Exchanges with and without the sword: slavery, politics-as-exchange and freedom in James M. Buchanan’s institutional economics

John Meadowcroft

Department of Political Economy, King’s College London, London, UK
E-mail: john.meadowcroft@kcl.ac.uk

(Received 2 September 2022; revised 24 May 2023; accepted 24 May 2023)

Abstract
James M. Buchanan’s politics-as-exchange retrospectively conceptualized formal institutions emerging from bilateral agreements to establish reciprocal rights and prospectively guided constitutional entrepreneurs to broker Pareto-superior reforms that had unanimous consent. Buchanan believed this conceptualization of politics-as-exchange was necessitated by his ontological–methodological individualism and would initiate a new era of consensual politics, but it is argued it led to illiberal conclusions that reflected dissonance between his Kantian individualism and Humean subjectivism. It meant, for example, that slavery was characterized as a bilateral agreement between very unequal parties and it is argued it logically implied abolition required the consent of slaveowners. But Buchanan’s ontology was compatible with the introduction into his framework of a right of exit that would have differentiated between exchanges with and without the sword to produce a consistent liberal constitutionalism.

Keywords: Buchanan; consent; exchange; exit; slavery

James M. Buchanan was an architect of the new institutional economics, a principal founder of public choice theory, and a leading exponent of modern social contract theory. Buchanan (2005: 58) was committed to a liberal constitutionalism that envisaged, ‘an idealized structure of social interaction is possible in which no person exerts power over another... Coercion by another person is drained out; individuals are genuinely “at liberty”’. Buchanan (1975a/2000, 1986a/2001, 1989a) theorized this liberal constitutionalism through a unique framework of politics-as-exchange in which formal institutions were retrospectively conceptualized emerging from bilateral agreements to establish reciprocal rights and prospectively alterations to those rights required the consent of all rights-holders.

But scholars both hostile and sympathetic to Buchanan’s project have been sceptical that he offered a convincing or appealing liberal theory. Carole Pateman (1988: 60–62), an influential critic of liberal constitutionalism, argued Buchanan’s social contract theory justified and formalized the subjugation of the weak by the strong and raised ‘some embarrassing questions about actual contracts in our society’. Hartmut Kliemt (2012: 440–441), one of the editors of Buchanan’s collected works, argued that at best Buchanan offered an account of ‘fictional consent’ that did not translate to real-world institutions.

These criticisms follow from the fact that Buchanan’s conceptualization of institutions resulting from bilateral agreements included exploitative and oppressive arrangements and his stipulation that institutional reform should have unanimous consent included the agreement of rights-holders who had created or benefitted from unjust arrangements. Buchanan’s liberal theory was compatible with ostensibly illiberal outcomes.

The linked examples of slavery and abolition highlight the pathologies of Buchanan’s framework as it applied to real-world institutions. Slavery may be understood as the most extreme institutional case...
that provides the most challenging test of Buchanan’s position. As one of Buchanan’s co-authors, Roger Congleton (2014: 54), wrote, ‘If even slavery is conceptually compatible with [Buchanan’s] contractarianism’, then it would seem within his theory no social state can ‘be considered illegitimate’. Slavery might also be considered a salient example given Nancy MacLean’s (2017) claims that Buchanan advanced a covert racist agenda, although the credibility of MacLean’s work has been significantly diminished by subsequent scholarship (Farrant, 2019; Farrant and Tarko, 2019; Magness, 2020; Magness et al., 2019; Marciano, 2020; Munger, 2018a).

Buchanan did not normatively justify slavery, neither as a theoretical construct nor as it has existed in any time or place, but within politics-as-exchange he retrospectively characterized slavery as a bilateral agreement in the state of nature between very unequal persons in which the weaker party sought to avoid an even worse fate (Buchanan, 1975a/2000: 77–78, 1987/2001, 1988a, 1991/1999, 1995a, 2003) and it has been argued prospectively it logically follows that abolition required the consent of slaveowners (Meadowcroft, 2020: 397–400; Mueller, 2003: 143; Reisman, 2015: 121–122).

Norman Barry (1984) and Randall Holcombe (2014, 2020) have argued important pathologies of Buchanan’s work resulted from an unresolved tension between his espousal of classical liberalism and his contractarianism: the former prioritized individual liberty and the institutions of limited government; the latter led to a conservative defence of the status quo, even one populated with unjust or illiberal arrangements.

In this article it is argued that the conflict within Buchanan’s work was in fact between his individualism and his subjectivism, or between the Kantian and the Humean aspects of his project. This dissonance between Kantianism and Humeanism is not unique to Buchanan – Chandran Kukathas (1989: 205) found a similar problem in the work of F. A. Hayek. Moreover, this tension reflected the fundamental division in liberal thought Kukathas (2003: 19) identified between the belief that a liberal society involved universal acceptance of the principles embodied in negative rights and the view that a liberal society was a framework within which different, even potentially illiberal, values and practices co-existed.

It is argued herein that Kukathas’ (2003) theory of the ‘liberal archipelago’, and also Robert Nozick’s (1974: Part III) ‘framework for utopia’, suggest the tension in Buchanan’s work may be resolved by introducing a right of exit into his institutional economics. This would necessitate differentiating between exchanges with and without the sword – bilateral agreements that are genuinely consensual and those only made because the weaker party is under duress and are then maintained by force. It will be argued that a right of exit can be grounded in Buchanan’s ontology and therefore be consistent with his individualism and subjectivism.

