The Unwritten Theory of Justice: Rawlsian Liberalism Versus Libertarianism

Barbara H. Fried
Stanford Law School

Research Paper No. 1957828

Stanford Public Law and Legal Theory Working Paper Series

Stanford Law School
Crown Quadrangle
Stanford, California 94305

This paper can be downloaded without charge from the Social Science Research Network Electronic Paper Collection http://ssrn.com/abstract=1957828
The Unwritten Theory of Justice: Rawlsian liberalism versus libertarianism

(Forthcoming, The Blackwell Companion to Rawls, Jon Mandle and David Reidy, eds., 2012)

When *A Theory of Justice* was published in 1971, utilitarianism was the game to beat in political philosophy, and Rawls made clear his intention to beat it. The appearance of Nozick's *Anarchy, State and Utopia* three years later singlehandedly enshrined libertarianism rather than utilitarianism in the popular imagination as the chief rival to Rawls's two principles of justice. Ever since, Rawlsian liberalism has had two parallel lives in political theory. The first—the version Rawls wrote—is a response to utilitarian’s failure to take seriously the separateness of persons. The second—the unwritten version ‘received’ by its general audience—is a response to libertarianism’s failure to take seriously our moral obligations to the well-being of our fellow citizens. This article considers how, had he written the second version, Rawls might have dealt with libertarians’ critique of ‘justice as fairness’ as fundamentally illiberal, and how his two principles might have been transformed in the process.
The Unwritten Theory of Justice: Rawlsian liberalism versus libertarianism

(Forthcoming, The Blackwell Companion to Rawls, Jon Mandle and David Reidy, eds., 2012)

Barbara H. Fried
William W. and Gertrude H Saunders Professor of Law
Stanford University
bfried@stanford.edu

When A Theory of Justice was published in 1971, utilitarianism was the game to beat in political philosophy, and Rawls made no bones about his intention to beat it. In his words (Rawls 1999a, 513), “Justice as fairness... offers, I believe, an alternative to the utilitarian view which has for so long held the preeminent place in our moral philosophy. I have tried to present the theory of justice as a viable systematic doctrine so that the idea of maximizing the good does not hold sway by default.” While other theories of justice—perfectionism, intuitionism and rational egoism—make cameo appearances in the book, Rawls’s case for ‘justice as fairness’ rests in significant part on its claimed superiority to utilitarianism, by virtue of its refusal to “justify institutions on the grounds that the hardships of some are offset by the greater good of others.” (1971, 15). The word “libertarianism” does not appear in the book, although he presumably has a Smithian version of libertarianism in mind in his brief discussion of “a system of natural liberty.” (1971, 72-75).
Three years later, Nozick published *Anarchy, State and Utopia* (hereinafter *ASU*). Part II of *ASU*, which lays out Nozick's Lockean theory of the ideal libertarian state, is explicitly framed as a rejoinder to Rawls's *Theory of Justice* (hereinafter *TJ*), and fully a third of its argument is given over to a close reading and critique of the book, in particular the Difference Principle. The publication of *ASU* singlehandedly made deontological libertarianism a political philosophy to be reckoned with in academic circles, at precisely the same time j it was on the ascendency in political circles. For the general readership of *TJ*, it also singlehandedly enshrined libertarianism rather than utilitarianism as the chief rival to ‘justice as fairness,’ and put the Difference Principle at the center of Rawlsianism. Forty years later, to most nonspecialists ‘Rawlsianism’ *is* the Difference Principle, and the most durable part of Nozick's argument has proved to be his critique of that principle. As mismatched as they are in aspirations and method, *ASU* and *TJ* have become the towering bookends of political theory, each the other’s chief foil.

By the time that Rawls was working on the revised edition to *TJ*, libertarianism was well entrenched in the political scene, and Rawls surely knew that to his general readership, ‘justice as fairness’ and the many variants it spawned (Dworkin’s resource egalitarianism, Sen’s capability theory, Anderson’s democratic equality) were generally viewed as egalitarians’ most cogent answers to libertarianism. Rawls, understandably, chose not to reformulate his argument to take cognizance of the changed political and philosophical landscape, and hence the revised version of *TJ*, like the original one, is framed as a rejoinder to utilitarianism,
making no mention libertarianism or libertarian principles, beyond the same brief
discussion of “systems of natural liberty.” (1999(a), 57-58, 62-64).

As a result, Rawls’s theory of justice has had two parallel lives in political
theory. The first—the version Rawls wrote— is framed as an alternative to
utilitarianism, and in particular utilitarianism’s failure to take seriously the
separateness of persons and individuals’ right to pursue their own projects in life.
The second—the version ‘received’ by its general audience— is framed as an
alternative to libertarianism, and in particular libertarianism’s failure take seriously
our moral obligations to the well-being of our fellow citizens.

Notwithstanding the received view of TJ and ASU as the opposing poles of
contemporary political philosophy, over the last thirty–five years a small cottage
industry has emerged, dedicated to showing how little is required, logically
speaking, to turn Rawls into Nozick and Nozick into Rawls. (Lomasky, 2005; Stick,
1987; Buchanan and Lomasky, 1984; Roemer, 1988; Gibbard, 2000) From one
perspective, this development is hardly surprising, given the shared, foundational
commitment of Rawlsianism and libertarianism to a moral individualism grounded
in the view that we are all “free and equal moral persons” with the right to pursue
our own conceptions of the good. (Rawls, 1971, 22-27). As Loren Lomasky (2005,
180) put it, “Is it possible to deny the fundamentally libertarian flavor of a theory in
which [maximal liberty for all] enjoys lexical priority?”

My aim here is to explore where and why, starting from that common
commitment, Rawls’s ‘justice as fairness’ and libertarianism come apart. I focus on
their respective treatments of the just distribution of wealth, as this is the point at
which the two theories most clearly diverge. Among traditional (right) libertarians, in addition to Nozick I consider the work of (among others) Loren Lomasky, Jan Narveson, Eric Mack, James Buchanan, Israel Kirzner, Richard Epstein, Gerald Gaus, Randy Barnett, John Simmons, David Schmidtz, John Hasnas and Murray Rothbard.ii I also consider the arguments for “left libertarianism” offered by Michael Otsuka, Peter Vallentyne, Hillel Steiner and others. (Unless otherwise indicated, “libertarianism” refers to right libertarianism alone.)

