AARHUS UNIVERSITY

Coversheet

This is the accepted manuscript (post-print version) of the article.

Contentwise, the post-print version is identical to the final published version, but there may be differences in typography and layout.

How to cite this publication

Please cite the final published version:

Pedersen, V. M. L. (2019). On the Anti-paternalist Project of Reconciliation. *Utilitas*, 31(1), 20-37.

<https://doi.org/10.1017/S095382081800016X>

Publication metadata

Title:	On the Anti-paternalist Project of Reconciliation
Author(s):	Viki Møller Lyngby Pedersen
Journal:	<i>Utilitas</i> , 31(1), 20-37
DOI/Link:	https://doi.org/10.1017/S095382081800016X
Document version:	Accepted manuscript (post-print)

General Rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

On the Antipaternalist Project of Reconciliation

VIKI MØLLER LYNGBY PEDERSEN

Department of Political Science, Aarhus University

Abstract How should antipaternalists deal with policies that seem to be simultaneously reasonable and paternalistic? In the literature, antipaternalists have sought to show that many policies that prevent people from harming themselves can be justified without appeal to the good accruing to the people interfered with; that is, without appeal to paternalistic reasons. However, while perhaps identifying sufficient nonpaternalistic reasons for supporting these policies under realistic circumstances, antipaternalists often fail, I argue, to identify *satisfactory* reasons that adequately reflect our underlying concerns pertaining to such policies. Included in those concerns are arguably the interests and well-being of the people whose choices are restricted by the policy in question. In this way, this paper reveals that the strategy of reconciling antipaternalism with seemingly paternalistic policies is beset by serious problems.

Keywords paternalism; antipaternalism; the project of reconciliation; the unconscionability doctrine

I. INTRODUCTION

The problem of paternalism, according to Feinberg, requires us to reconcile a liberal opposition to policies aimed at protecting people against self-regarding harm with the fact that many seemingly

reasonable policies apparently do just that.¹ Paradigmatic examples include laws requiring the use of seatbelts or motorcycle helmets in traffic. Confronted with such counterexamples, antipaternalists may either: (i) argue that the apparently paternalistic policies are in fact not so and hence that they can safely be endorsed; or (ii) reject the policies despite their initial appearance of being reasonable.

The so-called ‘project of reconciliation’² has been the source of a number of ingenious attempts to justify a range of seemingly paternalistic policies by invoking nonpaternalistic reasons in their favor. For example, Seana Shiffrin defends the unconscionability doctrine, which permits courts to refuse to enforce unconscionable contracts.³ Joel Feinberg defends motorcycle helmet laws.⁴ Elizabeth Anderson invokes a nonpaternalistic rationale for mandatory contributions to health insurance programs,⁵ and Paul Bou-Habib defends policies making it compulsory for imprudent people to insure themselves.⁶ Common to these arguments is that they seek to identify nonpaternalistic reasons that justify the relevant policies without appeal to the good accruing to the people interfered with.

¹ J. Feinberg, *Harm to Self* (New York and Oxford, 1986), pp. 25, 135, 142. See also P. de Marneffe, ‘Avoiding Paternalism’, *Philosophy & Public Affairs* 34 (2006), pp. 68-94, at 68; J. Kleinig, *Paternalism* (New Jersey, 1984), p. 82.

² This term is attributable to Peter de Marneffe, who finds inspiration from Feinberg’s characterization of the strategy. See de Marneffe, ‘Avoiding Paternalism’, p. 68; Feinberg, *Harm to Self*, p. 25.

³ S. Shiffrin, ‘Paternalism, Unconscionability Doctrine, and Accommodation’, *Philosophy & Public Affairs* 29 (2005), pp. 205-250.

⁴ Feinberg, *Harm to Self*, pp. 134-142.

⁵ E. Anderson, ‘What is the Point of Equality?’, *Ethics* 109 (1999), pp. 287-337, at 314, 316, 319, 329-330.

⁶ P. Bou-Habib, ‘Compulsory Insurance Without Paternalism’, *Utilitas* 18 (2006), pp. 243-263.

However, what exactly a satisfactory reconciliatory argument must look like is severely undertheorized in the literature. This paper proposes key desiderata pertaining to such arguments. Furthermore, I counter what appears to be a widely shared underlying assumption in the literature, namely that insofar as one may present a nonpaternalistic reason that is sufficient to justify the policy in question, one has shown that the policy is simply not paternalistic.⁷ I do this by drawing a distinction between sufficient and satisfactory reasons for policies. While antipaternalists undertaking the project of reconciliation may identify *sufficient* reasons for endorsing seemingly paternalistic policies under realistic circumstances, they often fail to provide *satisfactory* reasons that adequately reflect our concerns pertaining to such policies. I illustrate this problem of the procedure of the project of reconciliation by focusing on Seana Shiffrin's nonpaternalistic defense of the unconscionability doctrine. I argue that, even if Shiffrin's nonpaternalistic reason may allow antipaternalists to conclude that the state should not enforce unconscionable contracts in practice, it is implausible to exclude the paternalistic concern pertaining to such agreements from the justification of the doctrine. Thus, in my understanding, the policy should be considered paternalistic. I demonstrate this point by constructing a hypothetical case which reveals some serious deficiencies of antipaternalism in theory.

In light of this problem with option (i), antipaternalists may still resort to (ii) – that is, the strategy of jettisoning what initially appeared to be a reasonable policy. In many cases, however, this is not, I argue, a viable or attractive strategy. Accordingly, a third strategy presents itself as a plausible candidate, namely to endorse the policies on paternalist grounds. This means, of course, that the antipaternalist approach to such policies proposed by Feinberg would have to go. In my view, however, this is not regrettable.