Buchanan (1975a/2000: 11) wrote that he was, ‘in basic values, an individualist, a constitutionalist, a contractarian, a democrat’ – the values of liberal constitutional models of democracy presently challenged from within democracies by populist and extremist political movements and from outside by authoritarian and autocratic regimes. The ontological–methodological basis of Buchanan’s project may offer a more robust defence of constitutional democracy than theories that rely on normative or empirical claims open to dispute by critics. This article proposes to revise Buchanan’s institutional economics to produce a more consistent and compelling theory of liberal constitutionalism derived from Buchanan’s original framework. It is argued that, as Buchanan thought, a liberal society should allow the co-existence of diverse values and practices, but, contra Buchanan, individuals should be free to exit from arrangements they find intolerable. A successful liberal theory should recognize the importance of subjectivism but place it within an immutable framework of individualism.

After this Introduction, I set out Buchanan’s theory of politics-as-exchange that was the core of his institutional economics. The next section applies politics-as-exchange to the cases of slavery and abolition to show the practical implications of his framework. Politics-as-exchange is then placed in the context of Buchanan’s Kantian individualism and Humean subjectivism and the dissonance between these two constituent parts of his theory is elucidated. I next show that a right of exit like that found in the liberal theories of Kukathas and Nozick can be grounded in Buchanan’s ontology–methodology to produce a more consistent liberal constitutionalism. A final section concludes.
Politics-as-exchange: retrospective and prospective

Buchanan’s (1975a/2000, 1986a/2001, 1989a) institutional economics was founded on an original theory of ‘politics-as-exchange’. Bargaining in politics has been frequently understood in terms of exchange, notably in the context of distributional negotiations in clientelist and interest group politics (e.g. Robinson and Verdier, 2013; Weingast and Marshall, 1988), but Buchanan uniquely conceptualized formal institutions emerging from exchanges between individuals analogous to the economist’s understanding of markets emerging from the exchange of goods and services.

In Buchanan’s conceptualization, politics was a forum in which people came together to reconcile self-interested ends. The mutual reconciliation of different interests in politics was an exchange like those that took place in markets: ‘Politics is a structure of complex exchange among individuals, a structure within which persons seek to secure collectively their own privately defined objectives that cannot be efficiently secured through simple market exchanges’ (see also: Brennan, 2012; Buchanan, 1986c/1999: 461; Holcombe, 2020: 374–375; Marciano, 2009; Vanberg, 2004, 2018).

The idea of politics-as-exchange was present either implicitly or explicitly throughout Buchanan’s work. In setting out the conceptual framework of The Calculus of Consent, arguably the foundational text of public choice theory, for example, Buchanan and co-author Gordon Tullock wrote:

Men co-operate through exchange of goods and services in organized markets, and such co-operation implies mutual gain… At base, political or collective action under the individualistic view of the State is much the same. Two or more individuals find it mutually advantageous to join forces to accomplish certain common purposes. In a very real sense, they ‘exchange’ inputs in the securing of the commonly shared output (Buchanan and Tullock, 1962/1999: 18).

The parallel conceptualization of economic and political exchange informed Buchanan’s (2007: 17) statement that ‘the simple exchange of apples and oranges between two traders’ was ‘the starting point for all that I have done’.

Politics-as-exchange was applied retrospectively and prospectively. It was used retrospectively as a theoretical explanation of the existence of real-world institutions and prospectively to guide institutional reform consistent with liberal constitutionalism.

Retrospectively, all formal institutions were conceptualized as founded on exchange. A minimal state, a modern welfare state and an autocracy, for example, were all conceptually reducible to bilateral agreements that came to encompass the whole of society to constitute a social contract (Buchanan, 1975a/2000, 1977, 1987/2001).

These bilateral agreements were the process via which rights were created. Buchanan rejected the existence of natural rights or natural law. Rights were necessarily political, reciprocally assigned and recognized by individuals, and prior to those agreements no rights existed (Brennan and Buchanan, 1985/2000: Chapter 2; Buchanan, 1975a/2000: Chapter 1, 1977: Chapter 1; see also: Holcombe, 2020; Meadowcroft, 2011: 50–52; Vanberg, 2004).

Although Buchanan conceptualized all formal institutions arising from exchange, politics-as-exchange was not a literal description of the origins of existing institutions. Buchanan (1975b: 89) recognized no actual state had arisen via conscious, deliberate agreement. Rather, politics-as-exchange was a conceptual explanation of institutional emergence explicitly located in the Enlightenment tradition of Bernard Mandeville, David Hume and Adam Smith, wherein social order emerged spontaneously from human actions that did not intend this outcome (Buchanan, 1975a/2000: 47–48, 2005: Chapter 6; Levy and Peart, 2020: Chapter 2).

Unlike his 18th-century predecessors, however, Buchanan (1975a/2000: 211) thought an emergent order was as likely to be characterized by oppression and exploitation as freedom and prosperity: ‘I have no faith in the efficacy of social evolutionary process. The institutions that survive and prosper need not be those that maximize man’s potential. Evolution may produce social dilemma as readily as
social paradise’. Politics-as-exchange explained but did not normatively justify the existence of oppressive or exploitative institutions.

It was necessary to explain the emergence of ‘bad’ institutions given their predominance in the real world. Buchanan (1997: 225) argued, ‘No existing or proposed political constitution contains sufficient constraints or limits on the authority of the agencies of government over the activities of individuals and groups’ to protect citizens from predation via the state. Hence, ‘There is no liberal constitution in existence or in prospect’ and, ‘In this sense, all existing constitutions are failures’ (Buchanan, 1997: 225, emphasis in original).

The outcome of political exchange was determined by the equality, homogeneity and reasonableness of participants. Individuals with equal resources and capacities were likely to agree an equal assignment of rights, whereas people with very unequal capacities and resources would make agreements that reflected their inequality. Societies with sharp racial, religious or class divisions were more likely to produce outcomes that entrenched the power of one group over others compared to more homogenous societies or those without organized factions. Agreements anticipated to emerge from a state of nature in which ruthless individuals predated against others differed markedly from the more equitable outcomes imagined where strong norms of civility existed (Buchanan, 1975a/2000, 1979a/1999, 1981/2001, 1988b; Congleton, 2014: 53–54; Haeffele and Storr, 2018; Munger, 2020; Wagner, 2017: Chapter 7).

