

Counterfactual justifications of the state

Ralf M. Bader

Merton College, University of Oxford

ABSTRACT: By providing an interpretation of Nozick's justification of the state in *Anarchy, State, and Utopia*, this paper identifies and illustrates a form of justification that is distinct from traditional hypothetical, teleological and historical justifications.

I Nozick's justification of the state

In part I of *Anarchy, State, and Utopia* (ASU), Nozick attempts to refute the individualist anarchist and answer what he considers to be the fundamental question of political philosophy, namely whether there should be any state at all. In particular, he tries to justify the state by means of a hypothetical invisible-hand account, specifying a complicated process by means of which a minimal state could come into existence without violating any rights.

Nozick's account begins with individuals in the state of nature who enforce their own rights. This self-enforcement brings with it various inconveniences, leading to the formation of mutual-protection associations that evolve over time into professional protection agencies. Due to the network externalities characterising the market for the provision of protective services, a dominant protective agency establishes itself. This agency prohibits independents, i.e. people who are not clients of the agency, from enforcing their own rights, on the basis that they employ risky procedures that risk violating the rights of its clients. It thereby becomes an ultraminimal state that claims a monopoly on coercion within a certain territory. In order to be justified in prohibiting independents from engaging in risky self-enforcement, they must be compensated for the resulting disadvantage, which can be done by providing them with protective services. In this way, the rights of everyone within the territory will be protected and a minimal state will have arisen in a legitimate manner.¹

*For helpful comments, I would like to thank Bas van der Vossen, Joseph Carlsmith, Peter Jaworski, Fabian Wendt as well as audiences at Oxford and at the OSPP conference in Syracuse. I am grateful to Jerry Gaus and an anonymous referee for very detailed and helpful referee reports.

¹For a more detailed account cf. Bader: 2010, pp. 28-35.

1.1 A complete non sequitur

Nozick's justification² of the state is almost universally considered to be a failure. Not only do various critics find faults in assessing particular steps of his argument, for instance when it comes to Nozick's claims about procedural rights, network externalities as well as the compensation principle. In addition, it is generally held that Nozick's project is ill-conceived and that the very idea of trying to justify a state by appealing to a hypothetical invisible-hand explanation is misguided. As Simmons puts it, "given Nozick's orientation toward historical (or 'pedigree') evaluations of institutional arrangements, his justification of the state in terms of a purely hypothetical account of a minimal state's genesis might seem a complete non sequitur" (Simmons: 1999, p. 744). The thought is that, even if all the objections to the specific steps of his account could be overcome and the whole argument were to go through, Nozick would still not have managed to establish anything of significance.³

To begin with, it is hard to see how a hypothetical account could perform the requisite justificatory work. As Miller points out: "to say of a state that it *could* have arisen by [voluntary and permissible] means is actually to say very little" (Miller: 2002, p. 19). The fact that Nozick provides a hypothetical account is particularly puzzling, given that he develops and defends an entitlement theory of justice which is a distinctly historical theory (cf. Danley: 1978, p. 190). "The fact that a thief's victims voluntarily *could* have presented him with gifts does not entitle the thief to his ill-gotten gains" (ASU, pp. 151-152). If, as Nozick holds, the mere fact that someone could have voluntarily transferred property does not generate an entitlement, given that what matters is whether the person actually transferred the property, then how can legitimacy be conferred by the fact that an institution could have arisen without rights violations having taken place? Should it not equally be the case that what matters is whether the institutions actually did arise without violating rights? If distributions are justified in terms of the actual historical processes that generated them, then why is it not the case that political institutions are justified in a similar manner in terms of the actual historical processes that gave rise to them?

It might be suggested that all that is at issue is establishing the mere possibility of a legitimate state, i.e. showing that the state is not intrinsically immoral, in order to refute the individualist anarchist who claims that states are essentially illegitimate (cf. ASU, p. 6 and p. 119) and that hypothetical arguments do suffice for establishing such a possibility claim. Whilst this conclusion can in fact be achieved by means of a hypothetical account, this can be achieved in a much

²'Justification' and 'legitimacy' will be used interchangeably. For Nozick these amount to the state having the liberty right to use coercion to enforce rights, to punish wrongdoing, and to prohibit risky private enforcement (cf. ASU, pp. 133-134 and pp. 137-138).

³Cf. Schmidtz: 1990, Miller: 2002, amongst many others. For a more optimistic assessment cf. Gaus: 2011.

less roundabout manner by pointing out the (rather remote) possibility of a state arising as a result of unanimous consent. Accordingly, no appeal to complicated invisible-hand explanations would be required.⁴

Whilst Nozick's invisible-hand justification shows that a state can legitimately arise without unanimous consent, it is unclear as to why this should be of any particular import (though cf. ASU, pp. 293-294). Moreover, it is not even clear how strong Nozick's justification is, given that the justificatory force that derives from invisible-hand processes is relatively dubious since there is no guarantee that such processes give rise to improvements. (As Nozick acknowledges, an account that would explain the state's emergence in terms of a process of deterioration would not justify the state, cf. ASU, p. 5 footnote.)

Though the actions involved in an invisible-hand process are individually rational, they can fail to be collectively rational. The possibility of this type of collective action problem is particularly relevant given the path-dependency of the process leading to the dominant protective agency that Nozick envisages. "Clearly, it does not follow from the fact that people individually have reason to choose to sign up with the largest agency in their geographic region that collectively they have reason to applaud its emergence as a dominant agency" (Miller: 2002, p. 21). By contrast, this type of collective irrationality is unlikely to be found in the case of visible-hand explanations that require that everyone approve of the collective outcome. Unlike in the case of invisible-hand explanations where the individual transactions that are voluntarily and deliberately performed do not make reference to the collective outcome that emerges from them, in the case of a social contract the collective outcome is the very thing to which everyone agrees.

Thus, it is not clear why Nozick is providing an invisible-hand explanation. What purpose does this kind of explanation serve that could not be achieved by another kind of explanation, such as a consent-based account?

In short, why should one be in any way concerned about what could have arisen by an invisible-hand process? Both the hypothetical nature of the explanation as well as Nozick's insistence on using an invisible-hand process seem to be, at best, puzzling and, at worst, deeply misguided.

Moreover, it is not clear in what way a hypothetical account is meant to favour the minimal state in particular. After all, one can set up a legitimate more-than-minimal state by unanimous consent.⁵ Additionally, Nozick even provides an invisible-hand explanation of the more-than-minimal state in chapter 9 of *Anar-*

⁴Cf. "It is curious that Nozick gives no explicit attention to a Lockean contract as an alternative, more direct, route from the state of nature to a minimal state" (Miller: 2002, p. 16).

⁵Simmons argues that there is a conflict between, on the one hand, the consensualist strain of Nozick's approach, which would seem to make room for justifications of both minimal and non-minimal political arrangements, and, on the other, the minimalist strain that leads him to consider only the minimal state as being justified (cf. Simmons: 2005). The interpretation put forward in this paper shows how Nozick can assign a privileged position to the minimal state, whilst recognising that non-minimal states can be legitimate under special conditions.

chy, State, and Utopia, which he describes as being based on a principle of democracy and which amounts to “ownership of the people, by the people and for the people” (ASU, p. 290). Though any state that would have arisen in such a manner would be unappealing (at least to Nozick), it would not violate any rights and would not be illegitimate. The problem is thus that it is possible both for minimal as well as more-than-minimal states to arise in a legitimate manner via both invisible-hand as well as visible-hand mechanisms, making it unclear in what way the minimal state is meant to be privileged.

We are thus faced with three questions that put into doubt the very intelligibility of Nozick’s project:

1. How can a hypothetical account justify the state?
2. Why does Nozick invoke an invisible-hand explanation?
3. In what way is the argument meant to privilege the minimal state?

1.2 Rectifying past injustices

Prima facie, there is a stark incongruity between Nozick’s attempt to justify the state on the basis of a hypothetical account and his insistence that the justification of property distributions is a historical matter. Nozick’s entitlement theory of justice is concerned with how a distribution arose. In ideal circumstances,⁶ every action conforms to the principles of justice in acquisition and the principles of justice in transfer. The justice of a set of holdings will then derive from the actions that brought it about, insofar as what makes a distribution just will be the fact that it resulted from actions that conformed to these principles (rather than it being just in virtue of instantiating some specified pattern). In this way, the entitlement theory is a distinctly historical theory and it would seem that hypothetical facts are of no relevance to questions of distributive justice – all that matters is the actual facts about how the distribution came about. “Justice in holdings is historical; it depends upon what actually has happened” (ASU, p. 152).