⁷ See e.g. de Marneffe, 'Avoiding Paternalism', p. 74.

The paper is structured as follows. First, I briefly account for paternalism and present the common procedure of the project of reconciliation. Second, I introduce Shiffrin's nonpaternalistic argument for the unconscionability doctrine. Third, I draw a distinction between sufficient and satisfactory reasons. I argue that even if Shiffrin's nonpaternalistic argument is sufficient to justify the unconscionability doctrine on nonpaternalistic grounds, her argument is not satisfactory in that it does not adequately reflect our reasons for endorsing the doctrine. Fourth, I examine some alternative nonpaternalistic arguments for the unconscionability doctrine and argue that paternalistic considerations still play a decisive role in the justification. Finally, I conclude that there are serious problems attached to the procedure of the project of reconciliation, which seems to refute the plausibility of this antipaternalist strategy.

II. PATERNALISM AND THE PROJECT OF RECONCILIATION

This paper's argument relies on a reason-based understanding of the distinction between paternalism and antipaternalism. According to this, paternalism is plausibly understood as the position according to which the good or benefit to the agent interfered with is accepted as a good and relevant reason in favor of infringing the agent's autonomy. To say that reasons of this kind are good and relevant implies that they are valid and should count when evaluating the justifiability of policies and actions. They need not be decisive, however. Paternalists can accept that so-called paternalistic reasons are sometimes outweighed by other considerations. They accept paternalistic reasons *amongst others*; accordingly, it is possible to be a paternalist while arguing against actions

aimed at benefitting the agents interfered with in different cases. This understanding of paternalism is primarily gleaned from Feinberg's liberty-limiting principle of legal paternalism.⁸

In contrast, antipaternalists reject reasons that are paternalistic and thus accept a narrower scope of reasons for infringements with people's voluntary choices. According to the antipaternalist, prevention of self-regarding harm will never be a relevant reason in favor of interferences. Antipaternalists believe, Feinberg argues, that such reasons 'are morally illegitimate or invalid . . . by their very natures, since they conflict head on with defensible conceptions of personal autonomy'.⁹ Policies restricting people's voluntary choices that can only be justified by appeal to paternalistic reasons are therefore unacceptable to antipaternalists.

As described in the introduction, the procedure of the antipaternalist project of reconciliation has been to identify or develop sufficient nonpaternalistic arguments for some policies that seem to be both reasonable and paternalistic. By sufficient nonpaternalistic arguments, I mean arguments that suffice to justify the relevant policies without counting in their

⁸ Feinberg, *Harm to Self*, pp. 4, 25-26. Feinberg's principle focuses, however, on criminal prohibitions rather than, more broadly, infringements of autonomy. As a number of theorists have argued, Feinberg's focus is plausibly too narrow to embed the worries of antipaternalists. See e.g. J. Begon, 'Paternalism', *Analysis Reviews* 76 (2016), pp. 355-373, at 357; G. Dworkin, 'Paternalism: Some Second Thoughts', *The Theory and Practice of Autonomy*, ed. G. Dworkin (Cambridge, 1988), pp. 121-129, at 123.

⁹ Feinberg, *Harm to Self*, pp. 26. This is a rather strong form of antipaternalism. At the same time, the opposite paternalist position appears somewhat weak if one can be a paternalist without ever finding the paternalistic reason decisive. Accordingly, it may be argued that the class of antipaternalists should also include those who believe that the reason for preventing self-regarding harm is relevant and valid, but *always* outweighed by reasons of respecting individual autonomy. Whether the former or latter reason-based distinction between paternalism and antipaternalism is preferable is unimportant to my argument, which is compatible with both conceptions.

favor the benefits to the people interfered with.¹⁰ On the face of it, this procedure of the project of reconciliation looks plausible because it seems true that, as expressed de Marneffe, if ‘a general principle of antipaternalism is valid, then we should evaluate these policies by evaluating whether or not there is sufficient *nonpaternalistic* reason for them’.¹¹ If a sufficient nonpaternalistic justification is available, the seemingly reasonable, but apparently paternalistic, policies need not constitute counterexamples to antipaternalism as one may defend the policies in question without embracing paternalism. Richard Arneson expounds the thought behind the approach: ‘since a paternalist law is one that is justified, if at all, by paternalistic reasons, finding an alternative rationale that justifies the law defeats the counterexample’.¹² Accordingly, if successful, the project of developing sufficient nonpaternalistic arguments for seemingly paternalistic policies may be valuable for antipaternalists because it, in some way, shows that seemingly reasonable policies are not ‘intrinsically paternalist’ in that they *need not* be justified paternalistically.¹³

As I will argue in the following sections, this widely accepted procedure of the project of reconciliation is seriously flawed. I argue that the procedure sets the bar too low with respect to the criteria that relevant nonpaternalistic reasons should meet. In Sections IV and V, I illustrate these problems in the context of Seana Shiffrin’s nonpaternalistic defense of the unconscionability doctrine. First, however, I briefly recapitulate Shiffrin’s argument.

¹⁰ Here, I am greatly inspired by de Marneffe, ‘Avoiding Paternalism’, p. 74.

¹¹ de Marneffe, ‘Avoiding Paternalism’, p. 69.

¹² R. Arneson, ‘Joel Feinberg and the Justification of Hard Paternalism’, *Legal Theory* 11 (2005), pp 259-284, at 272. Relatedly, de Marneffe defends a definition of paternalistic policies which implies that ‘if there is sufficient nonpaternalistic reason for a policy, then it is not paternalistic’. De Marneffe, ‘Avoiding Paternalism’, p. 74.