Accordingly, in The Calculus of Consent Buchanan and Tullock (1962/1999) modelled relatively equal, homogenous and reasonable individuals agreeing constitutional rules that were essentially those embodied in the US Constitution at the American foundation. Buchanan and Tullock (1962/1999: 81) explained: ‘our analysis of the constitution-making process has little relevance for a society that is characterized by a sharp cleavage of the population into distinguishable social classes or separate racial, religious, or ethnic groupings’. An idealized liberal constitutional order was hypothesized to emerge only from a starting point of relative equality, homogeneity and reasonableness.

As noted in the Introduction, Pateman (1988: 60–62) argued Buchanan unwittingly exposed the pathologies of contractarianism by showing that absent pre-existing equality a social contract justified and formalized the domination of the strong over the weak. Buchanan, she contended, like Thomas Hobbes before him, showed that without prior equality, ‘the original pact will establish a society of masters and slaves’.

Pateman was correct that Buchanan showed that very unequal individuals would agree a social contract that preserved their pre-agreement inequality, but Buchanan’s intention was to show this relationship between the basis on which people entered agreement and the outcome agreed. Pateman’s objection was compelling only if it was believed that a social contract should produce an egalitarian outcome, which Buchanan also thought would only emerge from a starting point of equality.

Moreover, as noted above, Buchanan’s theory was explanatory not justificatory – conceptualizing institutions emerging from bilateral agreements did not grant normative legitimacy to those institutions. Politics-as-exchange reflected and illuminated a world of deep social divisions in which power was often brutal and irresistible, necessitating compromise with those who wielded it.

But the fact that even the most egregious institutions could be conceptualized in terms of bilateral agreements showed these arrangements ‘retain[ed] characteristic elements that we associate with the precepts of individualism’ (Buchanan, 1975a/2000: 71). If the basis of agreement could be identified in every institution, then Pareto-superior alternatives also always existed and welfare-enhancing reform was possible even in the most unpromising circumstances (Buchanan, 1975a/2000, 1987/2001, 1988a, 1989b/1991).

Prospectively, politics-as-exchange was to guide practical reform to move society towards arrangements more consistent with the ends of its individual members given the constraints of inequality, heterogeneity and unreasonableness. Buchanan (1988a: 77) did not believe people should wait for change to emerge spontaneously from the bottom-up: ‘Acquiescence before the inevitability of spontaneous evolution is a stance that holds, for me, little appeal’. It was the role of constitutional entrepreneurs to use the tools of political-economic analysis to broker new Pareto-superior exchanges.
Politics-as-exchange required institutional reform should have the agreement of all rights-holders and in this sense unanimous consent (Buchanan, 1966/2001, 1975a/2000: Chapter 9, 1977: Chapter 15, 2004). Unanimity was important ‘not because there is something sacrosanct about unanimity per se, but for the simple reason that it provides the only criterion through which improvements in rules and institutions can, in fact, be judged without the introduction of an explicit value scale’ (Buchanan, 1966/2001: 257 emphasis in original).

The prospective application of the principle of unanimity would herald a new era of consensual politics: politics would cease to be a competition for power over others and instead become a consensual enterprise. Consent was a necessary and sufficient condition for institutional reform (Buchanan, 1966/2001, 1975a/2000: Chapter 9, 1977: Chapter 15).

Buchanan recognized, however, that stasis would follow if every political decision required the explicit consent of every individual. To avoid stagnation, constitutional rules should be unanimously agreed, but this would include the agreement of less-than-unanimous decision-making rules to govern post-constitutional politics (Buchanan, 1975a/2000, 1977; Buchanan and Tullock, 1962/1999). Even in constitutional choice, however, if ‘“anti-social” or unreasonable individuals’ thwarted welfare-enhancing proposals then ‘the absolute unanimity rule must be broken’ (Buchanan, 1959: 134–135). This placed ‘an additional responsibility upon the political economist’ who ‘is forced to discriminate between reasonable and unreasonable men in his search for consensus’. For Buchanan (1959: 135) ‘the genuinely unreasonable individual can be readily identified’ because prospectively unreasonableness described opposition to change that was a strong Pareto-improvement. Third parties should not be permitted to block change that did not directly alter their rights; constitutional entrepreneurs should ignore unreasonable individuals who wilfully opposed a Pareto-superior proposal.

In practice, unanimity meant the use of compensation payments to make reforms Pareto-superior. Buchanan and Tullock (1962/1999: 90) wrote: ‘The unanimity test is, in fact, identical to the compensation test’. A compensation scheme that ensured all individuals benefitted from reform was equivalent to unanimous consent, though Buchanan (1977: Chapter 10) rejected the standard application of Kaldor–Hicks efficiency wherein the theoretical possibility of compensation made a reform Pareto-superior. The actual payment of compensation was necessary to establish unanimity, so the constitutional entrepreneur ‘must necessarily enter into the sometimes grubby world of debate, discussion, compromise, bargains, long-range deals, logrolling, package arrangements, and side payments’ (Buchanan, 1977: 140; see also: Brennan and Munger, 2014: 336).

It has been widely recognized, not least by Buchanan himself (1975b, 1988b, 2004), that the requirement of unanimity among rights-holders gave the status quo a privileged position because it maximized decision-making costs making change more difficult to bring about compared to non-unanimous decision-making. Even reforms that ameliorated an obvious wrong would be difficult to achieve if change could be blocked by rights-holders who were the beneficiaries of present arrangements or the perpetrators of a past injustice (Munger, 2018a, 2018b; Munger and Vanberg, 2023; Reisman, 2015: Chapter 6; Samuels, 2004).