Yet, the situation is drastically different in the context of non-ideal circumstances, where there is only partial compliance with the principles of justice. While acquisition principles specify how just holdings can arise in the first place, and while transfer principles specify how justly held holdings can be transmitted in a way that preserves justice, neither of these sets of principles tells us what is to be done when injustices arise and when holdings no longer conform to these principles. In such a situation, looking at the past is no longer sufficient. Once there is deviation from what is dictated by these principles of justice, the principles of rectification come into effect and specify what needs to be done in order for the injustice that has occurred to be rectified.

⁶The contrast between ideal v. non-ideal circumstances is to be understood in terms of full v. partial compliance with the relevant principles of justice.

Historical principles thus turn out to be only part of the full story. The entitlement theory switches in non-ideal circumstances from being a purely historical theory to a (partly) counterfactual theory. In addition to actual historical facts, one also has to appeal to counterfactual facts. “The principle of rectification presumably will make use of its best estimate of subjunctive information about what would have occurred (or a probability distribution over what might have occurred, using the expected value) if the injustice had not taken place” (ASU, pp. 152-153).⁷ Justice is then no longer a matter solely of what did happen but also of what would have happened, in particular a matter of what would have happened had the principles of justice in acquisition and transfer not been violated. The fact that distributions are justified not only in terms of what did happen, but also in terms of what would have happened, implies that hypothetical facts, in particular counterfactual facts, play a crucial role in the theory of justice in contexts in which injustices were committed.

Whereas a distribution is justified in ideal circumstances if it actually arose through a process in which no property rights were violated, it is justified in non-ideal circumstances if it would have arisen had circumstances been ideal. It is accordingly possible for a distribution that has arisen through a process involving unjust steps to nonetheless coincide with what would have resulted had no property rights been violated and to be justified in virtue of coinciding with this counterfactual distribution. That is, if the actual distribution coincides with the counterfactual distribution that would have been brought about in the absence of injustice, then, despite the fact that unjust steps led to the actual distribution, this distribution will be just on the basis that these injustices will have been rectified.⁸

1.3 Counterfactual histories

A historical theory of legitimacy that specifies which processes give rise to just political institutions will have to incorporate principles that are analogous to the principles of rectification. Such principles will specify what is to be done in non-ideal situations in which the historical principles have been violated and in which illegitimate institutions are in existence, telling us how to modify institutional structures to rectify past injustices. These principles will appeal to counterfactual facts about what institutional structures would have resulted if no rights had

⁷We will work with the simplifying assumption that there is a unique closest world in which the antecedent is satisfied (or that all closest worlds agree in the relevant respects), so that we always end up with would- rather than might-counterfactuals.

⁸The precise specification of the relevant counterfactual is a highly complicated matter. Identifying the correct rectification counterfactual is beyond the scope of this paper. The principle given in the main text is a rough first-pass approximation that is problematic in a number of respects but that will do for illustrative purposes (some of the relevant complications are discussed in section 2.6).

been violated on the part of these institutions.⁹ In the same way that counterfactual considerations become relevant in justifying distributions of holdings in non-ideal circumstances, they also become relevant in justifying political institutions in such circumstances.¹⁰ That is, in the same way that the principles of rectification require us to ask: ‘If no rights had been violated, what distribution of holdings would have resulted?’, we can ask: ‘If no rights had been violated, what institutional structures would have resulted?’¹¹

There is thus a general issue of how rectification is to take place in non-ideal circumstances that is equally applicable to property distributions as to political institutions. This issue is addressed by principles of rectification that are concerned with what needs to be done to right past wrongs. According to Nozick, these principles have an important counterfactual element, in that they require one to bring about that situation (conforming to the relevant principles of justification) that would have resulted had those wrongs not occurred, i.e. the situation that would have resulted under ideal conditions.¹² It is by determining what these counterfactual situations are like that potential explanations (i.e. explanations that do not fit the actual situation but would be correct explanations if things were different), in particular process-defective explanations that invoke processes that do not actually obtain but that would have obtained under suitable conditions, can justify their explananda and play a normative role (contra Wolff: 1991, p. 50 and p. 146 endnote 15).

Once this parallelism is recognised, Nozick’s justification of the state can be understood as a rectificationist justification of the minimal state. He is trying to provide a hypothetical account that explains what would have happened had there not been any rights violations. Part 1 of *Anarchy, State, and Utopia* constitutes a sketch of the relevant counterfactual history. The resulting political institutions are such that bringing them into existence would rectify the past injustices at the institutional level.¹³ In the same way that bringing about a set of

⁹The legitimising conditions for institutional structures require only that no rights be violated by the institutions in question – rights violations on the part of individuals do not form part of these conditions. (After all, it is the rights-violating behaviour of individuals that gives rise to the inconveniences of the state of nature, initiating the process that leads to the emergence of protective agencies.)

¹⁰Nelson has claimed that there is a disanalogy between parts 1 and 2 of *Anarchy, State, and Utopia* since the “arguments in pt. 2 of *ASU* depend on insisting that explanations (justifications) of distributions of resources *not* be process defective” (Nelson: 1986, p. 168 footnote 22). However, we can see that, far from being disanalogous, the structure of justification is the same in each case. In non-ideal circumstances it is process defective, i.e. counterfactual, whereas in ideal circumstances it is non-defective, i.e. historical.

¹¹As noted above, this is a simplistic first-pass approximation. A satisfactory theory of rectification will have to appeal to a much more complicated counterfactual.

¹²The transition involved in bringing actual institutions into conformity with those institutions to be found in the relevant counterfactual scenario is not allowed to violate side-constraints and might, moreover, have to satisfy certain procedural requirements.

¹³It is important to note that the sense in which a minimal state rectifies past injustices is re-

holdings that coincides with what would have resulted had there not been any rights violations ensures that a just distribution obtains, bringing about a set of political institutions that coincides with what would have resulted had there not been any rights violations ensures that just political institutions are in place. That is, a set of institutions can be justified if it would have arisen had no rights been violated, thereby making it possible for a state that results from an unjust process to be legitimate on the basis that it coincides with what would have resulted had no rights violations been perpetrated on the part of the state.

It is precisely such a coincidence of institutional structures under ideal and non-ideal circumstances that constitutes the essence of Nozick's process-defective justification of the minimal state (cf. ASU, pp. 7-8). The process described in part I of *Anarchy, State, and Utopia* is a process that explains and justifies the existence of a minimal state but that is defective in that this process is not actual. Yet, it is a process that could have brought about such a state and, more importantly, would have brought it about had the morally impermissible processes that actually occurred (or would have occurred in nearby possible worlds) not taken place. In this way, we can understand Nozick's claim that "[a] theory of a state of nature that begins with fundamental general descriptions of morally permissible and impermissible actions, and of deeply based reasons why some persons in any society would violate these moral constraints, and goes on to describe how a state would arise from that state of nature will serve our explanatory purposes, *even if no actual state ever arose that way*" (ASU, p. 7). For a state to be justified, it is not necessary that it did arise by the process that Nozick describes, since it is sufficient that it corresponds to what would have resulted under such ideal circumstances.

Accordingly, we can see why Nozick's justification of the state has to take the form of a hypothetical explanation. This is because the injustices that characterise the actual histories of political institutions make it necessary to consider counterfactual histories in which no injustices are committed, if one is to determine what needs to be done to arrive at a just set of political institutions that rectifies the past injustices. In other words, actual injustices perpetrated by states make it impossible to justify such states on the basis of historical justifications, thereby requiring a shift to counterfactual justifications.

Only counterfactual justifications can determine how the past injustices are to be rectified and hence determine what type of political institutions are justified and will have to be put in place. Given that actual circumstances are non-ideal, institutional structures need to be justified by showing that they correspond to what would have happened had circumstances been ideal. This means that, although

stricted to the institutional context, i.e. it is concerned with the claims individuals have to (not) be governed by certain (types of) institutional structures. That is, the establishment of a minimal state ensures that injustices will have been rectified as far as institutional structures are concerned. The rectification of injustices as they affect the well-being of individuals and their holdings is another matter that cannot be addressed by simply putting into place certain institutional structures. (Relatedly, the notion of 'ideal circumstances' is restricted to the institutional setting.)

no actually existing state can be justified in terms of its actual history, minimal states can be justified in terms of their counterfactual history, since, though such states did not actually arise in the right way, they would have arisen in the right way had circumstances been ideal.