¹³ Shiffrin, ‘Paternalism, Unconscionability Doctrine, and Accommodation’, pp. 208, 207, 212; see also de Marneffe, ‘Avoiding Paternalism’, p. 77.

III. SHIFFRIN AND THE UNCONSCIONABILITY DOCTRINE

Shiffrin defends the unconscionability doctrine that ‘enables a court to decline to enforce a contract whose terms are seriously one-sided, overreaching, exploitative, or otherwise manifestly unfair’¹⁴ by identifying a nonpaternalistic reason in its favor. The unconscionability doctrine has been characterized as paternalistic in that it repeals a voluntary agreement between two consenting persons because the terms of the contract seem detrimental to one party. In this way, as Shiffrin puts it, the doctrine prevents the voluntary agent ‘from making a *binding* contract that he actually wills on the grounds that, in some serious way, it does not promote his interests to be bound or to comply’.¹⁵ This paternalist classification of the doctrine has separated theorists into two groups: those who, in view of the apparent reasonableness of the doctrine, believe that an acceptable role for paternalism can be found; and those who give up the doctrine.¹⁶

In contrast to these positions, Shiffrin denies that the unconscionability doctrine is reserved to paternalists. She argues that the doctrine can be justified without appeal to paternalistic reasons. Her argument has three steps (where I will focus on the third). First, she clarifies how ‘the legal institution of contract requires, through their role in enforcement, the participation and cooperation of parties outside the agreement – that is, it requires the cooperation of the community, as it is embodied in the state’.¹⁷ Second, she argues that people are not entitled to the community’s

¹⁴ Shiffrin, ‘Paternalism, Unconscionability Doctrine, and Accommodation’, p. 205.

¹⁵ Shiffrin, ‘Paternalism, Unconscionability Doctrine, and Accommodation’, p. 211.

¹⁶ Shiffrin, ‘Paternalism, Unconscionability Doctrine, and Accommodation’, p. 207.

¹⁷ Shiffrin, ‘Paternalism, Unconscionability Doctrine, and Accommodation’, p. 221.

unconditional assistance in entering binding agreements. Third, she argues that the community can refuse to offer such assistance for reasons that are not paternalistic.¹⁸

The nonpaternalistic reason that Shiffrin identifies is, roughly, that the state has (or third parties in general have) ‘a *self-regarding concern* not to facilitate or assist harmful, exploitative, or immoral action’¹⁹ – an interest not to be co-responsible or complicit. Shiffrin illustrates this with reference to a similar case:

... it would be paternalist for me to hide your cigarettes to protect your health. Nonetheless, it would not be paternalist (and may be morally required) for me to refuse to buy you cigarettes or to refuse to retrieve them from a pilfering acquaintance if my motive for refusal is that I think that *I* should not perform substantial actions that contribute to your addiction or illness.²⁰

To some extent, this reason appeals to the harms that you incur by smoking. It does so, however, only in an indirect sense – that is, to the extent that *my* interests are involved. To see why the reason in question is not paternalistic, we can imagine that I know that if I do not buy you the cigarettes, another person will. Accordingly, my refusal in the situation makes no difference to the protection of your health. In spite of these added facts, it seems that I still have reason not to get involved, and *this* is the reason that Shiffrin appeals to in the context of the unconscionability doctrine. As she points out, ‘An analogous claim may be made with contract: there are some

¹⁸ Shiffrin, ‘Paternalism, Unconscionability Doctrine, and Accommodation’, p. 221.

¹⁹ Shiffrin, ‘Paternalism, Unconscionability Doctrine, and Accommodation’, p. 224 [emphasis added].

²⁰ Shiffrin, ‘Paternalism, Unconscionability Doctrine, and Accommodation’, p. 224.

agreements you have a right to form but no right to assistance in carrying them out and about which others may reasonably feel that they may or even must not assist'.²¹

This justification does not challenge people's right to make unconscionable agreements (or right to smoke, for that matter). It challenges their right to have the conduct assisted or facilitated by parties outside the contract. If the community has an interest in not assisting or facilitating 'harmful, exploitative, or immoral action', then the state's enforcement of unconscionable contracts would involve a setback to the interests of the community. This implies that the state may be morally justified in refusing to enforce unconscionable contracts with reference to the interests of third parties. This reason for the unconscionability doctrine does not appeal to the benefits achieved by the potential promisors and is thus not paternalistic.

Shiffrin's nonpaternalistic argument may be a sufficient reason for the unconscionability doctrine (in the sense that it suffices to warrant the doctrine without appeal to paternalistic reasons). It is not, however, a satisfactory justification for the unconscionability doctrine; at least I argue so below in Section IV. Specifically, I present a case in which it would be reasonable to prevent or interfere with a voluntary, but unconscionable, contract. That is, I contest that it is permissible for people to enter certain such agreements. This serves to cast doubt on whether the justifiability of the unconscionability doctrine should be based on the reason that Shiffrin appeals to: in the case in question, the factors driving Shiffrin's argument are absent, but our opposition to the contract in question remains. In all fairness, I should point out that Shiffrin apparently makes the assumption that people have a right to form unconscionable contracts in order to share as much ground as possible with opponents of the unconscionability doctrine.²² That

²¹ Shiffrin, 'Paternalism, Unconscionability Doctrine, and Accommodation', p. 224.