Buchanan (1988a: 85) claimed his insistence on unanimity simply recognized the incontrovertible fact that consensual politics must begin from the here and now: ‘quite simply, there is no other place from which to start. This existential acceptance of the status quo… has no explicit normative content and implies neither approbation nor condemnation by any criteria of distributive justice’. As a prospective theory politics-as-exchange was only concerned with the possibility of unanimous reform, not the origins or justice of the starting point of change.

To evaluate politics-as-exchange the next section will place the linked examples of slavery and abolition within the framework to show the difficulties that arise from the combination of its retrospective and prospective dimensions. As described in the Introduction, slavery is a particularly useful example because as the most extreme institutional case it provides the most exacting test of the practical implications of Buchanan’s theory.
Slavery and abolition within politics-as-exchange

Buchanan (1975a/2000: 77–78, 1987/2001, 1988a, 1991/1999, 1995a, 2003) repeatedly invoked slavery in his presentation of politics-as-exchange as a retrospective explanation of institutional emergence. Buchanan (1975a/2000: 78) wrote that where ‘personal differences are sufficiently great’ and ‘some persons have the capacity to eliminate others of the species’, then ‘something similar to the slave contract’ that gave the strong ownership of the product of the weak and enabled the weak to escape death was logically mutually advantageous and therefore could be conceptualized as a bilateral agreement. This contract was a strong Pareto-improvement in a state of nature characterized by deep inequalities and ruthless individuals willing to predate on others. The agreement established the rights of individuals as master and slave: ‘A contract of slavery would, as other contracts, define individual rights, and, to the extent that this assignment is mutually accepted, mutual gains may be secured from the consequent reduction in defense and predation effort’ (Buchanan, 1975a/2000: 78).

For Buchanan (2003: iv–v), slavery was an example of an institution that was obviously unfair but might be agreed as a least-worst option by individuals concerned with their utility: ‘A social order that includes the institution of slavery, in any form, violates elemental norms of fairness in a self-evident sense’, but nevertheless, ‘Slavery may indeed, in some forms, maximize expected utility’.

Slavery was anticipated to arise from the state of nature when callous individuals held sufficient power to force others to serve their ends and the enslaved accepted this fate in preference to death. The existence of slavery reflected the cold fact that, ‘In a strictly personalized sense, any person’s ideal situation is one that allows him full freedom of action and inhibits the behavior of others so as to force adherence to his own desires’, and in this sense, ‘each person seeks mastery over a world of slaves’ (Buchanan, 1975a/2000: 117).

The conceptualization of slavery as a bilateral agreement in which a person accepted enslavement to avoid an even worse fate at the hands of a more powerful and brutal adversary echoed the argument made by classical social contract theorists, notably Hugo Grotius (1625/2001: Book IIII, Chapters 4 and 7), Hobbes (1651/1985: Chapter XX) and John Locke (1690/1960: Chapter IV), that prisoners that they may receive life

The existence of slavery reflected the cold fact that, ‘In a strictly personalized sense, any person’s ideal situation is one that allows him full freedom of action and inhibits the behavior of others so as to force adherence to his own desires’, and in this sense, ‘each person seeks mastery over a world of slaves’ (Buchanan, 1975a/2000: 117).

Buchanan (1987/2001: 171) rejected the idea that slavery was a voluntary agreement, but he nevertheless thought that the decision to enter slavery was comparable to other bilateral agreements in which people accepted an assignment of rights that limited their freedom – such as the acceptance of compulsory taxation in a democratic polity. Buchanan (1987/2001: 171) asked: ‘Why is a slavery contract… different from any other contract? Why do all of us… think that the slave contract is not normally an ethically legitimate embodiment of voluntary exchange?’. Buchanan’s answer was that slavery was generally considered unique ‘because it does not allow for a viable post-contract exit option’. Buchanan argued, however, that slavery was one of many institutions that provided little practical opportunity for exit. In a majoritarian democracy, people were similarly ‘unable to escape the extractions demanded’. Although the coercion inherent to majoritarian democracy was rarely acknowledged, Buchanan claimed: ‘We are… [in a majoritarian democracy] all “slaves without masters”’. No single master forced people to work, but individuals were nevertheless subject to an irresistible power that extracted the proceeds of their labour.

Buchanan (1991/1999: 290–291; 1995a: 142–143) also argued slavery could be conceptualized as an exchange if the slave retained some measure of control over their present and future life. Buchanan (1995a: 142–143 emphasis in original) argued when an enslaved individual had some ‘minimal sphere of residual individual control’ that made possible future gains from trade between master and slave that were dependent ‘on the separate choice behavior of both the master and the slave’, then the observed ‘social state’ had the characteristics of an exchange. By contrast, an individual forced to work for another ‘with no behavioral discretion at all’ could not vary their effort or compliance,
and ‘was analogous to a mechanical device, say a clock’, leaving no basis for future negotiations to achieve mutually advantageous improvements.

Buchanan aimed to hypothetically establish slavery as an agreement, and, surely paradoxically, illuminate the agency of the enslaved by showing that enslavement followed an evaluation of anticipated costs and benefits and that future gains from trade between master and slave were often possible. As Geoffrey Brennan and Michael Munger (2014: 341) described, at the heart of Buchanan’s institutional economics was, ‘a restless and intentional human agency that seeks to discover all opportunities for mutual improvements’.

Buchanan (1975a/2000: 78) accepted that this analysis ‘may seem to represent a somewhat tortuous interpretation of slavery as an institution, but it is explicitly designed to allow the analytical framework developed here to be fully general’. It was necessary to present slavery as a bilateral agreement to show that all formal institutions could be placed within the exchange paradigm – even the most exploitative and oppressive.