In defending their respective theories of justice, both Rawls and libertarians (including self-identified deontologists like Nozick) move back and forth between prudential arguments—what will make peoples’ lives go better overall/produce a stable society-- and arguments from rights-based principles. I focus on the latter, for the simple reason that disagreements about the instrumental value of a given arrangement of property rights or distribution of wealth largely rest on disagreements about the facts. That peoples’ empirical assumptions tend to track their ideological predispositions is a significant, if depressing, fact about human nature. But for my purposes, if that is all that separates, say, Rawls and Nozick, there is no mystery about how to bring them together: change one or both of their empirical assumptions. In contrast, turning Rawls into a principled libertarian or Nozick into a principled Rawlsian egalitarian requires us to understand where, starting from a shared commitment to liberal individualism, they faced the same choices and chose differently.

For the most part, we don’t have to guess how libertarians would describe their points of divergence from Rawls. Much of the libertarian revival sparked by
ASU, like ASU itself, is framed as a rejoinder to Rawls. From the perspective of libertarians, Rawls has assumed the rhetoric of liberal individualism, but at every juncture has proved himself (in Loren Lomasky’s words) “more committed to his egalitarian redistributionist conclusions than he is to the [liberal individualist] premises that generate those results.” (Lomasky, 2005, 199) The net result, as far as libertarians are concerned, is that Rawls is a faux liberal individualist, and his disagreements with utilitarians are at most a friendly family squabble.

While Rawls did not address modern-day libertarianism in *TJ*, he did touch on the subject in other writings. But his comments are relatively brief, and focused exclusively on the Nozickean version, in particular Nozick’s derivation of the minimal state. (1993, 262-265; 1999b, 49-50). As a result, for the most part, we have to infer how he would have responded to the substantial libertarian critique of *TJ* that has amassed over the past four decades.

Finally, there is the question of method. In comparing Rawls to his libertarian counterparts, should we regard *TJ* as a procedural theory of justice or a substantive one? Rawls’s answer was, both: The two principles are to be tested by “find[ing] the arguments in their favor that are decisive from the standpoint of the original position,” from which vantage point there are “no given antecedent principles external to [the representative persons’] point of view to which they are bound.” But they are also to be tested “by comparison with our considered judgments of justice.” (1999a, 132) The general consensus at this point, however, is that the device of the original position is doing no independent work in deriving the
substantive principles of 'justice as fairness,' and at various points Rawls himself invites or offers the same assessment (1971, 141; 1993, 25-26; 1999a, 16).

On the other hand, given the central rhetorical role that the heuristic of rational choice plays, both in unfolding 'justice as fairness' and establishing Rawls's bona fides as a liberal individualist, one cannot ignore it.

Libertarians have pursued both procedural and substantive derivations of the just state, although (with the notable exception of Nozick) generally not in the same work. Most, however, have focused on the latter. As John Simmons (2005) has observed, contemporary libertarian political philosophers have said relatively little about how a state gains political legitimacy, and with it the right to coerce its citizens. The omission is surprising, because libertarian principles hardly imply indifference to the process by which individuals come under the coercive authority of the state, and most versions of libertarianism (whether they start from self-ownership, autonomy/liberty, concepts of mutual advantage) would seem to make individual consent a necessary precondition for legitimacy.

For anarcho-libertarians like Simmons (2005), Murray Rothbard (1977) and John Hasnas (2008), the explanation is straightforward. Believing that only explicit consent can justify subjecting an individual to the power of the state and that such consent cannot be obtained, they have concluded that the state cannot be justified.

The majority of libertarians, perhaps sharing anarcho-libertarians’ skepticism about consent but unwilling to accept their conclusion, have, sub silentio, given up on consent. Instead, they have argued that the state is legitimized insofar as its political arrangements, however arrived at, are just. (As discussed above, on
one view of TJ—probably the predominant view at this point-- this is precisely what Rawls has done. But a small band of libertarians has pursued the classical liberal social contractarian tradition, deriving the legitimacy of the state from hypothetical consent to its arrangements. Notable among these are David Gauthier (1986), James Buchanan (1975, 1984), Jan Narveson (1988) and Loren Lomasky (1987).

Nozick, is, perhaps ironically, the most famous expositor of both the second and third approach. Notwithstanding Nozick's waggish dismissal of the social contract as "not worth the paper on which it is not written," the Nozick of Part I of *ASU* pursues a procedural justification for the state that is essentially contractarian, although not labeled as such. The Nozick of Part II, abandoning procedural justifications for substantive ones, argues that the state is legitimate insofar as it conforms to just (Lockean) rights. In this respect at least, Rawls's and Nozick's projects parallel each other.

I start by comparing procedural derivations of the just state in Rawls and libertarians, and then turn to their respective substantive principles of justice.

A. **Constructing the Choice Position.**

Rawls (1999a, 156-57) explains the motivation for the original position as follows:

On the contract interpretation treating men as ends in themselves implies at the very least treating them in accordance with the principles to which they would consent in an original position of equality. For in this situation men have equal representation as moral persons who regard themselves as ends and the principles they accept will be rationally designed to protect the claims of their person.

Most contractarian libertarians would find little to quarrel with in this formulation. The disagreements all come from how Rawls on the one hand and
libertarians on the other construct the choice situation into which such persons are put and the resulting knowledge imputed to them.

By ‘persons,’ Rawls (1999a, 11, 152, 464) insisted he meant “determinate-persons,” each with his or her own “place in society, class position or social status, ... fortune in the distribution of natural assets and abilities, intelligence, strength and the like... psychological propensities .... [and] conception of the good.” This, more than any other feature of his contractualist thought experiment, was (in his view) what differentiated ‘justice as fairness’ from Harsanyi-like contractualist arguments for average utilitarianism, and establishes its superiority. To quote Rawls (1999a, 150, 152), “the utilitarian argument assumes that the parties have no definite character or will, that they are not persons with determinate final interests, or a particular conception of their good, that they are concerned to protect.” In short, “[t]hey are, as we might say, ‘bare-persons,’” each with the same chance of “being any one of a number of persons complete with each individual’s system of ends, abilities, and social position,” expressed by the “same deep utility function.”