²² See e.g. Shiffrin, 'Paternalism, Unconscionability Doctrine, and Accommodation', p. 210.

is, she might embrace this assumption merely for the sake of argument. Nevertheless, if the nonpaternalistic argument that Shiffrin presents adequately explained why we should endorse the unconscionability doctrine, it is hard to see why people should not be allowed to form unconscionable agreements even though they should not necessarily expect the state to facilitate such agreements.

IV. IS SHIFFRIN'S NONPATERALISTIC REASON SATISFACTORY?

As described previously, reconciliation theorists seek to avoid the objection that antipaternalists are forced by their principle to reject some seemingly plausible policies and actions. In the context of the unconscionability doctrine, the reason that Shiffrin proposes may suffice to explain why antipaternalists would recommend the doctrine. In fact, the circumstances may be such that the enforcement of unconscionable contracts would always give rise to a sense of co-responsibility or complicity; that is, would always involve setbacks to the interests of third parties. Accordingly, it may as good as always be possible to justify the unconscionability doctrine with appeal to this nonpaternalistic reason (and without appeal to the prevented harm to potential promisors). In that case, Shiffrin seems to have offered *sufficient* nonpaternalistic reason for the unconscionability doctrine. Thereby, her argument may seem to fend off those critics who, for example, were worried that unconscionable agreements between consenting parties would have to be enforced in compliance with the antipaternalist principle. Her argument is a response to the objection that antipaternalists must reject the unconscionability doctrine, even though the doctrine is, intuitively, plausible.

Even if the factors that Shiffrin appeals to would always be supportive of the unconscionability doctrine, her nonpaternalistic response nonetheless fails to address another

objection, namely that the nonpaternalistic reason is not *satisfactory*, meaning that it fails to adequately reflect our ultimate concerns pertaining to the doctrine. If our hostility to unconscionable agreements is not contingent on the concern that Shiffrin appeals to, our endorsement of the unconscionability doctrine should arguably not rely upon the presence of such concerns.

While paternalists and antipaternalists seem to agree that the state is permitted not to enforce unconscionable contracts, the agreement is based on different reasons, which Shiffrin's nonpaternalistic defense of the unconscionability doctrine illustrates. In deciding whether or not the state is allowed to not enforce unconscionable contracts, antipaternalists reject the independent relevance (or decisive relevance) of the interests or well-being of potential promisors. Instead, Shiffrin's nonpaternalistic argument takes the interests of those whose assistance is required to uphold the contracts in question to be crucial. When we refuse to enforce an unconscionable contract, Shiffrin argues, we seek protection *for ourselves* and cater to our own interests in not getting involved. If the harm prevented to potential promisors is relevant, it is so only in an indirect sense to the extent that it influences the interests of others.

In my view, this justification of the unconscionability doctrine is unsatisfactory in that it suffers from the shortcoming referred to above: it fails to address some of our crucial concerns about unconscionable agreements. In what follows, I seek to substantiate this point that our opposition to enforcing unconscionable agreements is not (or, at least not predominantly) grounded in the nonpaternalistic consideration that Shiffrin emphasizes. I do so by targeting the assumption of Shiffrin's nonpaternalistic defense of the unconscionability doctrine that people have a right to form unconscionable agreements. More specifically, I present a case in which Shiffrin's nonpaternalistic reason is not available, but we seem to have strong inclinations towards repealing

a contract between consenting agents. The reason that drives the intuition in this case plausibly also plays an important role in the context of the unconscionability doctrine. In this way, the case casts doubt on whether the nonpaternalistic argument that Shiffrin identifies encompasses the considerations that are relevant and important to the justification of the unconscionability doctrine. If not, we should not let the justification of the doctrine hinge on this nonpaternalistic argument.

Consider the following hypothetical case, which draws some inspiration from Irving Kristol's example of Gladiatorial contests in Yankee Stadium.²³ (The case I am about to suggest does not, however, take place in a sold-out Yankee Stadium, but will be settled without spectators.) Rita has signed a contract according to which she agrees to enter into gladiatorial combat against Robert. Robert is a professional and well-armed gladiator, and Rita will only be equipped with a blunt bread knife; but she sees this as a challenge that makes entering the game more tempting. Rita knows exactly what she is doing, and her choice can be characterized as voluntary. No one pressures her into complying with the contract. Everyone is indifferent. She can refrain from performing her part of the contract at any time. However, once the agreement is formed Rita will not do anything herself to break out of it. If nobody intervenes, she will take part in the gladiatorial combat and thus assume a very high risk of incurring grave and irreversible harm, indeed fatal such.

In this case, there is obviously no need for the state to facilitate or assist the unconscionable agreement between the two consenting agents in question. Accordingly, the issue is no longer whether the agreement should be enforced, but whether the state should intervene or

²³ I. Kristol, 'Pornography, Obscenity, and the Case for Censorship', *New York Times Magazine*, 28 March 1971, at p. 24. Feinberg refers to this case by Kristol in J. Feinberg, *Harmless Wrongdoing* (New York and Oxford, 1988), pp. 128-129.

prevent the combat from taking place to begin with. For the same reason, Shiffrin's nonpaternalistic justification of the unconscionability doctrine is not available in the case of Rita and Robert. According to Shiffrin, the unconscionability doctrine is justifiable because it protects others from having to assist or facilitate unconscionable agreements. As I have emphasized earlier, her defense of the unconscionability doctrine assumes that people have a right to form unconscionable agreements. What they do not have a right to is 'assistance in carrying them out'.²⁴ It is the former right – to form unconscionable agreements – that is contested if we believe that the state is justified in preventing or interfering with the combat.