Throughout his published work Buchanan made only one direct reference to the actual slave system that had existed in America, writing that the ‘immorality of slavery… provided the impetus’ for the Civil War (Buchanan, 1990/2001: 104). He did not address the prospective implications of politics-as-exchange for abolition, but it has been argued that slavery exposes the problems that follow from his insistence that reform should have unanimous consent irrespective of the nature or origins of the status quo. As rights-holders, slaveowners would be directly impacted by abolition, so within Buchanan’s framework they were not ‘unreasonable’ individuals whose opposition to reform could be ignored by constitutional entrepreneurs, and therefore it has been argued that logically their consent was required for abolition (Meadowcroft, 2020: 397–400; Mueller, 2003: 143; Reisman, 2015: 121–122).

The requirement that reform had unanimous consent was intended to provide an escape from the ‘violence trap’ where individuals prepared to use violence to maintain their socioeconomic position locked society in a destructive cycle of conflict in which social change could only occur through violence. A ‘deal with the devil’, that is, an agreement with the perpetrators or beneficiaries of present injustice, involving, for example, the payment of compensation, may enable a society to move to a new peaceful equilibrium and thereby initiate a new era of consensual politics (Cox et al., 2019; Hamlin, 2018; Munger, 2018b; Munger and Vanberg, 2023).

There is a precedent of such a deal to peacefully end slavery in the British government’s payment of compensation to slaveowners prior to abolition in its territories in 1833 that arguably avoided a conflict like the American Civil War (North, 1990: 85). It is, of course, morally troubling to pay slaveowners to end their inhumane practice, effectively recognizing the right to own slaves and rewarding the perpetrators of exploitation and oppression. The fact that this compensation has been an important source of inherited wealth and privilege in Britain, and remains so nearly two centuries later, may heighten such moral concerns (Draper, 2010; Hall et al., 2014). It might be a mitigating factor that this compensation was a transfer between beneficiaries of an unjust system, given that the taxpayers who funded it had knowingly consumed cheap goods produced by slave labour (Levy and Peart, 2005: 124–127, 2018: 14–15), but this cannot assuage all moral qualms. Nevertheless, if compensation enabled slaves to be freed while avoiding the horrors of war then it might still be judged a politically expedient, least-worst course of action.

But if the unanimity principle enabled slaveowners to hold-out against reform then it would perpetuate, rather than end, violence. As David Gauthier (1986: 190–191) wrote in the context of his own contractarian theory, slavery may be theoretically framed as a Pareto-superior agreement, but it should be remembered that in reality slaves would walk away from their masters if their chains were removed and guards dismissed. Given the violence inherent to slavery, if slaveowners could prevent abolition by withholding their consent then this entrenchment of the status quo would maintain a violent equilibrium, conceivably engulfing generation after generation of enslaved people.

1Franklin and Schweninger (1999) provide the definitive empirical account of attempts by slaves to escape from their owners in the American South (and also of attempts by owners to recapture runaway slaves).
In a direct response to Gauthier, Buchanan (1988a: 84) argued the coercion attributed to slavery originated in the anarchical situation that was ended by the agreement of the two parties to formalize their relationship as master and slave. This contract, Buchanan claimed, ‘permanently preserve[d] the preagreement advantage of the master’, with the weaker ‘enslaved only because of some inability to enforce more favorable terms of existence’, so that, ‘The slave does not, presumably, possess a viable exit option, one that would allow him to carry on with an independent and isolated existence’. Buchanan (1988a: 84–85) argued in the absence of a feasible alternative slavery should not be defined as coercive: ‘There is no benchmark of independent existence that will define the presence or absence of coercion… If a slave cannot survive independently, can he be said to be coerced?’ If the inability to live independently occurred because the slave ‘has been captured against his will’, then, ‘Despite our civilized sense that the master’s act of enslavement is unjust, hardheaded analysis here must conclude that independent existence for the slave was not feasible, and that any such existence was fantasy, given the presence of the potential master’. For Buchanan, a bilateral agreement of rights could only be judged coercive in comparison to a feasible alternative – a slave contract was not coercive because the master had removed the feasible alternatives that would enable such a comparison.

There is logic to Buchanan’s claim that the evaluation of a social state must be comparative with the appropriate comparisons being the feasible alternatives for the individuals therein. But the linked cases of slavery and abolition reveal troubling implications of this position. Buchanan’s retrospective conceptualization of slavery as a bilateral agreement logically meant slaveowners had a right to own slaves that then gave them a veto over abolition when it met the prospective requirement that reform should have the consent of all rights-holders. As such, within Buchanan’s framework individuals could only leave exploitative arrangements with the consent of their oppressors – a self-defeating conclusion for a liberal constitutional theory. The next section will set out the dissonance between the Kantian and Humean aspects of Buchanan’s theory that led to this unsatisfactory position.

**Kantian individualism and Humean subjectivism**

Buchanan (1977: 244) subscribed to ‘the basic Kantian notion that individual human beings are the ultimate ethical units, that persons are to be treated strictly as ends and never as means’. His institutional economics accordingly involved a commitment to moral equality that meant, ‘Each man counts for one, and that is that’ (Buchanan, 1975a/2000: 4) and it was therefore necessary to ‘refrain from imposing one’s own values on others’ (Buchanan, 2005: 69). As Brennan and Kliemt (2019a: 797) described, Buchanan inhabited a ‘world of economic Kantianism’ in which ‘economists should restrict their research to how individuals who respect each other as separate persons can pursue their several ends via mutual consent’ (see also: Kliemt, 2011; Levy and Peart, 2018, 2020).