But having populated the original position with “determinate-persons” motivated by rational self-interest, Rawls famously puts them behind the veil of ignorance, which prevents them from knowing what those determinate traits are. Rawls argues that his “determinate-person” behind the veil differs from the utilitarians’ ‘bare-person,” by virtue of the fact that he knows he is different from other individuals, though he doesn’t know how. As a consequence, while (contra utilitarianism) ethical considerations do not operate “in the characterization of the parties,” they operate in the characterization of the choice situation, by suppressing
the information required for parties to favor their own interests over others’.

(Rawls 1999a, 160, 166, 512) But to libertarians (and many others as well), it is a
distinction without a difference. As Nozick (1974, 228) famously responded in ASU,
turning Rawls on himself, the choosing self behind the veil satisfies Rawls’s ambition
(contra utilitarianism) to “‘take seriously the distinction between persons’…. only if
one presses very hard the distinction between men and their talents, assets, abilities,
and special traits.”

The resulting “bare deontological ego” (Loren Lomasky’s term) populating
Rawls’s original position may suffice to establish Rawls’s credentials as a moral
individualist in the limited sense that (unlike communitarian or other group-based
accounts of justice) it counts only individuals’ interests in evaluating institutional
arrangements. But from the perspective of libertarians, the gulf between that “bare
deontological ego” and the flesh-and-blood differentiated self at the heart of liberal
individualism is vast—indeed, unbridgeable. (Nozick, 1974, 228; Narveson, 1988,
132-33). As Lomasky (1987, 135, 138) put it:

> Everything pertaining to persons except their personhood, whatever that
could be when abstracted from ends, character, abilities, and relations to
material possessions, is thoroughly socialized…. The conclusion is thoroughly
illiberal.”

Rawls’s motivations for imposing the veil are, in the view of most
libertarians, thoroughly illiberal as well. Rawls defends the veil in (at least) two
different ways in TJ.

The first and most debated is that it prevents distributive shares from being
“improperly influenced by … factors so arbitrary from a moral point of view,”
including natural and social endowments and an individual's choice of ends. (Rawls, 1971, 59-60, 72). At this point, most libertarians conclude, 'justice as fairness' is simply not a liberal individualist theory. As Nozick, Lomasky and others have argued, Rawls here conflates two questions: whether people deserve their natural and social endowments; and whether they deserve to keep them and decide for themselves how to exercise them. The answer to the first question may well be ‘no.’ But because such endowments are what make individuals ‘determinate’ people rather than ‘bare-persons,’ “[a] robust liberalism must... maintain that contingencies become imbued with moral weight once they are intimately attached to the lives that persons actually live.” (Lomasky, 1987, 136, 138; Nozick, 1974, 216-227; Gauthier, 1974; Gauthier, 1986, 245-57).

Rawls's second justification is that without some sort of veil of ignorance, we cannot hope to reach agreement, because “we cannot reasonably expect our views to fall into line when they are affected by the contingencies of our different circumstances.” (1999a, 453, 465, 267-68; 1993, 53-54) The original position solves the problem by eliminating all such differences, thereby making “the deliberations of any one person ... typical of all.” Rawls (1971, 517) makes clear that eliminating the potential for disagreement is not just a happy byproduct of Rawls's having adopted the veil of ignorance for other reasons. It is, rather, one of the chief motives for adopting it: “Indeed, other things being equal, the preferred description of the initial situation is that which introduces the greatest convergence of opinion.”
At this point, libertarians (and many others) have concluded that ‘justice as fairness’ is not a contractualist theory either. Starting with the desideratum of unanimous agreement, Rawls, by his own acknowledgement, has reverse-engineered the original position to guarantee it. But surely, this puts the cart before the horse in any consent-driven theory of legitimation. If determinate persons, each calculating their rational self-interest with full knowledge of their actual circumstances in life, cannot reach agreement on a just state, then—as the anarcho-libertarians have concluded—so much the worse for agreement.

Finally, the consequences of imposing the Rawlsian veil are, from a libertarian perspective, thoroughly illiberal as well. By making “the deliberations of any one person ... typical of all,” Rawls turns the problem of social choice into a problem of individual choice under conditions of uncertainty, endowing that one representative person with preferences that reflect the aggregated preferences of all. As a result, what is ‘rational’ for the representative person to choose coincides with what an impartial observer, weighing every individual’s preferences equally, would choose—a result that reflects collectivist, not individualist, values. In Eric Mack’s words (2009, 133), “An abiding feature of liberal individualism—and more particularly moral individualism—is the deep-seated rejection of the idea of a shared substantive social end or hierarchy of ends to which all members of society are to be devoted.”

While it is difficult to turn Rawls’s original position into a choice situation that would be congenial to libertarians, the reverse is not necessarily true. By insisting on determinate, situated persons who know who they are, libertarian
contractarians face a host of difficult decisions that the original position allows Rawls to sidestep: What material goods, talents and social opportunities (including opportunities to exit the bargaining table) do we endow each of the hypothetical bargainers with? If we just start from where people happen to be as a matter of social fact, without regard to the justice of how they got there, why should their resulting hypothetical agreement carry any moral weight? And if we are going to construct a fictional set of goods, talents and social opportunities, where do we get them from? How do we deal with holdouts (strategic or otherwise)? What real-life constraints on bargains should we build into the imagined bargain and what suppress?

Depending on how libertarians resolve each of these issues, it is a relatively simple matter, if not to turn libertarians into Rawls, at least to move them quite far in that direction. Consider the following two examples.

**Constructing the parties' threat point.** Whether a given agreement is to the mutual advantage of all parties depends on what the next best alternative is for each of them if they fail to reach agreement. The Rawlsian original position avoids the need to engage this question by eliminating any diversity of interests and hence any possibility of bargaining. By contrast, social contractarians on the right have to answer it in order to derive any determinate conclusions. Doing so requires that they flesh out, in a variety of ways, the bargaining position each person finds him or herself in. What is the value of each individual's endowments with and without social cooperation? Do we have to suppress social endowments that were not justly acquired? What alternative forms of social cooperation are available to the parties
if they walk away from the bargaining table? How costly would it be to arrange those alternatives, and should we take those costs into account in fixing their threat point?

Depending on the answers to these and a host of other factors that (in the real world) determine individual bargainers’ threat points, libertarians may find themselves legitimating political arrangements ranging from the fantasy Nozickean state in which everyone retains the full value of their marginal product, to the status quo (whatever it happens to be), to a highly redistributive state not that different from Rawls’s. (Fried, 2003)

**The problem of holdouts:** In any realistic, large-number bargaining situation, no solution is likely to garner unanimous assent. Rawls dealt with that problem by eliminating all heterogeneity among the parties, thereby guaranteeing unanimity. Social contractarians on the right, committed to making the parties to the bargain thick, determinate selves, cannot avail themselves of that solution. Instead, at crucial points they have dealt with the problem of holdouts by eliminating the requirement of unanimous consent.