1.4 Invisible-hand explanations

Moreover, we can see why Nozick's justification has to take the form of an invisible-hand account. This is because the evaluation of the counterfactual requires one to assess not just any situation in which no rights are violated, but the closest possible situation in which there are no rights violations. The invisible-hand character of the explanation precisely ensures that this closeness condition is satisfied. What makes the particular hypothetical process that Nozick envisages relevant to what institutional structures are justified in the actual world, despite the fact that the process is defective in that it does not occur in the actual world, is its closeness to the actual historical process. In particular, it is the closest process to the actual process that does not involve violations of rights and as such determines which institutional structures are justified, given that the actual as well as alternative (closer) processes are ruled out on the basis that they involve rights violations.

The closeness of the counterfactual process that legitimises the minimal state is due to the fact that it is an invisible-hand process. Such processes appeal to filtering and equilibrium mechanisms to explain macro-level outcomes in terms of micro-level events that do not make reference to the macro-level outcomes (cf. Nozick: 1994, p. 314). The macro-level outcome will accordingly be an emergent feature of the system. The way in which the emergence of macro-level outcomes is explained by invisible-hand processes matters because the evaluation of the counterfactual that determines how past injustices are to be rectified consists in assessing the consequences of making various local adjustments that are required by the no-rights-violation condition and that amount to eradicating the particular injustices that are part of the actual process. The change from an illegitimate state in the actual world to a legitimate state in the selected counterfactual world thus consists in a series of local changes that correspond to implementing the counterfactual supposition that the relevant rights-violating actions that undermined the legitimacy of the actual state (as well as those relating to illegitimate states in closer counterfactual scenarios) had not taken place. We then evaluate the consequences of these local adjustments, thereby determining what would have happened under these idealised conditions. The macro-level outcome that results from these adjustments represents the kind of state that is justified and that needs to be brought into existence in order to rectify past injustices at the institutional level.

If Nozick is correct that a minimal state will emerge from a broad range of starting-points by a process that involves no rights being violated, then making these small adjustments is very likely to take us to one of these counterfactual pro-

cesses. The macro-level outcome consisting of the existence of a minimal state will result from the vast majority of close-by systems satisfying the no-rights-violation condition. This is because the process that Nozick envisages is highly stable, in that it does not rely on a particular starting-point, making the macro-level outcome resilient under micro-level variations. Since a broad range of different micro-level starting-points will all converge on the same macro-level outcome, it is very likely that one will end up with the same result, namely a minimal state, independently of which of the systems one should happen to be in. Accordingly, the closest world in which no rights are violated will be a particular instance of the type of process that Nozick describes, which means that the local alterations amounting to a change from non-ideal to ideal circumstances will have as their result the emergence of a minimal state.

By contrast, excepting special circumstances where a visible-hand process is already underway and only fails to be completed because of a rights violation (e.g. a situation where a state was about to be set up by unanimous consent but where this attempt was illegitimately thwarted), such a series of local adjustments will not lead to a visible-hand process, since this type of process requires a much more radical change. This is because visible-hand processes only arrive at the desired macro-level outcome from a small set of highly specific starting-points. This reliance on specific starting-points implies that we are not at all likely to end up at one of them by just making local adjustments that eradicate particular impermissible actions that undermine the state's legitimacy in the actual world. Instead, what needs to be done to arrive at such a starting-point is to implement a global change by positing a social contract. It is this lack of stability, as well as the global and specific nature of the required starting-points, that makes visible-hand processes more remote than the type of invisible-hand process envisaged by Nozick.

Accordingly, the closest process to the actual process that leads to a legitimate outcome will be an invisible-hand process rather than a visible-hand process. The reason why Nozick does not appeal to a social contract is thus not that such a contract would fail to justify the state. A state brought about by unanimous consent would be perfectly justified. The problem is rather that such a scenario is too remote and consequently not relevant for assessing the counterfactual as to what institutions would have arisen had no rights violations taken place. Even though visible-hand processes that bring about political institutions through unanimous consent legitimate the resulting institutions, they are substantially more remote from the actual world than invisible-hand processes. This ensures that it is the latter rather than the former processes that determine how the counterfactual is to be evaluated and thereby determine what is to be done to rectify past wrongs.

Moreover, not only do visible-hand processes require highly specific starting-points that are far removed from the actual world if they are to yield justified institutions, the particular outcome of any such process will depend on the specific

nature of the starting-point. Unlike in the case of invisible-hand processes where the same macro-level outcome results from a large range of different starting-points, there is no convergence in the case of visible-hand processes. For instance, in the case of consent theory, any institutional structure that people can consent to can be the justified outcome of a visible-hand process, making the outcome dependent on the specific starting-point. This fine-grained dependence on the specific nature of the starting-point risks leading to underdetermination, which would make it much more difficult to evaluate what institutions are justified.

1.5 Privileging the minimal state

The proposed interpretation of Nozick's justification as consisting in determining how past injustices at the institutional level are to be rectified thus explains why the justification has to take the form of a hypothetical explanation that is based on invisible-hand processes. In addition, it can explain in what way the hypothetical invisible-hand account favours the minimal state. If Nozick is right that a minimal state would have resulted if no rights violations (on the part of the state) had taken place, then it is such a state that is to be brought about, no matter whether a more-than-minimal state could have resulted in a legitimate manner in some other context. Any explanation of legitimate more-than-minimal states will be sufficiently remote to make it irrelevant.¹⁴

In particular, such explanations will either rely on visible-hand processes that do not explain the emergence of the state in terms of local non-political phenomena and thus will not be fundamental explanations but will require bringing in political notions to explain the emergence of the political realm (cf. ASU, p. 6). Or these explanations will rely on the kind of complicated process characterised in the demoktesis chapter (cf. ASU, pp. 280-290) which relies on highly elaborate and specific mechanisms and is too remote to be relevant. Although the process leading to demoktesis is an invisible-hand process, in the sense that the resulting state is an emergent outcome that arises without having been intended by anyone, the process is not stable due to its high specificity and does not proceed on the basis of mundane behaviour on the part of individual agents responding to ordinary incentives.

All of this ensures that, unless highly specific conditions are satisfied, one will not end up in a counterfactual scenario involving a more-than-minimal state when evaluating what would have happened had there not been any rights violations. While there are plenty of close possible worlds in which more-than-minimal states arise as a result of processes that involve rights being violated (the actual world being one of them), all those worlds in which such states come into

¹⁴Although more-than-minimal dominant protection agencies (that provide various services for their clients in addition to the protection of individual rights) can emerge, the provision of such additional services cannot be extended (without violating rights) to independents who did not voluntarily join the agency and subscribe to these services.

existence without rights being violated are remote and all the close worlds satisfying this legitimising condition are ones in which there are only minimal states. Given Nozick's moral framework, the main threat to his account does not derive from the possibility of ending up with a legitimate more-than-minimal state but instead from ending up with no state at all, and it is only by appealing to contentious empirical claims about network externalities that Nozick can ensure that his preferred evaluation of the counterfactual comes out true.

Thus, the claim that for Nozick “[a] minimal state, and only a minimal state, could arise by an invisible hand process” (Schmidtz: 1990, p. 89) is a mischaracterisation of his position. When concerned merely with what could happen, the minimal state is not privileged since more-than-minimal states can legitimately arise via invisible-hand processes (as well as by visible-hand processes). After all, this is clearly illustrated by the demoktesis chapter. Instead, as we have seen, the key issue is not what could possibly happen, what kind of state can arise, but rather what would have happened in the absence of rights violations. It is in this context that the minimal state is privileged since, on Nozick's view, the most plausible interpretation of the counterfactual specifying what kind of state would have arisen had no rights been violated will identify a world in which a minimal state arises by means of an invisible-hand mechanism.

1.6 Nozick's project

The nature of Nozick's project can, accordingly, be explained on the basis that the non-ideal circumstances in which we find ourselves require us to assess what happens in the closest possible world in which circumstances are ideal, making it necessary that the account takes the form of a hypothetical explanation based on an invisible-hand mechanism. This means that the suggestion that “Nozick could have presented his solution – less stylishly – without mentioning invisible hands, potential explanations, or hypothetical processes” (Wolff: 1991, p. 52) turns out to be mistaken. Rather than being inessential features of the solution, these elements are indispensable to a rectificationist justification. Moreover, we can see that there is no incongruity between parts 1 and 2 of *Anarchy, State, and Utopia*, but that they cohere nicely and constitute a unified and systematic project.