Buchanan also understood individual values and preferences to be necessarily subjective. In theorizing this subjectivism Buchanan (1989c/1991: 451–452) explicitly followed Hume’s rejection of the moral certitudes ‘previously offered by the dogmas of religion and reason’. There were no absolute moral truths or falsehoods supplied by either religion or philosophy, only subjective individual values, just as people had preferences for different colours of paint: ‘As an individual I may choose to paint my house either white or red’, and, ‘It becomes highly questionable, if not improper, to say that my selection of white as the color for my house is true while the alternative red is false’ (Buchanan, 1977: 69).

It was the responsibility of each individual, in concert with others, to devise their own ethical code to guide their conduct – to frame their own personal ‘constitution of our values’ (Buchanan, 1989c/1991: 452). Value relativism, and the possibility of resolving interpersonal conflict by agreement, made a social order founded on consent conceivable: ‘the recognition of the necessary relativism and individualism of values, along with the modelling of politics as the enterprise of resolving conflict among such values… makes the libertarian social order meaningful’ (Buchanan, 1986b/2001: 169–170).

Buchanan (1975a/2000: 4) believed agreement was the only evaluative standard consistent with his individualism and subjectivism: ‘A situation is judged “good” to the extent that it allows individuals to get what they want to get, whatsoever this might be, limited only by the principle of mutual
agreement’. Buchanan (1975a/2000: 3) argued ‘any attempt to lay down [external] standards becomes effort largely wasted at best and pernicious at worst’. It was effort wasted because there was no objective truth that could form the basis of such standards; it was pernicious because individuals enlisted in a political quest for truth were inevitably conscripted to work for the ends of others (see also: Berggren, 2016).

The principle of moral equality itself might be thought to be a normative standard external to individual actors that therefore implied the existence of an objective truth, natural law or natural rights. Indeed, Buchanan (1995a: 142) acknowledged that his individualism rested on what were essentially basic rights: ‘My basic proposition is that, in any social setting, individuals have rights, by which I mean that individuals retain personal control over actions along at least some minimal set of dimensions of behavioral adjustment’. For Buchanan (1995a: 142) ‘this elementary proposition’ of individual behavioural autonomy was ‘a positive description of social reality’ and hence a statement of social ontological reality rather than a normative assertion of natural rights or natural law. Buchanan’s statement that individuals had rights meant individuals had the capacity to create rights in their interactions with one another (Buchanan, 1995a: 142, 1977: Chapter 1).

Kantian individualism has been generally assumed incompatible with slavery wherein individuals are used as means for the ends of others (e.g. Anderson, 1999: 319; Ellerman, 1988: 1112–1113). Immanuel Kant (1796/1887: 119) himself rejected the idea that a person could agree to become a slave: ‘a Contract by which the one party renounced his whole freedom for the advantage of the other, ceasing thereby to be a person and consequently having no duty even to observe a Contract is self-contradictory and therefore of itself null and void’.2

As set out in the previous section, Buchanan (1975a/2000, 1987/2001, 1988a, 1991/1999, 1995a) rejected such an inference from his individualism because he believed forbidding individuals to agree a slave contract imposed external values on the individual participants in agreement that denied their agency and violated their moral integrity. But this reasoning led to the problematic conclusion that slavery was a bilateral agreement of reciprocal rights and abolition logically required the consent of slaveowners.

Buchanan thought his individualism and subjectivism were mutually reinforcing – both followed from his ontology–methodology and informed his view that the attempt to impose external values was fallacious and potentially tyrannous. But there was a subtle but crucial disjuncture between the two. Buchanan’s individualism implied people were ends not means and formal institutions were conceptually reducible to bilateral agreement; his subjectivism denied the existence of objective values transcendent of individuals and ruled out external limits on agreement. These two positions conflicted, however, because in the absence of any external constraint on agreement individuals could be imagined agreeing to arrangements in which they became means for the ends of others.

Buchanan’s individualism respected the integrity of the individual, but his subjectivism countenanced the most egregious violations of individual liberty. Individuals could enter institutions in which they became means for the ends of others if agreement happened to settle on such an outcome; and those arrangements could only be altered with the consent of all rights-holders, even those who had used violence to obtain rights over others.

As noted in the Introduction, Barry (1984) and Holcombe (2014, 2020) argued the crucial tension in Buchanan’s work was between his classical liberalism and his contractarianism. Barry (1984: 595) argued Buchanan’s failure to give his classical liberal Kantian ethics ‘normative bite’ in his contractarian theory led to nihilism wherein any institutional outcome could be imagined resulting from agreement, allowing, for example, ‘slavery to masquerade as freedom and theft to constitute a title of property’. Barry (1984) and Holcombe (2014, 2020) were correct that Buchanan failed to adequately infuse classical liberal moral considerations into his work, but the conflict within his work was in fact

2Though Kant (1796/1887: 198) did think a person could be legitimately enslaved as punishment for a crime by which they forfeited their rights.
between his Kantian individualism and Humean subjectivism – constituent parts of both his classical liberalism and contractarianism.

This problematic inconsistency at the heart of Buchanan’s institutional economics can be resolved, however. The next section will show that Buchanan’s ontology was consistent with a solution to the internal inconsistencies of his work derived from the liberal theories of Kukathas (2003) and Nozick (1974).

Exit, freedom and ontology

The dissonance between Buchanan’s individualism and his subjectivism reflected the fundamental division in liberal thought identified by Kukathas (2003: 19) between the belief that a liberal society involved the universal acceptance of negative rights, essentially Kantianism, and the more Humean view that a liberal society was a framework within which different, even potentially illiberal, values and practices co-existed. A successful liberal theory must reconcile these two positions.

Kukathas’ (2003) vision of a ‘liberal archipelago’, in which a liberal society was imagined as a sea with many different islands, representing the existence of multiple, diverse political associations, had similarities with Buchanan’s institutional pluralism. Unlike politics-as-exchange, however, Kukathas’ over-arching framework included the right to exit from any community or association. Kukathas recognized that a liberal social order necessitated the right to leave illiberal arrangements and practices, so that, ‘dissenters were not obliged to remain in associations which refused to tolerate them. There must be, at least in principle, the possibility of individual exit from illiberal communities or associations’ (Kukathas, 2003: 25).