Nozick’s derivation of the minimal state in Part I of *ASU* is a case in point. Nozick famously sets out to demonstrate how the minimal state could arise from a non-political Lockean state of nature without violating anyone’s Lockean rights. This requires that any rearrangement of those rights be consented to by the rights-holder. But the moment Nozick confronts equal and incompatible rights claims—in this case, the rights of each person to an independent protective agency of her choice—he gives up on consent, permitting the dominant protective agency
unilaterally to extinguish all other protective agencies. (Fried, 2011) Nozick’s motivation for doing so, it should be noted, is utilitarian: to achieve the social benefits of having one minimal state with coercive powers, rather than multiple warring proto-states. But if such aggregate welfarist considerations justify trampling on individual Lockean rights to create the minimal state, why not to solve other collective action problems within the state, and with very different distributive consequences than libertarians may wish for?iv

B. The content of liberty

Rawls (1999a, 157) grounded his substantive case for the two principles of justice in the Kantian imperative to treat all individuals as ends in themselves, and not merely means, in keeping with their moral status as free and equal beings. Once again, most libertarians would find little to quarrel with in that formulation. From that shared sentiment, however, the two camps diverge almost immediately, on the content of liberty in the first principle and the meaning of equality in the second.

Consider the following quintessentially Rawlsian defense of the lexical priority of liberty:

The moral adequacy of a society of project pursuers can be appraised by reference to the protection it affords to individualism: this is almost a universal truth. It is difficult to imagine a civil order in which individualism thrives but liberty rights are not a prominent component of that society’s moral grid.

Therefore it will not oversimplify much to cast the fundamental moral imperative of a well-ordered society as:

Each person is to have an equal right to the most extensive basic liberty compatible with a similar liberty for others.
The author of this passage is not Rawls but Loren Lomasky (1987, 100) who, as if to underscore the common intellectual provenance of libertarianism and Rawlsianism (and, one suspects, to tweak Rawlsians), helps himself in conclusion to Rawls’s first principle of justice.

In fleshing out the first principle, Rawls does achieve some common ground with libertarians (and also most clearly differentiates ‘justice as fairness’ from utilitarianism). In the end, however, that common ground is (in the view of libertarians) severely limited by Rawls’s construction of the ‘basic liberties.’

**The meaning of liberty.** The most expeditious way for Rawls to have packaged a social welfarist agenda in liberal individualist terms would have been to construe liberty as ‘positive liberty,’ in Isaiah Berlin’s famous taxonomy. He chose not to, adopting instead a ‘negative’ conception of liberty: freedom from interference with one’s pursuit of one’s own projects in life. (1999a, 198, 291; 1993, 327).

But having rejected substantive equality in favor of formal equality for purposes of the first principle, Rawls (1999b, 49; 1993, lvii) builds it into the second principle via the Difference Principle, which guarantees the “sufficient all-purpose means to enable all citizens to make intelligent and effective use of their freedoms.” Thus, argues Rawls, does “the two-part basic structure allow[] a reconciliation of liberty and equality.” (1999a, 179) It is a nice question whether Rawls’s division of labor between a negative conception of liberty and a substantive conception of equality gets him to a significantly different conclusion than if had simply given “liberty” a positive interpretation to begin with. But from the
libertarian perspective, it cedes most of the common ground Rawls might have been thought to establish with libertarianism by adopting a negative conception of liberty in the first place.

**What are the basic liberties?** Rawls’s (1999a, 53) list of the ‘basic liberties,’ amended slightly from *TJI* to account for criticisms from H.L.A. Hart and others, includes:

“[P]olitical liberty (the right to vote and be eligible for public office) and freedom of speech and assembly; liberty of conscience and freedom of thought; freedom of the person, which includes freedom from psychological oppression and physical assault and dismemberment (integrity of the person); the right to hold personal property and freedom from arbitrary arrest and seizure.”

As Lomasky notes (2005, 180), “conspicuously absent from this catalogue are economic liberties, including freedom of contract to buy and sell, to employ and be employed, or to accumulate and invest.” Throughout his writings, Rawls remained agnostic about the choice between private and collective ownership of the means of production, arguing the question must be settled at the (later) legislative stage, when the veil is partially lifted to allow knowledge of the “historical conditions and the traditions, institutions, and social forces of each country.” (1999a, xv-xvi; 1993, 298) But he expressly rejected the possibility that private property (other than personal property) was one of the basic liberties.

Whatever their other disagreements, all (right) libertarians are unified in their support of strong private property rights, although they get there by different routes. Lockean libertarians view private property rights (in one’s self, in the external things one acquires through one’s labor or other just means) as an
extension of self-ownership and at least coequal in importance. (Nozick, 1974; Hospers, 2003; Rothbard, 1982, 113-119). Kantian and consequentialist libertarians, in contrast, support them because and only insofar as they are instrumental in securing other desired ends (the ability to pursue one’s projects in life, mutual advantage defined in more utilitarian terms).

Rawls (1993, 298; 1999b, 49-50) addressed his very brief remarks on libertarianism only to the former. But the latter group has always been a significant strain of libertarian thought, and, among contemporary libertarians, is arguably much more influential than the Lockean tradition represented by Nozick. In an earlier generation, one would have included among the defenders of private property rights on instrumentalist grounds von Mises, Hayek, Milton Friedman and sometimes Ayn Rand. More recent proponents include James Buchanan, Israel Kirzner, Richard Epstein, David Gauthier, Gerald Gaus, Chandran Kukathas, Jan Narveson, Randy Barnett, Eric Mack and Loren Lomasky. The instrumentalist camp also stands in a more interesting relationship to Rawls, who purports to evaluate property rights—along with all other candidate primary goods-- by the same instrumental criterion: is it a good without which “persons cannot ... achieve their essential aims”? (Rawls, 1982, 172-73; 1993, 298)

That raises the possibility that all that separates Rawls and Kantian or consequentialist libertarians are their differing empirical assumptions about the instrumental value of property to liberty. If Rawls had been persuaded that a strong regime of private property rights was in fact essential to enable people to
pursue their own projects in life, or libertarians were persuaded that it was not, would their disagreements vanish? One assumes not, but why, exactly?