This account has the further advantage that it also explains why the transition from an ultraminimal state to a minimal state is unproblematic, despite the fact that this transition differs in nature from the other steps of the process, given that it requires that a compensation scheme be put in place, the setting up of which requires that the agents of the ultraminimal state are morally motivated.¹⁵ What

¹⁵Moreover, the aims and intentions of these agents might seem too close to the explanandum, namely the existence of a state, for this step to be straightforwardly classified as resulting from an invisible-hand mechanism (cf. Gaus: 2011, p. 122). Nozick addresses this concern by noting that the explanation “does not specify people's objective as that of establishing a state. Instead, persons view themselves as providing particular other persons with compensation for particular

is doing the work in this step of the argument is not an appeal to economic incentives and empirical claims about how people behave, but instead a claim about what is morally required. As Nozick notes, “the transition from an ultraminimal state to a minimal state morally must occur. . . . The operators of the ultraminimal state are morally obligated to produce the minimal state” (ASU, p. 52). This turns out to be unproblematic on the grounds that a failure to compensate would result in the violation of individual rights and would, accordingly, contradict the counterfactual supposition that no rights violations (on the part of the institutions) take place.

Nozick’s justification in this way relies on the interaction between two features, namely (i) moral constraints and (ii) empirical mechanisms, that jointly determine how the counterfactual is to be evaluated, i.e. what the worlds that are closest to the actual world that do not involve any rights violations are like. These two features are responsible for the two stages in the emergence of the minimal state. Empirical mechanisms underlying the invisible-hand process lead to the prohibition of independents and thereby to an ultraminimal state. The moral constraints in the form of the no-rights-violation condition, in turn, lead to the compensation of independents and thereby to a minimal state.

In addition, it becomes clear that the condition of there not being any rights violations should be understood as there not being any uncompensated rights violations. Accordingly, it turns out that the fact that prohibiting independents requires compensating them for the resulting disadvantage is unproblematic. In this way, we can see that the objection that Nozick’s favoured invisible-hand process itself involves rights violations and as such cannot generate a legitimate state does not succeed.

This objection takes the form of a dilemma insofar as it is held that either (i) the prohibition of independents is legitimate in which case no compensation is required, or that (ii) this kind of prohibition is illegitimate in which case compensation is due (cf. Holmes: 1981, p. 61; also cf. Wolff: 1991, p. 72).¹⁶ In the former case, we do not end up with a minimal state but are instead stuck with the ultraminimal state. In the latter case, we do get to the minimal state yet only via a process that involves rights violations, which ensures that the resulting minimal state is illegitimate.

In response, we can note that the second horn of the dilemma turns out to be unproblematic once it is recognised that a state is legitimate if there are no unrectified injustices. For a state to be justified, it does not need to have an unblemished history but only needs to be such that all its blemishes have been removed, insofar as injustices have been rectified and compensation has been provided to those

prohibitions they have imposed upon them. The explanation remains an invisible-hand one” (ASU, p. 119). In other words, the transition is the result of a series of local adjustments, each of which can be understood in non-political terms.

¹⁶According to Nozick, this is a false dilemma, cf. ASU, pp. 83-84. This type of objection has also recently been criticised by Hyams: 2004.

who have been prohibited from engaging in risky activities. Accordingly, Nozick can either hold that this dilemma is a false dilemma, insofar as compensation can justify boundary-crossings, making it the case that there is no injustice even though compensation is required since the provision of compensation is precisely what makes it the case that the boundary-crossing is justified, or he can accept that the boundary-crossing is not justified and accept that there is an injustice but then hold that this injustice is addressed if the independents are compensated for the resulting disadvantage, since compensation ensures that the resulting situation is legitimate despite having a blemished history.

The three questions we started out with can thus all be answered by the counterfactual interpretation that understands Nozick as being concerned not with a mere possibility claim about what could have happened but a counterfactual claim about what would have happened, in particular with what type of institutional structures would have arisen if no rights violations had been perpetrated on the part of these institutions:

1. How can a hypothetical account justify anything (given Nozick's emphasis on historical processes in establishing legitimacy in general)?
 - A hypothetical account can justify insofar as it is a counterfactual account concerned with what would have happened had there been no rights violations, thereby specifying what kind of situation would amount to a rectification of the past injustices that have actually occurred.
2. Why does Nozick provide an invisible-hand explanation (rather than, for instance, a consent-based explanation)?
 - Invisible-hand explanations ensure closeness to the actual world, thereby making them process-defective explanations that determine how the relevant counterfactuals are to be evaluated, whereas consent-based explanations only characterise remote possibilities.
3. In what way is the minimal state privileged (when other states could also have come about in a legitimate manner)?
 - The minimal state is privileged in terms of its closeness to the actual world, since any scenario that satisfies the condition that no rights violations occur and that involves a redistributionist state (whether one resulting from a social contract or a demoktesis-style process) will be much more remote than one involving a minimal state.

It is worth noting that this does not mean that Nozick's justification is a success. After all, there are numerous problems relating to particular steps in the invisible-hand account. These problems are both normative in nature, in particular when

it comes to his appeal to procedural rights and the application of the compensation principle in relation to the prohibition of risky activities, as well as empirical in nature, in particular his invocation of network externalities in explaining why a dominant protective agency would arise.¹⁷ In addition, the notion of rectification is riddled with problems and difficulties (cf. ASU, p. 152 and Davis: 1981). Nozick's brief remarks concerning rectification are completely underdeveloped and there are many difficult questions that need to be addressed if a rectificationist justification of the state is to stand on a firm footing. Yet, what has been shown is that the project that Nozick is engaged in is intelligible and that it in fact constitutes an interesting alternative to the well-known traditional methods of justifying political institutions, such as hypothetical consent accounts. Nozick is neither offering a purely historical nor a teleological justification but instead a counterfactual or rectificationist justification.

2 The general framework

We have identified an account according to which a state can be justified in non-ideal circumstances on the basis of counterfactual facts. It is possible for such facts to play a justificatory role because it is not the state's actual history that matters in non-ideal circumstances. What matters instead is the counterfactual history that would have unfolded and the kind of state that would have emerged had circumstances been ideal. Rectificationist principles allow us to deal with non-ideal circumstances and in this way supplement historical principles of justification that only apply in ideal circumstances.

2.1 Supplementing historical accounts

A complete (non-teleological) theory of justification needs to do two things:

1. It needs to provide an account as to how institutions are justified in ideal circumstances.

This function is performed by historical principles of justification.¹⁸

¹⁷Some problems, such as the regressive character of procedural rights (cf. Paul: 1981, p. 73), threaten to undermine Nozick's account altogether, whereas others merely restrict the range of cases in which the account is applicable and in which (minimal) states can accordingly be justified, for instance the problem that it is not possible to compensate a thoroughly committed anarchist for being prohibited from enforcing his own rights.

¹⁸These principles are the analogues of the principles of acquisition and transfer. For Nozick they are concerned with 1. consent, and 2. the prerogative to enforce compliance with certain rules and to prevent people from engaging in certain types of risky behaviour (given that they are compensated for the resulting disadvantage). That is, one can be forced to comply with rules and give up self-enforcement either if 1. one has consented to being forced (i.e. one is a member of a protective association), or if 2. one's non-compliance would constitute a certain type of risky

2. It needs to provide an account as to how institutions are justified in non-ideal circumstances.

This function is performed by rectificationist principles of justification.

A complete theory consists of both historical and counterfactual principles, covering ideal and non-ideal circumstances respectively. This implies that the characterisation of legitimacy will be disjunctive: institutional structures are justified iff they either (i) did arise in a certain way, or (ii) would have arisen had conditions been suitably ideal.

These two types of principles are not on an equal footing. In particular, counterfactual principles are dependent on historical principles and are unable to generate content by themselves. This is because what should happen in non-ideal circumstances is understood in terms of what would have happened in ideal circumstances. The counterfactual justification of a set of institutions is based on these institutions coinciding with those that would result from a historical process under ideal circumstances, which implies that the rectificationist project of explaining legitimacy in non-ideal circumstances in terms of legitimacy in ideal circumstances presupposes an account of the latter. Counterfactual principles, accordingly, are not independent but instead implicate historical principles.