A similar rendition of institutional pluralism appeared in Nozick’s (1974: 298) ‘framework for utopia’. In this thought experiment every individual imagined their best possible world and could choose to live in their imagined world or one conceived by another person, with those worlds able to attract sufficient people becoming viable. Importantly, in Nozick’s framework every person had the right to exit from any world – no one could be forced to live in a community or association against their wishes.

Buchanan (1975b, 1987/2001: 171–172, 2005: 58) recognized exit was a powerful guarantor of individual liberty, enabling people to escape unwanted associations and disincentivizing the exploitation of those able to leave a political community. His work on federalism, in particular, acknowledged the importance of opportunities for ‘internal exit’ to enable individuals to withdraw from some or all government services without moving geographically (Buchanan, 1995b, 1996; Buchanan and Faith, 1987; see also: Kurrild-Klitgaard, 2002).

For Buchanan, however, to assume a right prior to political agreement, even a right of exit from exploitative or oppressive institutions, imposed external values on the participants in agreement and therefore should be rejected. Buchanan himself might favour federal structures that facilitated internal exit, but he was steadfast in his view that the economist should only use their expertise to advise constitutional entrepreneurs on the implications of different institutions, not impose their own preferences on others (Buchanan, 1975a/2000: Chapter 1, 1977). Hence, Buchanan (1975b: 121) praised Nozick’s use of the economic model of clubs ‘with voluntary entry and exit’ to facilitate ‘the comparative evaluation of idealized social orders’ but argued Nozick’s framework was ultimately unpersuasive because it assumed the existence of rights without explaining how those rights came about (see also: Brennan and Kliemt, 2019b).

In reaching this conclusion, however, Buchanan did not consider the possibility that a right of exit need not follow from a normative or empirical truth-claim but could be consistent with from his own ontology. Buchanan’s (1979b/1999: 247) ontology ultimately rested on an account of human nature that contrasted ‘human beings’, who can, ‘within limits, shape the form of being that we shall be between now and the time of death’, with non-human animals like dogs that were conditioned and limited by their innate characteristics and external stimuli. A dog that grew up in a harsh environment would develop its more aggressive nature, while one raised with affection would become more placid, though both could be trained to change their behaviour. But a dog could not independently decide to change its character and become a different, better dog. Humans similarly possessed inherited
characteristics that were further shaped by their environments, but people also had the unique capacity to choose to learn new, different ways of life. Humans were natural and artefactual beings – possessing innate characteristics but also capable of self-authorship. Buchanan (1979b/1999: 247) wrote: ‘We construct our own beings, again within limits. We are artefactual, as much like the pottery sherds that the archaeologists dig up as like the animals whose fossils they also find’.

The desire for liberty was an innate human characteristic that followed from artefactual man’s universal capacity to imagine a different, better life. Buchanan (1979b/1999: 259, emphasis in original) wrote: ‘Man wants liberty to become the man he wants to become’. The aspiration and potential for self-authorship extended to the political realm: ‘the individual’s desire for liberty from the coercive power of others’ offered ‘the exciting and normatively relevant prospect of a world in which all participants are free to choose’ (Buchanan, 2005: 58). As Paul Lewis and Malte Dold (2020: 1167, emphasis added) explained: '[Buchanan’s] social ontology or account of human nature that portrays man as an artefactual being underwrites a commitment to a liberal social order in which people are granted the freedom to forge and reforge their characters as they see fit’.

This artefactual nature led to individual behavioural autonomy that continually produced reciprocal agreements and exchanges with others. Buchanan (1975a/2000: 8) wrote: ‘Men and women manage to walk along city pavements. With rare exceptions, they respect queues in supermarkets, in banks, and in airports’. When two people passed on the street they respected one another’s rights to walk along the pavement unmolested; when a person joined a queue in a supermarket, they respected the right of those in front to be served first. Social outcomes were emergent properties of interpersonal exchange, not the result of one person’s independent choice of a desirable social state, and such exchanges were ubiquitous and unavoidable (Buchanan, 1975a/2000: Chapter 1, 1977: Chapter 1, 1995a).

Although it is tempting to think of political agreement solely in terms of ink signatures on parchment documents, politics-as-exchange also concerned the decisions of individuals to respect one another’s rights in their most prosaic interactions. Buchanan’s individualism reflected the ontological reality that individual actions continuously created, affirmed and reassigned reciprocal rights (Buchanan, 1975a/2000: Chapter 1, 1977: Chapter 1, 1979b/1999).

Accordingly, Buchanan’s ontology implied limits on individual action even without the formal agreement of rights. Buchanan (1977: 22) wrote that his rejection of natural rights ‘does not imply that the legally defined rights of individuals, and the distribution of these rights, are arbitrarily determined by the political authorities’. The belief that rights were political did not mean rights were solely dependent on the existence of formal political institutions.

It followed that a person whose wallet was stolen in Central Park experienced ‘genuine coercion… because the victim, the thief, and external parties agree and accept the property rights. The wallet was the victim’s by right of assigned and acknowledged ownership’ (Buchanan, 1975a/2000: 56). It was generally understood that a wallet in a person’s pocket belonged to that individual, just as it was generally understood that a person stood in line ahead of others should be served first. In both cases, the relevant rights did not depend on the existence of formal institutions.

Buchanan’s account of the rights continually assigned and reaffirmed in the day-to-day interactions of artefactual individuals logically implied that taking a wallet in a street robbery, pushing ahead of a person in a queue and enslaving a person by placing them in chains, were violations of the moral integrity of the persons who were thereby harmed, irrespective of what formal rights existed. Buchanan’s ontology was therefore irreconcilable with the idea that agreements made following the coercive removal of feasible alternatives should only be altered with the consent of those who had used violence to assume power over others.

