's social contract, the institutions chosen are selected with our actual rationality in mind. If all goes well, then rational action within the chosen institutions will produce a just state of affairs. If our society actually possesses such institutions then they will be justified to us, for we both recognize them and identify with them: they only require we be rational yet, in doing so, allow us to be part of a genuine moral order.

CONCLUSION

Moral theorizing is hard. If done right, then we end up with a set of moral reasons and their requirements that are something we can legitimately aspire to. Aspiration, though, demands more than one might initially think. For aspiration requires not only something above and beyond where we currently are, but something not so far above and beyond where we currently are such that we could not possibly imagine achieving what our theorizing says we should achieve. This, we have seen, is the central tension at the heart of the social contract: not only must we recognize the requirements of the contract as bona fide moral proposals, but we must also be able to identify with these demands as speaking to us and our own rationality. The paradigmatic examples of the social contract—pursued by John Rawls and James M. Buchanan respectively—failed to live up to these ideals. But even though this is so lessons can be learned and a better model of the social contract attained—this, indeed, is the aspiration of this chapter.

NOTES

1. Not only is Buchanan included among the great contemporary philosophers, but he had significant correspondence and scholarly interaction with some of them as well. For Rawls on Buchanan, see Rawls ([1963]1999, 74; 1971a, 70, 197, 266–67); [1974] 1999, 225; and 2001, 16, 97). For Buchanan on Rawls see Buchanan (1972; 1976; 2000; 2003), and Buchanan and Faith (1980). For letters from their correspondence see Levy and Peart (2008, ch. 19). And for further analysis see Amadae (2003, 149–52; 2011; 2015, 182–87).

2. These are not the only tensions present in Buchanan's social contract theory. Another inconsistency is whether Buchanan thinks the social contract requires actual or hypothetical consent. This issue is touched on in §5 below.

3. It is worth noting that Buchanan sees his veil of uncertainty as being very similar to Rawls's veil of ignorance. I argue below that this is not the case. See Buchanan ([1975] 2000: 10n4, 91n18, 221n8).

4. For an excellent recent survey of empirical approaches to analyzing constitutions, see Tarko (2015).

REFERENCES

- Amadae, S. M. 2003. *Rationalizing Capitalist Democracy: The Cold War Origins of Rational Choice Liberalism*. Chicago: University of Chicago Press.
- Amadae, S. M. 2011. "James M. Buchanan, John Rawls, and Democratic Governance." In *Deliberative Democracy: Theory and Practice*, edited by Robert Cavalier, 31–52. Pittsburgh, PA: Carnegie Mellon Press.
- Amadae, S. M. 2015. *Prisoners of Reason: Game Theory and Neoliberal Political Economy*. Cambridge: Cambridge University Press.
- Beard, Charles. 1913. *An Economic Interpretation of the Constitution of the United States*. New York: The Macmillan Company.
- Brennan, Geoffrey and James M. Buchanan. 1981. "The Normative Purpose of Economic 'Science': Rediscovery of an Eighteenth Century Method." *International Review of Law and Economics* 1 (2): 155–66.
- Brennan, Geoffrey and James M. Buchanan. 1983. "Predictive Power and the Choice Among Regimes." *The Economic Journal* 93 (369): 89–105.
- Brennan, Geoffrey and James M. Buchanan. (1985) 2000. *The Reason of Rules: Constitutional Political Economy*. Indianapolis: Liberty Fund.
- Buchanan, James M. (1959) 1999. "Positive Economics, Welfare Economics, and Political Economy." In *The Calculus of Consent: The Logical Foundations of Constitutional Democracy*, 191–209. Indianapolis: Liberty Fund.
- Buchanan, James M. 1972. "Rawls on Justice as Fairness." *Public Choice* 13 (1): 123–28.
- Buchanan, James M. (1975) 2000. *The Limits of Liberty*. Indianapolis: Liberty Fund.
- Buchanan, James M. 1976. "A Hobbesian Interpretation of the Rawlsian Difference Principle." *Kyklos* 29 (1): 5–25.
- Buchanan, James M. (1979) 1999. "Politics without Romance: A Sketch of Positive Public Choice Theory and its Normative Implications." In *The Calculus of Consent: The Logical Foundations of Constitutional Democracy*, 45–59. Indianapolis: Liberty Fund.
- Buchanan, James M. (1987) 2001. "Justification of the Compound Republic: The *Calculus* in Retrospect." In *Choice, Contract, and Constitutions*, 68–75. Indianapolis: Liberty Fund.
- Buchanan, James M. 2002. "John Rawls, *Justice as Fairness: A Restatement*." *Public Choice* 113 (3): 488–90.
- Buchanan, James M. 2003. "Justice Among Natural Equals: Memorial Marker for John Rawls." *Public Choice* 114 (3–4): iii–v.
- Buchanan, James M. and Roger Faith. 1980. "Subjective Elements in Rawlsian Contractual Agreement on Distributional Rules." *Economic Inquiry* 18 (1): 23–38.
- Buchanan, James M. and Dwight R. Lee. 1982a. "Politics, Time, and the Laffer Curve." *Journal of Political Economy* 90 (4): 816–19.
- Buchanan, James M. and Dwight R. Lee. 1982b. "Tax Rates and Tax Revenues in Political Equilibrium Some Simple Analytics." *Economic Inquiry* 20 (3): 344–54.
- Buchanan, James M. and Gordon Tullock. (1962) 2004. *The Calculus of Consent: The Logical Foundations of Constitutional Democracy*. Indianapolis: Liberty Fund.