This means that counterfactual principles do not constitute a complete theory by themselves. Rather than replacing historical principles, they supplement them. They do this by showing how we are to make sense of justification in non-ideal circumstances in which states have blemished histories in terms of institutions coinciding with those to be found in counterfactual scenarios in which the historical principles are satisfied. They are thus not a substitute but a supplement that needs to be included in a complete theory.

2.2 Incomplete theories

Historical theories of justification will be incomplete without principles of rectification. A purely historical account that only specifies the mechanisms by means of which a legitimate state can come into existence, such as consent theory, tells us which institutional structures are justified (namely those that have received the appropriate form of consent) and which ones fail to be justified (namely those not resting on consent). Yet, by itself, it is an incomplete theory (in the same way that an entitlement theory consisting only of principles of acquisition and transfer is an incomplete theory) that does not provide adequate guidance in non-ideal circumstances where the historical principles have not been complied with. It only determines that there is injustice but does not tell us how it is to be addressed, in particular how the past wrongs are to be rectified. It thereby ignores that there

behaviour that can be prohibited, given that one is compensated for resulting disadvantages (i.e. one is an independent employing risky procedures).

are injustices that call for rectification, i.e. that there are particular steps that are required to address the particular injustices that obtain.

Non-ideal circumstances do not give rise to a blanket prescription to replace extant unjust institutions by institutions that conform to the historical principles. Simply instituting a political system that is considered to be justified according to the historical principles is not the appropriate response, since this does not address the past injustices that are in need of rectification. Insisting in the context of non-ideal circumstances that justification can only be based on actual consent amounts to ignoring the need for rectification. It is like insisting that all transfers of holdings have to be voluntary even when some holdings are illegitimate. What is required when a thief has illegitimately acquired certain goods is that the stolen goods be returned to their rightful owners. This transfer obviously does not have to be voluntary and the thief can instead be forced to return the goods. Voluntary transfers preserve justice and ensure that no new injustices are introduced. Yet, voluntary transfers are no longer what is called for when faced with unjust holdings. Similarly, while consent justifies political institutions in ideal circumstances, actual consent is not what is needed when institutions are unjust. Instead, what is required in each case is that the extant injustices be rectified.

Rectifying past injustices might even be required for the historical mechanisms of justification to be applicable. This is because a tainted past can undermine the normative force of consent. In particular, consent is only binding and only gives rise to justified institutions if the choice-situation is suitably untainted. This means that it can be necessary in non-ideal circumstances to address past injustices first before a consent-based justification can become applicable.

This problem can be illustrated by a situation in which an unjust state is in existence, but in which the citizens are given the opportunity to form a new state by means of a social contract based on unanimous consent. For instance, a dictator could allow the people to create a new state, where failure to reach agreement will mean reverting to the status quo, i.e. to the dictatorship. Since the status quo point of the bargaining problem is the unjust state, the incentives for reaching agreement are substantially different from those that would be faced if agreement were sought in the context of a justified status quo point. The problematic nature of this bargaining situation impugns the resulting social contract. As a result, it is far from clear whether unanimous consent in such a situation would (fully) justify the resulting set of institutions.

Not only is it unlikely that there will ever be unanimous consent, it is also not clear that such consent would classify as genuine consent, given that past injustices taint the choice situation. Consent theory thus only constitutes a self-standing theory in ideal circumstances. A complete theory, however, must also tell us how one is to deal with non-ideal circumstances, how past injustices are to be rectified and how one is to get to a situation in which the historical principles are applicable.

Dealing with non-ideal circumstances is precisely what is achieved by rectificationist principles. They determine what is to be done in non-ideal circumstances, i.e. how one is to deal with the injustices that are to be found at the level of institutions. They do this by showing which institutional structures are such that bringing them into existence would amount to a rectification of past wrongs. In this way, they show how institutions can be legitimate, despite failing to conform to the historical principles, insofar as they stand in the relevant correspondence to the institutions that would have emerged in ideal circumstances. Accordingly, one needs to supplement historical principles by principles of rectification to arrive at a complete theory that is applicable in both ideal and non-ideal circumstances.

2.3 Degrees of justification

When it comes to providing guidance as to how to get from institutions that lack justification to ones that are justified, it is important to have a comparative notion of justification that admits of degrees.¹⁹ Recognising degrees of justification enables a theory to provide guidance as to how a non-ideal situation is to be improved bit by bit, allowing one to approximate a state of affairs in which institutions are fully justified. This is particularly relevant when it comes to making improvements where the ideal cannot be reached (either in principle or in practice) but can only be approximated.

The rectificationist account naturally generates a comparative notion of justification, insofar as rectifying particular injustices amounts to partial rectification which corresponds to the institutions in question being partially justified, whereas rectifying all injustices amounts to the institutions being fully justified. In non-ideal circumstances, justification is understood in terms of a coincidence of actual institutions with those in the relevant counterfactual scenario. Whilst coincidence with a scenario in which no rights violations occur and in which the historical principles are perfectly complied with amounts to full justification, coincidence with a scenario in which only certain rights violations fail to occur and in which the historical principles are only imperfectly complied with can be construed as partial justification. In this way, comparative facts regarding the extent to which the counterfactual scenario is idealised generate an ordering of institutional structures. Institutional structures can accordingly be more or less justified, depending on the extent of unrectified rights violations that they involve.

In determining the extent of rights violations, one can distinguish different ways in which rights can be violated and attach different weights to different violations (for instance, one might like to discount or set aside what Thomson calls mere infringements of rights (cf. Thomson: 1990, p. 122)), such that the

¹⁹Certain understandings of justification will require one to think in terms of degrees of illegitimacy since only illegitimacy will admit of a scalar construal.

extent becomes a weighted function of certain types of violations. Weightings allow us to construct not just a quantitative measure of the deviation from the ideal case that is based merely on the number of rights violations, but also a qualitative measure that considers the significance of the deviations.²⁰

The degree of rectification is then determined by the proportion of (weighted) injustices that are rectified. A certain set of injustices are rectified if the actual institutions coincide with those in the closest counterfactual scenario that is based on the local adjustments amounting to eradicating these injustices (without introducing any new ones). We can thus order counterfactual scenarios in terms of the range of local adjustments that generate them. With which counterfactual scenario the actual institutions coincide determines comparative facts about the extent to which injustices are rectified. Past injustices are fully rectified if the actual institutions correspond to those in the counterfactual scenario in which no rights violations occur. By contrast, if no counterfactual scenario based on supposing that certain rights violations had not occurred contains the institutions to be found in the actual world, then no injustices are rectified, i.e. the only institutionally-coinciding counterfactual scenario is that in which none of the actual injustices failed to occur, i.e. the actual world.

How many (weighted) rights violations are to be found in the actual history determines the degree of injustice.²¹ Together, the degree of rectification and the degree of injustice determine the degree of justification. This corresponds to the degree of injustice of the relevant institutionally-coinciding counterfactual scenario. We can thus order institutional structures to be found in various counterfactual scenarios based on the extent to which these scenarios deviate from ideal circumstances. The closer they are to the ideal case, the smaller the (weighted) number of unrectified rights violations and hence the more justified the institutions are. Institutional structures are maximally justified if they correspond to those found in the counterfactual scenario that involves the minimal violation of rights. That is, they are fully justified if the institutional structures to be found in the actual world completely coincide with those of the ideal counterfactual scenario, which can happen either if there are no past injustices and hence no need for rectification, i.e. in ideal circumstances this counterfactual scenario is the actual scenario, or if there are injustices but they are fully rectified.

²⁰Since the requirement not to violate rights is distinct from the duty to rectify past rights-violations, one does not end up with a utilitarianism of rights, i.e. what is at issue is not bringing about fewer violations of side-constraints but rectifying past injustices to a greater extent. Moreover, weightings can be applied even if one considers rights to be side-constraints, given that the importance or urgency of rectification need not derive from the status of the right that has been violated.

²¹Comparisons across different states require the scales to be normalised.

2.4 Moral principles and empirical mechanisms

Which institutions are justified on the basis of counterfactual facts in a particular non-ideal context depends on two factors:

1. The moral framework specifying the historical principles of justification as well as the principles of rectification.
2. The descriptive account of the empirical mechanisms that explain how the behaviour and interactions of individuals lead to the formation of institutional structures.