The view that we do not have any good reasons for barring Rita from entering into the combat seems, to me, implausible. I suspect that many liberals would agree.²⁵ The principle of consent arguably has limits.²⁶ If I am right, the case of Rita and Robert does not necessarily undermine Shiffrin's nonpaternalistic argument as *a* reason for the unconscionability doctrine. It illustrates, however, that in cases of unconscionable agreements between consenting agents, we have other crucial considerations. Those considerations are not reflected in Shiffrin's argument, but are arguably of relevance and importance to the justification of the unconscionability doctrine. If people should not always be allowed to form unconscionable contracts, this concern significantly contributes to justifying the state not enforcing the relevant agreements. In addition, the concern challenges the notion that we would not want the unconscionability doctrine were it not for the nonpaternalistic reason that Shiffrin identifies. In this way, the case of Rita and Robert

²⁴ Shiffrin, 'Paternalism, Unconscionability Doctrine, and Accommodation', p. 224.

²⁵ However, it should be pointed out that one *Utilitas* referee does not share my intuition in this case.

²⁶ Cf. V. Tadros, 'Consent to Harm', *Current Legal Problems* 64 (2011), pp. 23–49.

underpins that it is unsatisfactory to let the justification of the unconscionability doctrine rely on Shiffrin's argument.

It is, however, not given that the paternalistic justification is salient in the above case, in that other nonpaternalistic arguments may justify intervention. Accordingly, the case may not represent a hypothetical counterexample to antipaternalism in general. Nevertheless, in Section V, I examine a number of possible nonpaternalistic arguments for intervention and argue that it is still possible to identify cases in which we are moved by paternalistic considerations. If I am right, the paternalistic argument justifies the lack of a right to form certain unconscionable agreements.

V. ALTERNATIVE NONPATERNALISTIC REASONS?

In explaining what is wrong with unconscionable contracts, antipaternalists may, as pointed out above, appeal to nonpaternalistic reasons other than the one that Shiffrin invokes. If they succeed in this, my concerns about Shiffrin's argument cannot be seen as part of a broader objection to antipaternalism.

In a famous but puzzling passage of *On Liberty*, John Stuart Mill rejects that people should be free to form a certain type of unconscionable contract, namely voluntary slavery contracts.²⁷ Mill's briefly outlined reason for interfering with a person's voluntary act in this way is based on the following considerations for this person's liberty: 'The principle of freedom cannot require that he should be free not to be free. It is not freedom to be allowed to alienate his freedom.'²⁸ Various authors have since pointed out that, in this particular case, Mill seems to resort to

²⁷ J. S. Mill, *On Liberty* (New York, 1956), p. 125.

²⁸ Mill, *On Liberty*, p. 125.

paternalism.²⁹ However, in an effort to reconcile Mill's intolerance of voluntary slavery contracts with his antipaternalistic principles, Feinberg examines four nonpaternalistic arguments for disallowing slavery agreements.³⁰ If these arguments work, they may justify interference with the unconscionable agreement between Rita and Robert as well.

Feinberg's first argument is a soft paternalistic justification according to which it is plausible to assume that no one would voluntarily sell him- or herself into slavery. Along the same lines, it may be assumed that no sane, poorly armed amateur would enter gladiatorial combat against a well-armed professional. Nevertheless, as Feinberg argues with respect to voluntary slavery contracts, while this may be true, this truth 'is not an *a priori* one but rather one that must be tested anew in each case by the application of independent, non-circular criteria of mental illness'.³¹ Hence, the soft paternalistic reason does not rule out that voluntary agents should be allowed to e.g. permanently enslave themselves or take part in suicidal gladiatorial combat.³²

Feinberg develops this argument by pointing out that the state may be justified 'in presuming nonvoluntariness conclusively in every case as the least risky course'.³³ The rationale is that, *in practice*, it is both difficult and costly to test voluntariness – and that it would be worse

²⁹ See e.g. J. Feinberg, 'Legal Paternalism', *Canadian Journal of Philosophy* 1 (1971), pp. 105-124, at pp. 116-117; R. Arneson, 'Mill versus Paternalism', *Ethics* 90 (1980), pp. 470-489, at p. 486; G. Dworkin, 'Paternalism', *The Monist* 1 (1972), pp. 64-84, at pp. 75-76.

³⁰ Feinberg, 'Legal Paternalism' pp. 117-119.

³¹ Feinberg, 'Legal Paternalism', p. 117; see also Dworkin, 'Some Second Thoughts', p. 126.

³² In fact, Feinberg points out some motives that people may have for voluntarily consenting to become slaves. See Feinberg, 'Legal Paternalism', p. 117.

³³ Feinberg, 'Legal Paternalism', p. 119.

(mistakenly) to allow nonvoluntary slavery than to forbid voluntary slavery contracts (as a precaution). Again, the argument holds that self-regarding and voluntary slavery agreements are unobjectionable in theory. For the same reason, the above soft paternalistic justifications are of no avail to antipaternalists in the above hypothetical case in which we *know* per assumption that Rita is acting voluntarily. The arguments imply that Rita should be permitted to fight.

Feinberg's second proposal concerns the protection of third party interests and is thus related to Shiffrin's nonpaternalistic reason for the unconscionability doctrine. In contrast to Shiffrin, Feinberg does not appeal to the costs of facilitating or assisting unconscionable agreements, but rather to the costs of witnessing or remedying the effects of such agreements. According to Feinberg, '[t]here are certain risks ... of an *apparently* self-regarding kind that men cannot be permitted to run, if only for the sake of others who must either pay the bill or turn their backs on intolerable misery'.³⁴ It is an open question what the more specific content of 'the bill' in the slavery case is. Feinberg mentions, for example, the costs involved in witnessing runaway slaves being forced back to their owners.³⁵ But in the case of Rita, there will be no such costs in that enforcement is not needed to uphold the agreement. Feinberg may be right, however, that unconscionable agreements can have certain negative financial or emotional effects on the interests of third parties and that these warrant interference by the state.