- Freeman, Samuel. (1990) 2007. "Reason and Agreement in Social Contract Views." In *Justice and the Social Contract: Essays on Rawlsian Political Philosophy*, 17–44. Oxford: Oxford University Press.
- Gaus, Gerald. 1996. *Justificatory Liberalism: An Essay on Epistemology and Political Theory*. Oxford: Oxford University.
- Gaus, Gerald. 2011. *The Order of Public Freedom: A Theory of Freedom and Morality in a Diverse and Bounded World*. Cambridge: Cambridge University Press.
- Gaus, Gerald and John Thrasher. 2015. "Rational Choice and the Original Position: The (Many) Models of Rawls and Harsanyi." In *The Original Position*, edited by Timothy Hinton, 39–58. Cambridge: Cambridge University Press.
- Gauthier, David. 1986. *Morals by Agreement*. Oxford: Oxford University Press.
- Gauthier, David. 2013. "Twenty-Five On." *Ethics* 123 (4): 601–24.
- Imbeau, Louis M. and Steve Jacob. 2015. *Behind a Veil of Ignorance? Power and Uncertainty in Constitutional Design*. New York: Springer.
- Kirchgässner, Gebhard. 2014. "The Role of Homo Oeconomicus in the Political Economy of James Buchanan." *Constitutional Political Economy* 25 (1): 2–17.
- Kogelmann, Brian. 2015. "Modeling the Individual for Constitutional Choice." *Constitutional Political Economy* 26 (4): 455–74.
- Levy, David M. and Sandra J. Peart. 2008. *The Street Porter and the Philosopher: Conversations on Analytical Egalitarianism*. Ann Arbor: University of Michigan Press.
- Perrson, Torsten and Guido Tabellini. 2005. *The Economic Effects of Constitutions*. Cambridge, MA: MIT Press.
- Rawls, John. (1958) 1999. "Justice as Fairness." In *Collected Papers*, edited by Samuel Freeman, 47–72. Cambridge, MA: Harvard University Press.
- Rawls, John. (1963) 1999. "Constitutional Liberty and the Concept of Justice." In *Collected Papers*, edited by Samuel Freeman, 73–95. Cambridge, MA: Harvard University Press.
- Rawls, John. (1967) 1999. "Distributive Justice." In *Collected Papers*, edited by Samuel Freeman, 130–53. Cambridge, MA: Harvard University Press.
- Rawls, John. 1971a. *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Rawls, John. 1971b. "Justice as Reciprocity." In *Collected Papers*, edited by Samuel Freeman, 190–224. Cambridge, MA: Harvard University Press.
- Rawls, John. (1974) 1999. "Some Reasons for the Maximin Criterion." In *Collected Papers*, edited by Samuel Freeman, 225–31. Cambridge: Harvard University Press.
- Rawls, John. (1993) 2005. *Political Liberalism*. New York: Columbia University Press.
- Rawls, John. 2001. *Justice as Fairness: A Restatement*. Cambridge, MA: Harvard University Press.
- Schmidtz, David. 2011. "Nonideal Theory: What It Is and What It Needs to Be." *Ethics* 121 (4): 772–96.
- Sen, Amartya. 2009. *The Idea of Justice*. Cambridge, MA: Harvard University Press.
- Smith, Adam. (1776) 1981. *An Inquiry into the Nature and Causes of the Wealth of Nations*. Indianapolis: Liberty Fund.
- Southwood, Nicholas. 2010. *Contractualism and the Foundations of Morality*. Oxford: Oxford University Press.

- Stemplowska, Zofia and Adam Swift. 2012. "Ideal and Nonideal Theory." *The Oxford Handbook of Political Philosophy*, edited by David Estlund, 373–89. Oxford: Oxford University Press.
- Stemplowska, Zofia and Adam Swift. 2014. "Rawls on Ideal and Nonideal Theory." *Blackwell Companion to Rawls*, edited by Jon Mandle and David A. Reidy, 112–27. Chichester: John Wiley & Sons.
- Tarko, Vlad. 2015. "The Challenge of Empirically Assessing the Effects of Constitutions." *Journal of Economic Methodology* 22 (1): 46–76.
- Thrasher, John. 2014. "Uniqueness and Symmetry in Bargaining Theories of Justice." *Philosophical Studies* 167 (3): 683–99.
- Thrasher, John and Gerald Gaus. 2017. "On *The Calculus of Consent*." In *The Oxford Handbook of Classics in Contemporary Political Theory*, edited by Jacob Levy. Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780198717133.013.25.
- Tullock, Gordon. (1959) 2004. "Problems of Majority Voting." In *Virginia Political Economy*, 51–68. Indianapolis: Liberty Fund.
- Valentini, Laura. 2012. "Ideal vs. Non-Ideal Theory: A Conceptual Map." *Philosophy Compass* 7 (9): 654–64.
- Vallier, Kevin. 2015. "A Moral and Economic Critique of the New Property-Owning Democrats: On Behalf of a Rawlsian Welfare State." *Philosophical Studies* 172 (2): 283–304.
- Vanberg, Viktor and James M. Buchanan. 1989. "Interests and Theories in Constitutional Choice." *Journal of Theoretical Politics* 1 (1): 49–62.