A rectificationist justification is concerned with evaluating what structures would be generated by these mechanisms in the context of the counterfactual scenario in which the past injustices that occurred in the actual world and that need to be rectified do not take place but where instead institutions arise in conformity with the historical principles of justification. Since the resulting structures would arise in compliance with the constraints and requirements specified by the normative framework, they would be justified in those counterfactual situations and accordingly constitute the ideal case to which the non-ideal actual institutions need to be made to conform. Moral principles and empirical mechanisms jointly determine what the relevant ideal counterfactual scenario looks like and how the non-ideal actual scenario is to be made to coincide with it.

Rights, compensation principles and rectification principles constitute the moral framework.²² This determines, on the one hand, which actions, if any, leading to the situation in question generate (uncompensated) injustices that need to be rectified. On the other, it determines the constraints that institutional structures need to satisfy in order for them to be justified in virtue of their historical genesis, and thereby fixes the conditions that counterfactual processes must satisfy in order to generate legitimate outcomes that can be candidates for determining what needs to be done in order for past injustices to be rectified.

In addition to depending on one's understanding of morality, this type of justification also involves a dependence on the empirical mechanisms that generate the relevant institutions. These mechanisms determine how the counterfactual is to be evaluated. In particular, they determine the closeness facts and thereby determine which of the counterfactual processes satisfying the moral conditions are

²²The compensation principle is to be distinguished from the principle of rectification. (Davis: 1977, p. 220, for instance, conflates these principles.) The former is concerned with a special range of cases in which a boundary crossing can be justified by compensating the person for the disadvantages that result from prohibiting him or her from performing certain risky activities, whereas the latter is concerned with what needs to be done to right past wrongs, whereby this rectification in no way makes it the case that the boundary crossings are justified. The compensation principle plays a role within the ideal counterfactual history, whereas the principles of rectification connect the counterfactual history to the actual history.

the ones the outcomes of which are to be brought about. Whilst all counterfactual processes satisfying the moral conditions are on a par as far as the normative framework is concerned, the descriptive part of the theory differentiates amongst these candidates in terms of their closeness to the actual world.

As a result, one ends up with an interesting dependence on positive social science, insofar as one needs to appeal to empirical mechanisms to determine which set of institutions would have arisen within the moral constraints. Since it is necessary to specify both moral constraints and empirical mechanisms in order to evaluate this counterfactual, the normative project of justifying the state becomes intertwined with the empirical project of explaining the state. This is because justification becomes a function of explanation in ideal circumstances. The rectificationist principles connect the explanation as to how institutions arise in ideal circumstances with the account as to which institutions are justified in the actual world.

This implies that one's views about issues such as network externalities and public goods will have a direct impact on the nature of the institutions that will be justified. For instance, different sets of institutions will be justified, whilst holding Nozick's moral framework fixed, depending on whether one accepts the empirical claim that a dominant agency would arise, which is based on Nozick's understanding of the network externalities involved in the provision of protection services. If the claim about network externalities is accepted, then the counterfactual will take us to a situation in which a minimal state arises. Yet, if this claim is rejected, then the counterfactual will take us to an anarchic situation in which a plurality of different protective agencies compete, without any of them being dominant and without any of them having a (de facto) monopoly. This means that Nozick's commitment to individual rights could equally lead one to anarchism as to the minimal state, if one were to adopt a different view about the empirical question as to how protective agencies operate and interact.

In the same way that one can vary the account of the empirical mechanisms to end up with different institutional structures being justified, one can also vary the account of morality. This method of justification can, accordingly, be generalised beyond Nozick's framework by modifying the background understanding of rights. A modification of the understanding of rights that is operative will affect the nature of the political institutions that can be justified in this manner, given that the counterfactual scenario picked out by the rectificationist principle is dependent on what moral constraints are in place.

For instance, if an alternative moral framework that includes positive rights were to be adopted, then this could necessitate a transition from a minimal state to a more-than-minimal state, in the same way that the transition from an ultra-minimal state to a minimal state is required due to the compensation principle within the context of Nozick's theory. In particular, if a minimal state were to violate the positive rights of its citizens – for instance if it should fail to provide

them with the level of welfare/resources to which they have a rightful claim – then this would constitute a wrong that would have to be rectified. The way in which it would have to be rectified (at the institutional level) would involve bringing into existence the institutions that are to be found in ideal circumstances in which no rights are being violated by the state. Since these counterfactual circumstances involve more-than-minimal states that ensure that the positive rights of their subjects are fulfilled, it would follow that the rights-violating minimal state would have to be transformed into this type of more-than-minimal state in order for the injustice resulting from the violation of positive rights to be rectified.²³

The fact that the rectificationist account generates a comparative notion of justification enables it to deal with cases in which rights fail to be compossible, as happens in the case of positive rights that are not co-satisfiable.²⁴ There is a concern that if the compossibility condition is not satisfied, then one ends up with all no-rights-violations counterfactuals being vacuously true. If rights are not compossible, then rights violations are unavoidable, which means that there is no ideal scenario in which the moral constraints are fully complied with. Accordingly, the no-rights-violations antecedent will be impossible, thereby ensuring that any counterfactual with this antecedent will be vacuously true and that any set of institutions is, accordingly, trivially justified.²⁵

In order to avoid this problem, one has to require the non-vacuous truth of the relevant counterfactual. Since there will be no non-vacuously true no-rights-violations counterfactual, no institutional structures will be fully justified. But this does not imply that all institutions will be equally unjustified. Instead, one can make various comparative claims. In particular, one can evaluate what institutions would have arisen in the situation in which the extent of rights violations is minimised. Rather than evaluating what would have happened had no rights been violated, one evaluates what would have happened had rights been violated

²³In the case of positive rights, the distinction between rectification at the level of institutional structures and the level of property distributions that was drawn above is particularly clear, and it would seem that, in addition to transforming the state into a more-than-minimal state, a rectification of property holdings would have to take place, which presumably would require the state to make up for arrears.

²⁴Some consider compossibility to be a necessary requirement on a system of rights, especially those defending a choice theory of rights (cf. Steiner: 1977). An interest theory of rights, by contrast, is likely to lead to non-compossibility (cf. Waldron: 1989, p. 503).

²⁵The requirement that the no-rights-violation condition of the counterfactual be satisfied not by supposing that the relevant rights had simply been alienated, waived or forfeited is particularly important in this context, since this condition could otherwise be satisfied even when dealing with non-compossible rights. That is, a rights violation can be avoided in two ways: 1. by means of the perpetrator acting differently, or 2. by means of the victim having different rights. The counterfactual relevant for rectification identifies scenarios in which the rights-violating action is not performed or does not succeed, i.e. the rights are held fixed in evaluating the counterfactual, rather than a scenario in which the action is performed but does not classify as a rights violation. In short, the condition of there not being rights violations is to be satisfied by getting rid of violations rather than by getting rid of rights.

to the minimal extent possible. This counterfactual then determines which institutional structures are maximally justified and hence what needs to be done to rectify all rectifiable injustices.²⁶

2.5 Hypothetical justifications

Rectification principles involve counterfactuals. This means that hypothetical facts have to be brought in to justify political institutions (except in ideal circumstances in which there are no injustices that need to be rectified and in which the justification is purely historical). Whilst being hypothetical in nature, these principles differ radically from traditional hypothetical accounts for justifying the state. In particular, traditional accounts are not understood as supplementing historical principles in order to address non-ideal circumstances, but are instead taken to constitute complete theories that are meant to replace historical accounts. Such hypothetical principles are thus construed as independent principles that apply unrestrictedly, and not as restricted principles that are parasitic on certain logically prior principles that they are meant to supplement.

I. JUSTIFICATORY FORCE

The rectificationist type of justification is partly hypothetical in nature. However, it is not vulnerable to the objections that are usually raised against hypothetical justifications. Any historical justification, such as a consent-based justification, needs to be actual – “hypothetical contracts do not supply an independent argument for the fairness of enforcing their terms. A hypothetical contract is not simply a pale form of an actual contract; it is no contract at all” (Dworkin: 1976, pp. 17-18). The fact that hypothetical contracts lack normative force in the actual world implies that traditional hypothetical consent models are at risk of collapsing into teleological models. Since hypothetical contracts are not actually binding, they can, at best, indicate that a certain state of affairs is rational or desirable.²⁷ Hypothetical explanations will then lack justificatory force in their own right and will only be able to play a role insofar as they point to teleological justifications (cf. Schmidt: 1990). In this way, hypothetical accounts that, on the face of it, seemed to be forms of what Simmons calls transactional justifications collapse into generic justifications (cf. Simmons: 1999, p. 764). The appeal to hypothetical scenarios then becomes merely a heuristic device. Since idealised

²⁶Additionally, any plausible theory that allows for there to be conflicts of rights should identify a procedure or mechanism for resolving such conflicts. Accordingly, one should not assess the situation in which no rights are violated, but rather the situation in which there are no unresolved rights violations.