The likely answer on Rawls’s side is that the priority he puts on political and civil liberties over economic liberties in the end does not depend on empirical assumptions either as to the ends that persons in the original position would choose for themselves or the instrumental value of different ‘goods’ in the pursuit of those ends. Rather, it reflects Rawls’s own conception of the good. In the text of TJ itself (1999a, 139), Rawls equivocates about whether his list of “basic liberties” and the lexical priority he assigns to them rests on the presumed desires of others or a perfectionist notion of what they ought to desire. In the years following the issue of the first edition, however, he unambiguously came down on the side of the latter. (1999a, xiii; 1980, 525, 527, 547).

From the libertarian perspective, of course, this gives away the store. As Loren Lomasky (2005, 183-84) put it, Rawls’s willingness to give priority to his own “higher order” conception of the good cannot be squared with his ostensible commitment to respect equally the “different and indeed incommensurable and irreconcilable conceptions of the good” that individuals possess. (Rawls 1993, 303; Rawls 1982; Rawls 1971). The latter would seem to require that we value wealth accumulation, just as much as political participation, as a legitimate project in life if that is in fact what individuals want for themselves.

The opposite perfectionist notions are no doubt work at work to some extent, sub rosa, in libertarians’ ostensibly instrumental justifications for laissez faire capitalism and strong private property rights. Of course, as Loren Lomasky
(2005, 191) and other libertarians have argued, it is possible that a redistributive, social welfare state will frequently lead to worse outcomes especially for the least well-off. But it is possible it would not, and I think it is a fair surmise that libertarians’ empirical hunches, as much as Rawls’s, are following their political predilections and not the other way around.

The lexical priority of liberty. In his famous response to Rawls’s first principle, H.L.A. Hart argued that Rawls’s commitment to the “most extensive total system of liberty compatible with a similar liberty for others” ignored the reality that one person’s exercise of her basic liberties may impinge on another’s basic liberties. Rawls, responding to this objection in “The Basic Liberties and Their Priority” (1993, 291), revised the desideratum to a rhetorically more modest “fully adequate scheme” of liberties, and explicitly acknowledged that liberties must “be limited when they clash with one another,” language that made its way into TJ2 (1999a, 54)

One could argue that both of these qualifications were implicit in Rawls’s original formulation of the “most extensive total system of liberty compatible with a similar liberty for others.” The real problem for Rawls was not to acknowledge such clashes but to show they can be resolved by considerations that sound in individual rights rather than utilitarian concerns. This problem is hardly unique to Rawls; it inheres in all deontological schemes (most definitely including libertarianism). Rawls (1993, 336) explicitly deferred the solution to the constitutional stage, and most libertarians have de facto deferred it as well. That omission is relevant to the present inquiry, because, were they to address it, both Rawlsians and libertarians
likely would find themselves driven to utilitarian considerations to resolve the ubiquitous clash of “rights,” thereby bringing the two not only closer to utilitarianism but to each other as well.

The little that Rawls does say about reconciling clashing liberty interests supports that prediction. Rawls (1993, 295) minimizes the problem by arguing that basic liberties are not “infringed when they are merely regulated, as they must be, in order to be combined into one scheme as well as adapted to certain social conditions necessary for their enduring exercise.” But the distinction between “infringing” liberties and “regulating” them is a verbal one only. And turning to the problem of regulating speech, Rawls’s commonsense resolution of conflicting interests (1993, 335-36, 340-56) is not obviously distinct from what most people would intuit to be the optimal aggregative solution.

In the libertarian camp, Nozick’s confused attempt in Part I of *ASU* to adjudicate between identical warring rights to self-protection in the state of nature provides its own cautionary tale. In the end, Nozick extricates himself by recourse to the decidedly non-libertarian principle that might makes right, itself justified by the decidedly non-libertarian desire to maximize social good by forcing warring factions into one minimal state. (Fried, 2011).

**C. The meaning of equality**

Rawls’s Difference Principle has provoked more criticism than any other aspect of ‘justice as fairness.’ That criticism has come from all philosophical quarters. But the aspect of the difference principle that has preoccupied welfarist and egalitarian critics—whether the strict leximin constraint on inequality would
(or should) be chosen in the original position—has been at most a minor irritant to libertarians. Their primary objection is to the baseline of equality built into the original position. (Rawls, 1999a, 55, 130) That starting point rules out the possibility that the real-world self-interest of the better endowed will be reflected in the ex ante calculus of ‘mutual advantage’. In this respect, then, libertarians’ complaints about the difference principle reduce to their complaints about the original position.

There is, however, a deeper issue raised by the Difference Principle that is not subsumed under libertarians’ objections to the original position: whether it can be defended as a plausible ex ante choice even from Rawls’s original position. As many have noted, Rawls’s maximin solution is rational only if we assume extreme (indeed, infinite) risk aversion. Rawls (1999a, 149, 512-13) repeatedly denied he was resting the case for the Difference Principle on that assumption. But it is difficult to see how else to interpret comments like “the two principles are those a person would choose for the design of a society in which his enemy is to assign him his place,” or that the Difference Principle would be chosen by someone who “has a conception of the good such that he cares very little, if anything, for what he might gain above the minimum stipend that he can, in fact, be sure of by following the maximin rule.” (Rawls, 1999a, 132-34) And it is difficult to see what, other than infinite risk aversion, would lead someone in the original position to select the Difference Principle out of self-interest.

The maximin solution is, however, a rational choice from the perspective of those who find themselves actually to be the worst off ex post. Rawls’s official
position, of course, is that the choice of principles is to be made from an ex ante and not an ex post perspective. But he frequently elides the two, and at least on occasion unambiguously adopts an ex post perspective, including in the passage that provoked a strong rejoinder from Nozick in ASU. Positing two individuals, A (the more favored) and B (the least favored), Rawls (1971, 103) answers the hypothetical complaint of A that “he is required to have less than he might since his having more would result in some loss to B” by noting that A is still better off than he would have been under a system of noncooperation. Nozick (1974, 196) famously responded that the same defense could be given for any scheme of cooperation, and that the question, left unanswered by Rawls, was why A wasn’t justified in holding out for a scheme that gave him a larger share of the benefits than he would receive under the difference principle.