A right of exit like those in Kukathas’ (2003) ‘liberal archipelago’ and Nozick’s (1974) ‘framework for utopia’ would operationalize these ontological considerations. Artefactual individuals with behavioural autonomy continuously make and remake agreements with one another that imply the ability to independently decide to enter and exit different institutional arrangements.

A right of exit would necessitate distinguishing between exchanges with and without the sword – between exchanges that are genuinely consensual and those that result from coercion. In practice this
would necessitate theoretical consideration of whether individuals would remain within an institution if they were able to walk away.

Importantly, a right of exit would not establish a precedent allowing third parties (such as governments) to unilaterally alter further rights because interventions in bilateral agreements beyond a right of exit would remain incompatible with Buchanan’s ontology and therefore would be ruled out. Allowing slaves to be freed without the consent of their masters would not unleash a Leviathan without limit on its power. Rather, exit would extend the ontological boundaries of politics-as-exchange one step beyond the understanding that all value resided with the individual participants in agreement to also recognize that those individuals should be free to leave arrangements they find intolerable.

To be fully general politics-as-exchange should explain the existence of oppressive and exploitative institutions, recognizing the limits placed on human freedom by deep inequalities, social division and the cruelty of many who hold power. But it should not do so by falsely claiming slaves consent to their enslavement. Buchanan thought consent was a necessary and sufficient condition for institutional reform that was essentially synonymous with ‘agreement’, but his conceptualization of political exchange without a right of exit drained the concepts of consent and coercion of meaning.

Of course, a right of exit would not magically solve all the challenges of basing politics on consent. There remain important questions as to whether individuals should have a right of entry to alternative associations and whether third parties have a duty to rescue the oppressed that are beyond the scope of the present article (see: Crowder, 2007; Fried, 2003; Kukathas, 2012). But a right of exit would move Buchanan’s liberal constitutionalism from a position of apparent compatibility with oppressive and exploitative institutions to instead expose arrangements that rest on coercion.

Buchanan’s work was remarkably consistent throughout his career – characteristically his account of slavery within politics-as-exchange did not change over many decades of scholarship. Unusually, however, Buchanan’s (1965/2000) first published comments on slavery in his foreword to Tullock’s The Politics of Bureaucracy described slavery in markedly different terms to his work that followed. Here Buchanan (1965/2000: 66) wrote that slavery was an example of ‘the exchange process, with the things exchanged being “slaves”’. This claim that slavery was an exchange of persons, not an exchange between people, implied it was intrinsically coercive. This understanding was absent from Buchanan’s later work, but the inclusion of a right of exit within politics-as-exchange would expose those illiberal institutions in which people are exchanged as if they are things.

Conclusion
James Buchanan’s politics-as-exchange retrospectively theorized institutions emerging from the bottom-up as a result of bilateral agreements and prospectively guided constitutional entrepreneurs to broker institutional reform that had the consent of all rights-holders to initiate a new era of consensual politics. Theorizing politics as an exchange illuminated the individualist foundations of existing institutions and also revealed the limits on human freedom imposed by the social constraints of inequality, heterogeneity and unreasonableness.

But the linked examples of slavery and abolition demonstrate that illiberal conclusions can follow when the principle of unanimity is applied irrespective of the nature or origins of the status quo. When the retrospective conceptualization of slavery as a bilateral agreement meets the prospective requirement that reform should have the consent of all rights-holders then logically slaveowners have a veto over abolition and the principle of unanimity perpetuates a violent equilibrium.

These pathologies of Buchanan’s theory resulted from dissonance between his individualism and his subjectivism, or between the Kantian and the Humean aspects of his project. Buchanan’s individualism demanded respect for the integrity of the individual, but his subjectivism allowed individuals to become means for the ends of others. Without a right of exit like those in Kukathas’ (2003) ‘liberal archipelago’ and Nozick’s (1974) ‘framework for utopia’, Buchanan’s subjectivism could countenance illiberal outcomes with little prospect of emancipation. But a right of exit can be grounded in
Buchanan’s ontology, recognizing that artefactual individuals with behavioural autonomy to make and remake reciprocal agreements of rights should be able to leave arrangements they find intolerable. This would necessitate differentiating between exchanges with and without the sword – between political agreements that rested on consent and on coercion.

With a right of exit, politics-as-exchange may inform an ontological defence of constitutional democracy consistent with Buchanan’s (2005: 58, emphasis in original) conviction that a social order is possible in which ‘we can all be free’. A liberal constitutional order should respect the moral integrity of the individual, appreciate the subjective nature of human experience and prevent individuals and groups using political power to impose their truth-judgements on others. This social order can be conceptualized emerging from bilateral agreements, but only when individuals are free to exit from exploitative and oppressive arrangements. A successful liberal theory must place subjectivism within an immutable framework of individualism.

Acknowledgements. An earlier version of this article was presented at the Public Choice Society Meetings, Nashville, TN, USA, on 11 March 2022. I thank my discussant, Jonathan Plante, and the attendees for helpful comments. I also thank Samuel DeCanio, Paul Lewis, Liya Palagashvili, Virgil Storr, Richard Wagner and the journal’s four anonymous referees for their comments on earlier versions.

References


Munger M.C. (2018b) Starting from where we are: the importance of the status quo in James Buchanan. In Wagner RE (ed.), *James M. Buchanan*. Basingstoke: Palgrave, 39–64.


---

**Cite this article:** Meadowcroft J (2023). Exchanges with and without the sword: slavery, politics-as-exchange and freedom in James M. Buchanan’s institutional economics. *Journal of Institutional Economics* 1–15. https://doi.org/10.1017/S1744137423000176