²⁷Cf. “What we ‘could agree to’ has prescriptive force for the Hobbesians, not because make-believe promises in hypothetical worlds have any binding force but because this sort of agreement is a device that (merely) reveals the way in which the agreed-upon outcome is rational for all of us” (Hampton: 2007, p. 482; also cf. p. 486).

agents would consent to those principles or institutions that are independently justified, entertaining hypothetical situations can help us to identify what reasons there are and what institutions are justified, without providing any independent support.^{28, 29}

Unlike traditional hypothetical accounts, the appeal to hypothetical scenarios that forms part of a rectificationist justification is not a mere heuristic device that helps us to identify which institutions have the relevant features that are required for them to be teleologically justified. Instead, the principles of justification themselves are specified counterfactually, i.e. these principles have counterfactual content. In order to determine what they require, one needs to determine what happens, in particular what institutional structures emerge, in the hypothetical circumstances in which agents do not violate the moral requirements specified by the historical principles. What type of institutional structure needs to be brought about is thus a function of counterfactual facts. In this way, counterfactual facts have justificatory force. They determine which institutions are justified and what is required if actual injustices are to be rectified.

2. IDEALISATION

Relatedly, the notion of 'ideal' circumstances differs from that which features in standard hypothetical accounts. Since the rectificationist account is not concerned with considering what idealised agents would choose or consent to in order to identify a privileged set of institutions, the idealisation does not pertain to the nature of the agents and the conditions in which they choose principles or create institutions. There is no need to idealise away their biases and imperfect information, nor to suppose that they are fully rational or 'reasonable'. Instead, what is being idealised is the history that generates certain institutions, given that one is concerned with what would have happened had there not been any injustice. The idealisation, in this way, concerns the behaviour of the agents (in particular, the behaviour of those acting on the part of the institutions, i.e. agents of the state). Their behaviour is idealised insofar as one supposes that they do not violate the normative requirements that bind them.³⁰

²⁸Given that it would seem that in order to determine whether idealised agents would consent to certain principles one would already have to know whether these principles are justified, hypothetical accounts would appear to presuppose the very thing that they are meant to establish.

²⁹An alternative understanding of hypothetical consent is based on a constructivist approach, according to which certain principles turn out to be justified precisely because they would be selected in the idealised circumstances. Appealing to hypothetical situations in that case becomes a matter of determining what the correct principles are. This type of approach, however, can (at best) only justify principles, but it cannot underwrite the authority of the state (cf. Stark: 2000).

³⁰To idealise the behaviour of the agents need not imply supposing that they are virtuous. All that is required of the relevant ideal scenario is that no rights violations occur, which need not imply that no rights violations are attempted. What exactly is required to implement the counterfactual supposition that no rights violations occur depends on how rights violations are construed, in particular on how much of the causal chain they implicate, i.e. on whether the

Given that rectification principles supplement rather than replace historical principles, idealisation is only required if the actual history diverges from the historical principles. By contrast, facts about the actual history drop out entirely in the case of hypothetical consent accounts. Such accounts do not assign any significance to historical principles – the fact that people actually consented or failed to consent is of no significance. What matters is not actual consent but idealised consent, such that if idealised agents would not consent to a certain institutional set-up, then those institutions will not be justified even if they have been generated on the basis of actual consent. Only if people actually happen to be ideal in the relevant way will their consent suffice for justification, and even in that case what is doing the work is not the fact that they consented but the fact that they are ideal.

3. HISTORY-DEPENDENCE

The fact that the rectificationist account is not a traditional hypothetical account concerned with what suitably idealised actors would choose or agree to implies that what is justified is not a fixed set of institutional structures that are meant to be applicable across the board. Rather than being concerned with the intrinsic suitability of certain institutions, what is at issue is the question as to which institutions are justified in a particular context. Evaluating the counterfactual requires one to take into consideration the closeness of various hypothetical scenarios to the actual world. The relevant ideal counterfactual scenario is thus a function of the particular historical circumstances. If the actual histories differ, then the ideal counterfactual histories can diverge, which, in turn, makes it possible for different institutions to be justified. What classifies as ideal is, accordingly, not constant and fixed. Which institutional structures are justified can thus vary from context to context (though, if Nozick is right, there will not be much variation in practice, since the relevant hypothetical histories will in most cases converge on the minimal state). It is worth noting that this context-variation is not to be explained in terms of the differential suitability of various institutions to present circumstances, but in terms of different societies having different histories.

The kind of reasoning that is traditionally invoked in hypothetical consent arguments does not, however, have to be discarded altogether. Instead, it can play a role in determining what reasonably rational people would have consented to, and hence can help us to determine what would have been likely to emerge in sufficiently ideal circumstances. In other words, this kind of reasoning can inform our understanding of the ways in which the empirical mechanisms would operate in the ideal counterfactual scenario. The advantage of the rectificationist approach is that it does not require excessive idealisation of the agents. Nor is there a need to apply a veil of ignorance, which would bring with it the risk

rights violation is only constituted by the harmful effect, or also by the action that produced it, or even by the agent's intention.

of making the deliberative position too attenuated to generate any substantive content. Instead, one only needs to make the empirical claim that people on the whole respond in reasonable ways to the incentives that they face (whereby these incentives, as well as what classifies as reasonable, can be culturally determined and can vary from context to context).

2.6 Evaluating counterfactuals

In order to determine which institutional structures are justified in a certain context, one needs to evaluate a complicated counterfactual about what would have happened had the relevant historical principles of justification not been violated. Evaluating this kind of counterfactual brings with it a number of difficulties that might be thought to cast doubt on the adequacy of rectificationist justifications.

The first concern is that a rectificationist account is hopeless due to epistemic problems, insofar as the evaluation of such counterfactuals cannot be settled with any degree of determinacy. If it cannot be determined what the rectificationist principle requires in a given situation, then any such account would fail to be applicable and would not yield substantive insights.

In response we can note that, even though problems arise when it comes to evaluating the counterfactuals relevant for rectification, these difficulties are not especially significant in the case of institutional structures (in fact, they are notably less prominent than in the case of property distributions). In particular, the invisible-hand explanation ensures that the resulting outcome is stable, in that it will arise from a large range of initial starting-points. Accordingly, a significant amount of convergence can be expected when concerned with the nature of the institutional structures that will arise in close-by worlds in which there are no (uncompensated) rights violations perpetrated on the part of these institutions, thereby making the evaluation of the relevant counterfactuals relatively straightforward. By contrast, no such convergence can be expected in the case of property distributions, which means that the rectification of violations of property rights faces more serious epistemic obstacles.^{31,32}

³¹Property distributions also result from invisible-hand mechanisms, yet the type of stable features that they involve and in terms of which there is convergence are macro-features, such as those studied by economists, e.g. efficiency. The particular micro-realizations, i.e. who gets what, are not stable, however. Additionally, the rectification of violations of property rights is rendered especially problematic by the fact that property can be transformed, destroyed, improved, and damaged, which makes room for situations in which it is very difficult or even outright impossible to return holdings in the relevant condition to their rightful owners.

³²The sharp separation of rectification regarding property and rectification regarding institutional structures is an oversimplification due to the fact that the territoriality of states implies an inseparable connection with property. As a result, even though there is convergence in terms of institutional structures, there need not be any convergence when it comes to the territories that are governed by the institutional structures in question. There are, however, two mitigating considerations. First, there is likely to be more convergence in the case of territory than in the

Moreover, the comparative conception that recognises degrees of rectification helps us to address epistemic concerns. For practical purposes we do not need a precise evaluation of the counterfactual. All that is required is identifying the rough shape of justified institutions. This limited knowledge allows us to move in the right direction. Even if we do not have a fully determinate characterisation of which institutions are (fully) justified, we can know that certain institutional structures are not justified and that changing them in various ways will lead to a more justified set of institutions. Accordingly, despite only having imperfect information that prevents us from perfectly realising justified institutions, we can nonetheless improve things by approximating the fully justified state of affairs. The applicability of the rectificationist account thus does not impose excessively onerous epistemic requirements.