But as Nozick (1974, 196-97*) also noted in a lengthy footnote, there is a deeper, structural problem with Rawls’s argument: how to make sense of it from the “perspective of the original position.” In the original position, there would be no A and B; there would be only the representative person with some uncertain chance of being A and some uncertain chance of being B. If that representative person chose the difference principle, on balance, as the principle most advantageous to his future self, given the possibilities of turning out to be A or B, to whom is he now complaining? And when is now? Nozick concluded that the only sensible way to construe A’s and B’s complaints is to imagine that A and B have stepped out of the original position and are viewing matters ex post, after the winners and losers are known. From that perspective, B, knowing he is the loser and A the winner, will
understandably be unmoved by the statement that he should have less so that A might prosper more.

But if B's argument from the ex post perspective is decisive for Rawls, his contractualist thought experiment is not merely superfluous; it is misleading. For in that case, the notion of impartiality that motivates the Difference Principle is not that we should stand as equals in choosing the basic structure of society that will further each of our own interests, judged ex ante, but that we should stand as equals (at least as regards primary goods) at every moment of our lives. If it is a strain to fit the first notion of impartiality into the framework of mutually disinterested, rational choice, it is impossible to fit the second. The moral force of B's argument from the ex post position is that we must share what we have with the least among us, not because it is in our actual self-interest to do so and not because, from behind the veil, we would have calculated it to be in our self-interest as a form of insurance against the worst outcomes we might face, but because it is what we owe to others who turn out to be less fortunate than ourselves. Whether that duty is best realized through a maximin rule or instead by some other redistributive metric, it expresses a powerful moral perspective. But the perspective is not that of liberal individualism, conventionally construed. It is the perspective of those who (in Rawls's words) “agree to share one another's fate,” as would members of family who “commonly do not wish to gain unless they can do so in ways that further the interests of the rest.” (1971, 102, 105)

Libertarians, not surprisingly, have pounced on that statement and similar ones in *TJ* as evidence that Rawls is a liberal individualist in self-description only.
Judged by the commitments that actually motivate him, he reveals himself to be a communitarian, who conceives of society not (as libertarians would insist) as the mere aggregation of whatever individuals comprise it, with no identity or interests of its own, but instead as the highest organism, the “just social union of social unions,” without which we are all “mere fragments.” (1999a, 464)

Lomasky (2005, 199), assessing the resulting divide between libertarians and Rawls, concludes that the reason that Rawls does not end up a libertarian is “not because Rawls’s theoretical underpinnings are fundamentally hostile to libertarian perspectives…. [T]hey are in fact hardly more than a hair’s breadth away from yielding a recognizably libertarian position.” It is instead that Rawls “is more committed to his egalitarian redistributionist conclusions than he is to the premises that generate those results. Whenever he enters into wide reflective equilibrium, opposition to libertarianism is one of those relatively fixed points unlikely to be dislodged…. [H]is continued inability to come to terms successfully with libertarianism is due to an internal tension between his methodology and his convictions. One or the other has to give; invariably, it is former.” As John Stick (1987, 387) notes, the same might be said of the other side as well: “The existence of these argumentative resources [to turn Rawls into Nozick and back again] suggests that the disagreements of Rawls and Nozick are founded upon the political content of their views, and not upon differences in method.”

D. And yet....

At this point, it is easy to see where Rawls and libertarians come apart and hard to see how to get them to agree on much of anything except by persuading one
or the other to change their fundamental moral commitments. But a number of other considerations put that conclusion in doubt.

On Rawls’s side, Rawls makes two concessions to the non-ideal world that could push ‘justice as fairness’ towards conclusions much more congenial to libertarians.

The first is deference to the strains of commitment. Whatever principles we might otherwise regard as just, Rawls argues, must be tempered in light of the possibility that they will have “consequences [the parties] cannot accept” or can do so “only with great difficulty.” (1999a, 153) It is possible to understand this argument as addressed to the rational person choosing from the original position—that is, arguing it would be rational for representative persons, choosing principles \textit{ex ante}, to take into account peoples’ likely ex post reactions to really bad outcomes. If so, the argument collapses into the argument for the ex ante rationality of extreme risk aversion, with all the problems that attend that argument. But the more likely concern behind ‘the strains of commitment,’ I believe, is that however rational it might be \textit{ex ante} for people to take a given gamble, one cannot count on them to take their licks quietly \textit{ex post}, when things in fact turn out very badly. Whatever the moral arguments for and against bailing people out of the adverse consequences of their informed choices, Rawls’s ‘strains of commitment’ sensibly suggests that we may want to do so as a matter of prudence, to maintain the stability of the social contract over time. Judged from this perspective, the Difference Principle dominates both average utilitarianism and libertarianism, because it protects
would-be losers against the worst possible outcomes that the other two theories of justice would permit. (Rawls, 1971, 176-78; 1993, 17; 1999b, 49-50)

This argument for the Difference Principle surely proves too much, because the threat to stability that Rawls is most worried about here can be avoided by guaranteeing a basic minimum in the context of utilitarianism, libertarianism or a sufficientarian version of egalitarianism, without going all the way to the Difference Principle. But more importantly, concerns about the “strains of commitment” can as easily be turned on ‘justice as fairness’ as its rivals. Nozick led the way here, noting that while Rawls worries about the potential resentment those who turn out to be worst-off would feel about any distributive principle less generous to them than the Difference Principle, he ignores the potential complaints of those who turn out to be best off that they are being bled dry for the sake of the worst off. (Nozick, 1974, ; Rawls, 1999a, 470-71) Lomasky (2005, 185-86) suggests one might also worry that everyone other than the worst-off would resent the draconian effects of the strict Paretianism built into the Difference Principle. There are other things one might worry about as well-- e.g., how those who care more about economic liberties than political liberties would react to a conception of liberty that protects the latter and not the former, or how those whose conception of the good life is to amass great wealth would respond to the Difference Principle, which deprives them of the opportunity to do so.

The larger point here is that the Rawlsian concern with “strains of commitment” belongs to the real world, not the ideal one. As a result, it opens up the whole of ‘justice as fairness’ to reevaluation in light of how people who do not
necessarily share Rawls's idealized conception of justice would actually respond to living in a world that has institutionalized it. It’s hard to say what the upshot of that reevaluation would be, but one cannot rule out the possibility that it would push Rawlsians to revise Rawls’s two principles substantially in the direction of libertarianism.