In my view, this third party cost argument by Feinberg does not form the genuine basis of our resistance to the agreement between Rita and Robert. And whether it warrants the prevention or interference with unconscionable agreements between consenting agents seems, at

³⁴ Feinberg, 'Legal Paternalism', p. 119.

³⁵ Feinberg, *Harm to Self*, p. 75.

least, to be a very contingent matter.³⁶ To test it properly, let us revise the case so that there are no such costs and see whether our reluctance to let Rita fight persists.³⁷ For example, let us suppose that Rita has no relatives or friends who would be deeply affected by her choice.³⁸ Moreover, we may assume that the state requires that the parties of the contracts purchase insurance that will cover potential costs which may result from Rita's misery (e.g. to finance hospital treatment). Despite these stipulations, it indeed seems to be the case that we should still be concerned with the plight of Rita. Her very probable loss of well-being seems to have an independent pull on us. If I am right that the case for interfering with Rita's decision to enter the combat is persuasive in this thought experiment, and I have excluded the relevant third party considerations, it seems that antipaternalists may still face a hypothetical counterexample. It is an implication of their view that they would have to permit an unconscionable agreement in a situation where this would be unreasonable.

The third nonpaternalistic reason that Feinberg suggests is a legal moralistic one. This argument focuses on the exploitation involved even in *voluntary* slavery contracts. Most agree that it is deeply immoral to take another person as a slave, which, from a legal moralistic point of view, may explain why people have a duty not to do so. The argument is not paternalistic in that it focuses on preventing people from acting immorally by becoming slave owners rather than

³⁶ This general objection to Feinberg's argument is stated in V. Pedersen & S. Midtgaard, 'Is Anti-Paternalism Enough?' *Political Studies*, doi: 10.1177/0032321717734515 (2017).

³⁷ Note that I have already excluded potential spectator problems, which include, for example, concerns such as those Feinberg raises that attending the fight will lead to brutalization among the spectators. Feinberg, *Harmless Wrongdoing*, pp. 131-132.

³⁸ Not that I am convinced that this would in itself justify the state in preventing Rita from harming herself.

preventing people from harming themselves by becoming slaves. An alternative legal moralistic justification for interference is, however, that selling oneself into slavery is in itself an immoral act.³⁹ In parallel with the agreement between the slave and the slave owner, we may prevent the gladiatorial fight with reference to the immoral acts of Robert or Rita.

The moralistic spirit of the arguments is, however, as Feinberg acknowledges, not in sync with the liberal way in which antipaternalists prefer to argue.⁴⁰ The reason for this is, according to Feinberg, that if we test the moralistic argument on the most plausible interpretation of Mill's 'harm principle', the argument does not suffice to justify intervention. Based on Mill's doctrine about liberty, Feinberg argues that limitations of individual freedom should only be justified with reference to harms (or offenses) to others. However, some instances of other-regarding harm are not justified with reference to this principle. Specifically, the principle does not justify limitations of consenting parties' freedom to form voluntary agreements. The reason is, according to Feinberg, that the harm principle is qualified by the 'Volenti maxim', according to which '[t]o what a man consents he may be harmed, but he cannot be wronged; and Mill's "harm principle," reinterpreted accordingly, is designed to protect him and others only from wrongful invasions of their interest'.⁴¹ Even if some immoral actions are accepted as 'genuine evils', the intention to prevent such evils is, as Feinberg puts it, 'rarely if ever enough to offset the

³⁹ This latter legal moralistic argument should not be confused with moral paternalism in that it appeals to 'the mere immorality of the conduct interfered with' rather than 'the harm done to the agent's character'. G. Dworkin, 'Paternalism', *The Stanford Encyclopedia of Philosophy*, E. N. Zalta (ed.), <<https://plato.stanford.edu/archives/win2017/entries/paternalism/>> (2017).

⁴⁰ Feinberg, 'Legal Paternalism', p. 188; Feinberg, *Harm to Self*, p. 81.

⁴¹ Feinberg, 'Legal Paternalism', p. 108.

presumptive case for liberty'.⁴² For this reason, antipaternalists should plausibly oppose moralistically justified interventions.⁴³

Feinberg considers a fourth non-paternalistic rationale that also rests on the problematic role of the slave owner, namely what he calls the Kantian argument.⁴⁴ It is an argument from the idea that people have certain inalienable rights; they are not allowed irrevocably to give up or trade away their fundamental freedoms. Relational egalitarians have recently appealed to this argument. As Elizabeth Anderson points out, denying individuals such freedoms may initially appear paternalistic. The relational egalitarian justification for assigning the inalienable rights in question, however, is not based on or related to the interests or good of the rights bearers, but on the unconditional duties of others. The rationale is, as Anderson explains it, that 'every individual has a worth or dignity that is not conditional upon anyone's desires or preferences, not even the individual's own desires. This implies that there are some things one may never do to other people, such as to enslave them, even if one has their permission or consent.'⁴⁵ In the case of Rita and Robert, the gladiatorial combat may be prevented with reference to some corresponding unconditional duties of Robert's, whose participation in the unconscionable agreement may be seen as disrespectful to Rita and conflicting with how she should be treated in view of her equal moral worth. Although Rita consents to the fight, it is impermissible for Robert to perform what amounts to an easy kill of Rita. Robert violates Rita's non-waivable rights by doing so.