For instance, in Nozick's theory, any reduction in the size of more-than-minimal states will amount to a greater rectification of past injustices and hence lead to a situation that enjoys a higher degree of justification. By getting closer to a minimal state, one will get closer to the institutions to be found in the relevant ideal counterfactual scenario. This holds true even if there is some uncertainty as to whether the relevant ideal scenario might contain not a minimal state but rather no state at all or only an ultraminimal state. Settling this issue is only required when it comes to completely rectifying past injustices, not when approximating the institutions that are fully justified.

A second concern is that the counterfactual that tries to identify the institutional structures in the closest world in which no rights violations take place runs into difficulties due to non-identity problems. This is because the identity of the people who actually exist may be dependent on the rights violations that took place, such that different people would have existed had these rights violations not occurred. Insofar as the ideal situation contains different people from the non-ideal actual situation, it would not seem to be possible to make sense of the idea that existing people have claims to rectification that can be addressed by bringing the latter situation into conformity with the former.

In order to address non-identity problems, one needs to complicate the antecedent of the counterfactual in such a way that identities are held fixed. This follows from the general principle that one needs to distinguish rights violations from the side-effects to which they give rise, i.e. distinguish facts that involve rights violations from facts that merely depend on rights violations. Since rectification only requires one to address the former, the relevant counterfactual will hold fixed mere side-effects. Although it is often said that rectification amounts

case of movable material goods, due to the limited transferability of territory as well as due to the existence of various geographical focal points that naturally tend to become territorial boundaries. Second, the wide-spread convergence at the institutional level ensures that how geographical space is carved up into different territories is not of all that much significance, at least when we are concerned with the claims that individuals have to (not) be governed by certain types of institutional structures.

to undoing the effects of injustice, this characterisation is too broad since there is no need to undo side-effects of injustice that do not involve rights violations but are only contingently causally connected to rights violations. Applied to the case at hand, this means that, rather than considering the closest scenario in which no rights violations take place, one has to evaluate the closest possible world in which the same people exist and in which there are no rights violations. Put differently, one has to require that rectification proceeds on the basis of an identity-preserving counterfactual.³³

For instance, if x violates y 's rights, whereby one of the side-effects of this rights violation is that z_1 rather than z_2 exists, then the counterfactual that simply assesses what would have happened had no rights been violated will select a world in which z_2 exists. By contrast, the ideal same-people counterfactual will select a world in which there is no rights violation and the 'side-effect' that z_1 rather than z_2 exists is brought about in some other way (that does not involve any rights violations). As long as the identity of z (i.e. whether z_1 or z_2 exists) is not essentially implicated in the wrong-doing but is only contingently causally connected to the rights violation, one can separate out the injustice from the effects that the injustice has on z 's identity. If the injustice that needs to be rectified can be separated out from contingently causally connected processes that have an effect on identities, then one can rectify the injustice whilst holding the latter processes fixed. In this way, we can 'localise' the effect of the rectification of the injustice and then determine the overall effect that results from this rectification together with identity-affecting facts that are being held fixed. For instance, if the rights violation occurs at t_1 , and if the identity-affecting action occurs at t_3 , then one can evaluate what would have happened at t_2 had no rights been violated as well as what would have happened at t_4 given the ideal situation at t_2 together with the identity-affecting action at t_3 . The combined effect of the localised rectification together with the identity-affecting process determines the relevant ideal counterfactual situation with which the actual world is to be made to coincide.

There might be some injustices that are not rectifiable due to non-identity problems, i.e. cases where the identity-affecting fact is not merely a side-effect but involves a rights violation. The antecedent of the ideal same-people counterfactual will then not be satisfiable, thereby making it the case that any such counterfactual will be vacuously true. Yet, even if there should be such cases, one can still use the proposed account to determine what is to be done in order to rectify those injustices that are rectifiable, i.e. those injustices that have corresponding non-vacuously true rectification counterfactuals. This means that one should restrict the antecedent of the counterfactual to those particular injustices that are rectifiable. Rather than assessing what would have happened had

³³Since what matters is that the claims of those who do exist are met, i.e. that they are not suffering any disadvantages as a result of past injustices, all that is required strictly speaking is that the people who actually exist constitute a subset of the people in the counterfactual scenario.

no rights violations occurred, one assesses what would have happened had no rectifiable rights violations occurred.

3 Conclusion

A counterfactual or rectificationist justification of political institutions thus constitutes an interesting (though as yet underdeveloped) alternative to traditional methods of justifying the state that treats the justification of institutional structures as being analogous to that of property distributions, subsuming these domains under a unified theory of justification.³⁴ This kind of justification is applicable in the non-ideal circumstances in which we find ourselves where the actual processes that led to existing political structures are far removed from what would be required for a purely historical justification to be applicable and where the prospects for genuine unanimous consent are beyond being slim.

³⁴Whether the proposed account is ultimately viable depends on whether this parallelism can be substantiated. This paper has focused on exploring this approach, identifying the kind of theory that can be constructed and the desirable features that it has, rather than on providing a foundational theory that undergirds this parallelism.

References

- [1] BADER, R. M. *Robert Nozick*. Continuum, 2010.
- [2] DANLEY, J. R. An examination of the fundamental assumption of hypothetical process arguments. *Philosophical Studies* 34 (1978), 187–195.
- [3] DAVIS, L. Nozick's entitlement theory. In *Reading Nozick: Essays on Anarchy, State, and Utopia*, J. Paul, Ed. Rowman & Littlefield Publishers, 1981, pp. 344–354.
- [4] DAVIS, M. Necessity and Nozick's theory of entitlement. *Political Theory* 5, 2 (1977), 219–232.
- [5] DWORKIN, R. The original position. In *Reading Rawls*, N. Daniels, Ed. Basic Books, 1976, pp. 16–52.
- [6] GAUS, G. Explanation, justification, and emergent properties: an essay on Nozickian metatheory. In *The Cambridge Companion to Nozick's 'Anarchy, State, and Utopia'*, R. M. Bader and J. Meadowcroft, Eds. Cambridge University Press, 2011, pp. 116–142.
- [7] HAMPTON, J. Contract and consent. In *A Companion to Contemporary Political Philosophy*, R. Goodin, P. Pettit, and T. Pogge, Eds., 2nd ed., vol. 2. Blackwell Publishers, 2007, pp. 478–492.
- [8] HOLMES, R. Nozick on anarchism. In *Reading Nozick: Essays on Anarchy, State, and Utopia*, J. Paul, Ed. Rowman & Littlefield Publishers, 1981, pp. 57–67.
- [9] HYAMS, K. Nozick's real argument for the minimal state. *The Journal of Political Philosophy* 12, 3 (2004), 353–364.
- [10] MILLER, D. The justification of political authority. In *Robert Nozick*, D. Schmidtz, Ed. Cambridge University Press, 2002, pp. 10–33.
- [11] NELSON, A. Explanation and justification in political philosophy. *Ethics* 97, 1 (1986), 154–176.
- [12] NOZICK, R. *Anarchy, State, and Utopia*. Basic Books, 1974.
- [13] NOZICK, R. Invisible-hand explanations. *The American Economic Review* 84, 2 (1994), 314–318.
- [14] PAUL, J. The withering of Nozick's minimal state. In *Reading Nozick: Essays on Anarchy, State, and Utopia*, J. Paul, Ed. Rowman & Littlefield Publishers, 1981, pp. 68–76.

- [15] SCHMIDTZ, D. Justifying the state. *Ethics* 101, 1 (1990), 89–102.
- [16] SIMMONS, A. J. Justification and legitimacy. *Ethics* 109, 4 (1999), 739–771.
- [17] SIMMONS, A. J. Consent theory for libertarians. In *Natural Rights Liberalism from Locke to Nozick*, E. Frankel Paul, F. D. Miller, Jr., and J. Paul, Eds. Cambridge University Press, 2005, pp. 330–356.
- [18] STARK, C. Hypothetical consent and justification. *The Journal of Philosophy* 97, 6 (2000), 313–334.
- [19] STEINER, H. The structure of a set of compossible rights. *The Journal of Philosophy* 74, 12 (1977), 767–775.
- [20] THOMSON, J. J. *The Realm of Rights*. Harvard University Press, 1990.
- [21] WALDRON, J. Rights in conflict. *Ethics* 99, 3 (1989), 503–519.
- [22] WOLFF, J. *Robert Nozick - Property, Justice and the Minimal State*. Stanford University Press, 1991.