Rawls’s willingness to permit any inequalities that benefit the worst off likewise leaves the implications of the Difference Principle hostage to real-world human psychology. *Pace* G.A. Cohen, the incentives necessary to induce the well-off to work 40 rather than 30 hours a week or shift from painting to doctoring for the benefit of the least well-off are unlikely to account for the major inequalities we observe in (say) contemporary American society. But the picture might change if we took a broader view of labor decisions, to include the sorts of entrepreneurial risks that have arguably contributed the most to raising everyone’s (including those who are least well-off) standard of living in the industrialized world over the last three centuries. If Adam Smith and his followers were in the end proved right as an empirical matter about the universal advantages of an unregulated market economy, the Difference Principle would commit Rawls to join in their endorsement of laissez-faire capitalism, and with it a distributive scheme that most libertarians could happily live with.  

On libertarians’ side, a number of factors could well nudge the programmatic implications of libertarian principles much closer to Rawlsian egalitarianism than libertarians themselves have assumed.

The possibility that has gotten the most attention over the last two centuries
is whether and how to restrict individuals from appropriating the commons for their private use. The widely varying answers libertarians have given to this question have led (in their own accounts) to widely disparate distributive outcomes. At one end of the spectrum, libertarians like Rand, Rothbard, Narveson and Mack would allow private appropriation by the first comer, without any obligation to leave a “fair share” for others. At the other extreme, “joint-ownership” libertarians (e.g., Grunebaum, 1987) argue either that no use may be made of natural resources without the unanimous consent of others or resources maybe used but not appropriated for exclusive use.

Most libertarians fall in between these two extremes, recognizing the right to unilateral appropriation but subject to some version of Lockean proviso that we leave “enough, and as good” for others. They have, however, disagreed about what is in the commons, what rights one acquires by virtue of a legitimate act of appropriation, and what its required to leave “enough, and as good” for others.

Right libertarians have typically read the ‘commons’ very narrowly, limiting it to the material natural resources, and read the rights of appropriation broadly, at the extreme giving first appropriators full ownership over natural resources on the basis of token labor. Interpretations of the Lockean proviso, however, have varied more widely even within the right libertarian camp. Nozick (1974, 176-77) famously concludes that later-comers are adequately compensated by the institution of private property, which is made possible by earlier generations’ appropriations. Sufficientarian libertarians like Simmons (1992) and Lomasky (1987), in contrast, construe “fair share” to require appropriators to leave others
sufficient resources for a decent life.

At the other end of the political spectrum, ‘left libertarians’ have interpreted the “fair share” obligation in a fashion that leads to far more egalitarian results. At the modest end would be Georgist proposals for the state to expropriate land rents and redistribute them equally to all citizens. At the more extreme end, contemporary left libertarians like Michael Otsuka, Peter Vallentyne and Hillel Steiner have read “the commons” broadly, to include not only tangible natural resources but also social capital (language, culture, knowledge, functioning markets, etc.) and (in Steiner’s case) even the gene pool from which our genetically determinate selves emerge, and have interpreted the requirement to leave “enough, and as good” to authorize the state to levy a confiscatory tax on virtually all incomes, and distribute the proceeds so as to offset differences in natural endowments. The result is a redistributive scheme at least as egalitarian as Rawls’s, and in some cases quite likely more so. (Vallentyne and Steiner, 2000; Fried, 2004).

A number of other less-discussed factors could well push libertarianism in the direction of Rawlsianism. The first, reflected in Nozick’s Principle of Rectification, is the moral imperative most Lockians feel to undo the fruits of past injustices. As even Nozick conceded, given the historical record, complying with this duty in the world we actually live in could well have radical implications (starting with returning America to the Native Americans) that swamp in importance all the other distributive implications of libertarianism.

Second, as Lomasky notes (2005, 195-96), Nozick is a decided outlier among prominent libertarians in denying any societal obligations to the worst-off. Most
libertarians would guarantee a basic minimum to those who cannot obtain it for themselves, and many regard the refusal to do so as indefensible. In Lomasky's words, “To the extent that successful civility requires the provision of aid, welfare rights merit a place alongside liberty rights.” (1987, 126) This concession is enormously significant, both in closing the gap between Rawlsianism and libertarianism and in neutralizing what is, for most people, the most powerful selling point of ‘justice as fairness’: its guarantee that no one will be left destitute in the name of principle, be the principle average utilitarianism or libertarianism.

Third, libertarians who reject the moral purity of anarchy for the pragmatic advantages of a coercive state have just begun to confront the compromises with principle that may be necessary to secure those advantages. Having achieved the minimal state by force in Part I of ASU, the Nozick of Part II regains religion on the sanctity of individual consent, famously asserting that compulsory taxation is a form of slavery. But most non-anarchic libertarians recognize that the collective action problems that necessitate a state in the first place also necessitate giving the state the power to levy compulsory taxes, at a minimum to fund public goods; the power to solve the problems of externalities that cannot realistically be solved by private contract; and the authority to operate through majoritarian or supermajoritarian rule, at the cost of minority interests. (Mack, 1986; Lomasky, 1987; Schmidt, 1991; Epstein, 1985; Buchanan, 1968, 1975; Buchanan and Tullock, 1962; Sugden, 1990; Gauthier, 1986). At the extreme, as Herbert Spencer and other 19th century proponents of laissez-faire foresaw and Progressive-era social democrats set out to prove, there is no amount of redistributive mischief that cannot be justified by
reference to such necessities. (Spencer (1981 [1843], Fried, 1998). Pragmatic libertarians have spent much of the last century trying to work out principled limits to each of these necessary concessions. Without prejudging the ultimate success of these efforts, I think it is fair to say, once again, that one cannot rule out the possibility that at the end of the day, libertarians will find themselves moving much closer to social welfarists on pragmatic grounds.

Finally, there is the question of what Rawls himself might make of his two principles of justice, were he to revisit them now at forty years’ distance. Rawls left a number of clues that together suggest he might well have revised them significantly, in a direction congenial to libertarians. Two potential changes in particular are worth noting.

The first concerns individual responsibility. Rawls was famously equivocal in *TJ* about whether individuals’ distributive share should be ‘ambition-sensitive’—that is, reflect their work effort and other voluntary choices made against the backdrop of an otherwise just world. The Difference Principle, however, seems to resolve the question in the negative, basing individuals’ distributive shares on the positions individuals find themselves in, without regard to how they got there. (The Difference Principle would of course allow greater return to greater work effort to the extent necessary to induce effort that redounds to the benefit of all. But the inequalities that result are a forward-looking, instrumental concession, not a backward-looking measure of desert.)