⁴² Feinberg, *Harmless Wrongdoing*, p. 5.

⁴³ Feinberg, *Harmless Wrongdoing*, pp. 6, 130-131.

⁴⁴ Feinberg, 'Legal Paternalism', p. 118.

⁴⁵ Anderson, 'What is the Point', p. 319.

It is, however, uncertain whether the Kantian argument is sufficient with respect to justifying restrictions on voluntary agreements on non-paternalistic grounds for two reasons. First, according to Feinberg, the Kantian argument suffers from the same shortcoming as the moralistic argument.⁴⁶ The reason is outlined above: when people consent to harm, they are not wronged. Accordingly, the argument is uncongenial to Mill's harm principle, and the need to prevent the 'evil' in question is an inferior and plausibly insufficient reason to justify limitations of voluntary choices. Second, as argued by Søren Flinch Midtgaard, what you can ask people to give up without referring to their or other's interests or well-being might be limited. As Midtgaard clarifies, the relational egalitarian justification is 'non-interest-based'. It is 'grounded not in a concern for the well-being or interests of the rights holders (i.e. those who hold certain non-renounceable rights), but in an appropriate respect for the persons holding the rights, or for their worth'.⁴⁷ According to Midtgaard, it is questionable whether the intrinsic value appealed to in the relational egalitarian justification actually warrants the relevant curtailments of liberty.

However, it might be possible to adjust the relational egalitarian argument in a way that makes it interest-based *and* nonpaternalistic. Antipaternalists may, for example, argue that while we should not prevent Rita from harming herself, it would still be appropriate and justifiable to prevent Robert from harming Rita with reference to Rita's interests or well-being. No matter what, I find it possible to reconstruct the case of Rita and Robert in a way that excludes the possibility of appealing to the outlined arguments relying on Robert's problematic acts or role.

⁴⁶ Feinberg, 'Legal Paternalism', p. 118.

⁴⁷ S. F. Midtgaard, 'Non-Renounceable Rights, Paternalism, and Autonomy', *Utilitas* 27 (2015), pp. 347-364, at p. 362.

Consider the following case of Rita and Robot Robert, which is very similar to the case of Rita and Robert. The only difference between the two cases is that in the case of Rita and Robot Robert the professional, well-armed gladiator is a robot that Rita has built and programmed herself.⁴⁸ Imagine that technological development makes it possible to produce humanlike robots that look and behave completely like human beings. The risk that Rita is running in this case corresponds to the risk in the former case. Again, if nobody intervenes, Rita will take part in the gladiatorial combat against Robot Robert and thus assume a very high risk of incurring massive irreversible harm, if not losing her life. Here, the Kantian-inspired argument is excluded as the robot cannot be said to act immorally or disrespectfully towards Rita. The robot is a means by which Rita is harming herself. Accordingly, the combat between Rita and Robot Robert is a case of purely self-regarding harm. The intuition that Rita should be barred from fighting is persistent, though, and just as strong as in the case of Rita and Robert. This would seem to suggest that a salient intuition across the two cases concerns the importance of preventing grave self-regarding harm.

One question that remains to be answered is whether we should not endure this case of a person voluntarily consenting to gladiatorial combat. Here, it is of course possible (and consistent with antipaternalism) to take the position that, in the absence of nonpaternalistic reasons to which we can appeal, it is reasonable to let people expose themselves to great danger as it is always more important to respect people's voluntary choices than to prevent them from potentially destroying

⁴⁸ Furthermore, Rita has bought the parts for the robot herself. It is not the case that someone else owns the robot and lends it to Rita for the fight.

their future interests.⁴⁹ As I mentioned in the introduction, antipaternalists may choose to bite the bullet and accept the fight in the absence of nonpaternalistic reasons to appeal to.

Nevertheless, in my view, this persistent approach seems to conflict with common sense in two ways. First, to prevent Rita from incurring massive harm would be to prevent an outcome that is uncontroversially bad. Insisting that it is, in principle, always wrong to interfere with voluntary choices to prevent outcomes even of this kind seems to involve an intuitively implausible weighing of important considerations. That said, I am aware that our intuitions regarding the same policies and actions might differ. For the same reason, it is difficult to draw decisive conclusions from arguments that appeal to the intuitive plausibility of policies and actions. Although I acknowledge this, I question the idea that ignoring Rita's self-regarding harm would not be an embarrassment at least to those antipaternalists who endorse the unconscionability doctrine. If I am right, the case of Rita and Robot Robert represents a hypothetical counterexample to defenders of antipaternalism who must refrain from interference in a case where this would be, intuitively, plausible.

VI. SUFFICIENT VS. SATISFACTORY REASONS

In Section IV, I conceded that antipaternalists are not committed to rejecting the unconscionability doctrine. In light of this, one may ask why it matters that they do not, as I have just shown, seem to provide a satisfactory justification. In other words, why is it important whether the justification provided actually reflects our crucial concerns attached to the relevant situations, if the nonpaternalistic reason is sufficient to justify the unconscionability doctrine? Whether or not we accept paternalistic reasons, this may seem to lead to the same practice.

⁴⁹ See also Dworkin, 'Some Second Thoughts', p. 126.

However, as in most other aspects of political philosophy, we should not only care about practice.⁵⁰ In the previous section, I hope to have substantiated that we ought to pay attention to paternalistic concerns as they arise with respect to the interests and well-being of potential promisors when we evaluate policies, including the unconscionability doctrine. If I am right that the harm to potential promisors is a proper and independent consideration that not only does but also should worry us in cases of unconscionable agreements, then it is implausible to exclude this consideration from the justification of the unconscionability doctrine. My argument for the unsatisfactory nature of the nonpaternalistic reasons arguably reveals some deficiencies of antipaternalism, at least on a theoretical level. And the plausibility of the antipaternalist project of reconciliation relies on whether a general principle of antipaternalism is persuasive. For the project of reconciliation to make sense, we should arguably not want the policies whether or not reconciliation theorists identify sufficient nonpaternalistic reasons in their favor.

Nonpaternalistic reasons, including Shiffrin's, may, however, reflect some of the worries we have, leading us to embrace the unconscionability doctrine. In line with Shiffrin's argument, the apparently unfair and exploitative nature of the contracts in question may indeed mean that assisting in upholding them would involve a setback to those upholding the contracts. Accordingly, the alternative paternalistic reason that I have emphasized here may not be exhaustive of the concerns grounding the doctrine. If this is true, it may seem as if my argument is subject to an objection similar in nature to the one I have deployed against antipaternalism. It would seem to be so in virtue of the paternalistic argument's incomplete or lacking nature just pointed to.

⁵⁰ For defenders of the view that moral rights and principles should be tested on a higher level of abstraction, see e.g. Arneson, 'Joel Feinberg', pp. 273-274; G. A. Cohen, *Rescuing Justice and Equality* (Cambridge, 2008), p. 268.

However, appearance is misleading here. While antipaternalists reject paternalistic reasons (or reject these as ever being decisive), no reasonable paternalist would rule out the very likely existence of relevant nonpaternalistic reasons. Therefore, if we should not only worry about the interests and well-being of potential promisors, but also have other, nonpaternalistic, considerations, then this is not a problem for paternalists. Paternalists can accept that such considerations are also relevant. In fact, even if the unconscionability doctrine might to some extent in certain cases counteract the paternalist aim of benefitting the person interfered with, paternalists are not committed to abstaining from interference or declining to sustain contracts in such situations. As Feinberg argues, it is completely consistent with paternalism to accept that, sometimes, nonpaternalistic reasons should make us override or set aside paternalistic considerations.⁵¹ For these reasons, paternalists find themselves better equipped to provide satisfactory reasons for policies protecting people against self-regarding harm even if the paternalistic reason in itself only reflects a substantial part of our proper concerns pertaining to such policies.⁵²

⁵¹ Feinberg, *Harm to Self*, p. 25.

⁵² It may be argued that this correction or clarification is redundant in that no one denies that paternalists can accept nonpaternalistic reasons. There is, however, a tendency to assume that paternalists must endorse policies that benefit people. For example, according to Shiffrin, '[t]hose who endorse paternalism would have to defend a broader range of interferences and omissions—including cases like Intervention [where an uninvolved party intervenes to prevent an agreement] and Fines [where the uninvolved party imposes fines or in other ways complicates the implementation of the contract, again because he believes it harms one of the parties]'. See Shiffrin, 'Paternalism, Unconscionability Doctrine, and Accommodation', pp. 231, 225. While I agree with Shiffrin that most paternalists plausibly would endorse the restrictions in question, I disagree that they would necessarily *have to*.

VII. CONCLUSION

In this paper, I have cast serious doubt on the procedure of the project of reconciling apparently paternalistic and seemingly reasonable policies with antipaternalism. Reconciliation theorists may often fulfill the project's internal aim of identifying reasons that suffice to justify the relevant policies and actions on nonpaternalist grounds. Seana Shiffrin's nonpaternalistic argument for the unconscionability doctrine seems to accommodate this purpose. The objection I have raised in this paper is, however, that reconciliation theorists should be more ambitious with respect to what their nonpaternalistic reasons should live up to. In this latter regard, I have argued that Shiffrin's nonpaternalistic defense of the unconscionability doctrine fails in that her argument does not adequately reflect our ultimate concerns pertaining to the doctrine. When the reasons identified by reconciliation theorists are unsatisfactory in this way, we should arguably not let our endorsement or rejection of seemingly reasonable policies and actions rely on them. Sometimes the policies sought to be justified nonpaternalistically are plausible partly because of the paternalistic reasons in their favor. In such instances, we can construct hypothetical cases in which antipaternalists must jettison the policies in question even though this would be unreasonable. The last resort for antipaternalists, namely abandoning what appears to be a reasonable policy, would, at least in the dialectic context of this paper, seem to constitute an embarrassment to reconciliation theorists, and to raise doubts about the plausibility of the antipaternalist principle. The upshot is that, when confronted with apparently paternalistic and sound policies, neither of the two antipaternalist options outlined in the beginning of this paper offer much hope.⁵³

⁵³ **Acknowledgements** For helpful comments to various versions of this article, I am grateful to: the members of the Political Theory Section, Department of Political Science, University of Aarhus; participants in the 'Europe on Distributive Justice' conference, 14-16 June 2017, at the same institution; participants in the 'Festival of Ideas 2'

vikip@ps.au.dk

workshop, 30-31 March 2017, Law & Philosophy, Universitat Pompeu Fabra, in Barcelona. I am especially thankful to Søren Flinch Midtgaard, Kasper Lippert-Rasmussen, Serena Olsaretti, Andrew Williams and Jens Damgaard Thaysen for detailed written comments and helpful conversations. I would also like to thank the editor and two anonymous referees for their insightful and constructive reviews.