Rawls’s apparent position on this issue is anathema to libertarians, who regard the right to the peaceful enjoyment of the fruits of one’s labor as “the core
prescriptive postulate of libertarianism.” (Mack, 2011, 673) But libertarians are hardly alone in this regard. Beginning with Ronald Dworkin’s ambition-sensitive resource egalitarianism, most contemporary liberal egalitarians have distanced themselves from Rawls on this point. As G.A. Cohen (1989, 933) said of Dworkin’s departure, in an oft-quoted passage, it “has, in effect, performed for egalitarianism the considerable service of incorporating within it the most powerful idea in the arsenal of the anti-egalitarian right: the idea of choice and responsibility.”ix

In the Preface to *TJ2* (1999a, xiv-xv), Rawls suggests that if he were writing *A Theory of Justice* anew, he would move much closer to the luck egalitarian position, and stress equal citizenship rather than equality of material goods as the motivation for redistributing wealth. As Andrew Williams (2006) notes, depending on one’s empirical assumptions, the shift to ambition-sensitive egalitarianism on the left could well give libertarians most of what they want in terms of outcomes, and would certainly give them the portion of economic rights they most care about on principle.

Second, in writings after *TJ1*, Rawls backed away from the Difference Principle significantly. Beginning with “The Idea of an Overlapping Consensus” in 1987, he suggested that the core moral intuition behind the Difference Principle was a sufficientarian version of egalitarianism that guaranteed “all citizens sufficient material means to make effective use of [their] basic rights,” a goal that he acknowledged could be achieved through a wide range of schemes, of which the Difference Principle was just one example. (1993, 156-57; 1997, 774) By the time he wrote the Preface to *TJ2*, Rawls acknowledged it might not even be the most
plausible example, and gave what was at best half-hearted and somewhat misleading defense of its relative virtues. (1999a, xiv)

If one imagines that Rawls circa 2012 in fact would have ditched the Difference Principle for some version of ambition-sensitive sufficientarianism, it is not out of the question that his disagreements with many libertarians would be reduced to haggling over numbers—that is, over exactly what level of a basic minimum is sufficient to enable a decent life.

E. Conclusion.

From the selfish perspective of contemporary readers, it is regrettable that the other *Theory of Justice*—the one framed as a response to the libertarian critique of the actual *Theory of Justice* -- will remain unwritten. One can only speculate how engaging with that critique might have moved Rawls to revise ‘justice as fairness.’ But there seem to be two obvious directions in which he could have gone, both foreshadowed above. Whether either of them would have left him with what he hoped to achieve—“a viable systematic doctrine” to rival that of utilitarianism—is more doubtful.

The first would have been to ditch any perfectionist notion of the good, in favor of the “incommensurable and irreconcilable” ends that determinate individuals actually possess, and to revise the list of primary goods accordingly, to include those things that are essential to attain those ends, whatever (within reason) those ends might be. It seems likely that the resulting principles of justice would have included much broader protections for property rights—at least insofar as they were the result of individual effort—and have ditched the Difference
Principle entirely, in favor of some guaranteed decent minimum for all. Faced with that revised version of ‘justice as fairness,’ one imagines that most libertarians would concede, however grudgingly, that it is a plausible interpretation of liberal individualism.

The second would have been to abandon any pretence of seeking terms of cooperation that would be to the mutual advantage of determinate selves, and substitute two principles, both unabashedly reflecting Rawls’s perfectionist notion of the good: strong protections for civil and political liberties, coupled with broad-based redistribution of wealth in accordance with some social welfarist metric. It seems likely that the two principles would look much like the actual ‘justice as fairness,’ with one notable exception: the Difference Principle would have to be jettisoned in favor of a redistributive scheme that strikes a more plausible compromise between egalitarian and social welfarist ends. It is an open question whether that amended version of ‘justice as fairness’ would in the end look all that different from average utilitarianism, once one takes into account (in calculating utility) the diminishing marginal utility of income and what one assumes is a widespread preference for political and civil liberties. But whether it did or not, it would frankly acknowledge that, like utilitarianism, the redistributive portion of its program rests on other-regarding, not mutually disinterested, values. Whatever libertarians made of this amended version of ‘justice as fairness’ on the merits, one imagines they would regard that acknowledgement as a victory.

Notes
I use the word “deontological” to distinguish the libertarian revival spearheaded by \textit{ASU} from the more pragmatic Hayekian and public choice defenses of laissez faire, which were already part of the mainstream academic discourse by the time \textit{ASU} was published. 

I omit consideration of the virtue-based (or teleological) arguments for libertarianism most associated with Ayn Rand and (among contemporary writers) Douglas Rasmussen and Douglas DenUyl. 

Rawls (1971, 15, 103) \textit{does} in fact engage the question when he argues that the difference principle is necessary to elicit the willing cooperation of the worst-endowed. But as discussed below, it is not obvious how this argument fits in with the original position, which presupposes that no one knows whether they are the worst-off or not, and hence would be making the demands on themselves. 

I return to this issue at pp. . 

The one exception Rawls makes is for political liberties, which he argues are sufficiently important to ‘guarantee that the political liberties, and only these liberties, are secured by... their ‘fair value’.’ (1999a, 179). 

“\textit{The natural effort of every individual to better his own condition ... is so powerful a principle, that it is alone... not only capable of carrying society to wealth and prosperity, but of surmounting a hundred impertinent obstructions with which the folly of human laws too often incumbers its operation.}” (Smith, 1981 [1776], bk IV, ch. v; ii , 42-43) 


On the evolution of Rawls’s thinking post-\textit{TJ1}, see Sterba (2003), 62-70. 

In addition to Dworkin, a list of “ambition-sensitive” egalitarians would include Amartya Sen, Philippe van Parijs, Richard Arneson, Eric Rakowski, G.A. Cohen, John Roemer, Peter Vallentyne and Brian Barry. For further discussion, see Knight and Stemplowska, 2011. 

I say misleading, because Rawls omitted the one comparison in which the weaknesses of the Difference Principle in most people’s minds are brought to the fore: Rawls’s two principles of justice versus the first principle plus average utilitarianism with a sufficientarian minimum. 

References 


