Wage Exploitation and the Nonworseness Claim: Allowing the Wrong, To Do More Good

David Faraci
Durham University

ABSTRACT: Many believe that employment can be wrongfully exploitative, even if it is consensual and mutually beneficial. At the same time, it may seem third parties should not do anything to preclude or eliminate such arrangements, given these same considerations of consent and benefit. I argue that there are perfectly sensible, intuitive ethical positions that vindicate this “Reasonable View.” The view requires such defense because the literature often suggests that there is no theoretical space for it. I respond to arguments for the clearest symptom of this obscuration: the so-called nonworseness claim that a consensual, mutually beneficial transaction cannot be “morally worse” than its absence. In addition to making space for the Reasonable View, this serves my dialectical goal of encouraging distinct attention to first- and third-party obligations.

KEY WORDS: wage exploitation, nonworseness claim, sweatshops, paradox of deontology, consequentialization

Owner offers full-time employment to Employee, for which the latter will receive compensation insufficient to meet his basic needs (food, shelter, etc.). Owner could offer Employee better compensation without lowering the net benefit of their interaction or eliminating Owner’s profits. Indeed, Employee’s compensation represents only a small fraction of the profits his work generates for Owner. Nevertheless, Employee, who has no feasible alternative for even approaching his basic needs, correctly concludes that the position on offer is his best option—including over any options he would have if Owner’s business did not exist. Employee, therefore, consents to work for Owner. No third parties are harmed by this interaction.

I find the following view about this case intuitively plausible. This is a case of wage exploitation: Owner is taking advantage of Employee in paying him unfairly low wages.1 Owner acts wrongly in doing this, even though she had no prior obligation to benefit Employee; both Owner and Employee consent to the interaction, it makes them both better off, and there are no negative externalities.2 Nevertheless, third parties should not do anything that would lead Owner to rescind the offer of employment, given these same considerations of consent and mutual benefit. Call this thought—that an interaction like Owner and Employee’s can be wrongfully exploitative, yet third parties should not do anything (expected) to prevent or eliminate it—the “Reasonable View.”
In this article, I argue that the Reasonable View is indeed reasonable, in the sense that there are perfectly consistent, intuitive ethical positions that vindicate it. The Reasonable View requires such defense because, as we’ll see, the current literature on wage exploitation often suggests, implicitly or explicitly, that there is no theoretical space for it. This is a mistake.

I proceed as follows. In section 1, I discuss how the contemporary literature (focusing primarily on the sweatshop literature) obscures the theoretical space for the Reasonable View. I focus on arguably the clearest symptom of this obscuration: the apparent plausibility of the nonworseness claim, according to which a consensual, mutually beneficial transaction cannot be “morally worse” than its absence (Wertheimer 1999, 289–93, though note that Wertheimer himself rejects the claim). Proponents do not defend the nonworseness claim as an upshot of a particular moral theory, such as the view that welfare and autonomy are all that matters morally. Rather, it is meant to play a central role in a purportedly theory-neutral defense of wage exploitation, and thus a rejection of the Reasonable View.

In section 2, I explore the intuitive case for the nonworseness claim and argue that it relies on three errors. It fails to appropriately distinguish the evaluative from the deontic. It fails to adequately attend to the potential moral significance of the fact that different options are available to different agents. And it conflates value monism with value pluralism. Once these errors are exposed, it becomes clear that a perfectly sensible form of deontology is consistent with the Reasonable View.

Given my arguments, it can be tempting to conclude that only deontologists can embrace the Reasonable View. Indeed, it has been argued that the Reasonable View should be rejected because it faces “the paradox of deontology” (Ferguson 2016). I include a principled response to this objection in section 2. In section 3, I develop an example of a consequentialist theory consistent with the Reasonable View. This serves as a further response to the “paradox of deontology” objection, further illuminates the structure of the Reasonable View, and demonstrates that even consequentialists can endorse it.

In section 4, I consider the dialectical, theoretical, and practical implications of my arguments. My discussion shows that there is no theory-neutral argument for or against the Reasonable View. Whether interactions like Owner and Employee’s involve wrongdoing depends on both substantive ethical questions and empirical ones. It also serves my primary dialectical goal, which is to encourage everyone to more carefully distinguish their treatment of first- and third-party obligations. This is the point of the article’s subtitle: we must attend to the possibility that we should sometimes allow wrongdoing, in order to do more good.

1. WHATEVER HAPPENED TO THE REASONABLE VIEW?

Until about twenty years ago, moral condemnation of sweatshops was nearly universal in the United States. As Powell and Zwolinski point out at the start of their (2012) defense of sweatshops, Ian Maitland could still comfortably title his (1996) article “The Great Non-Debate Over International Sweatshops.” Since then, numerous economists and philosophers have offered defenses of sweatshops in both the

---

1. WHATEVER HAPPENED TO THE REASONABLE VIEW?

Until about twenty years ago, moral condemnation of sweatshops was nearly universal in the United States. As Powell and Zwolinski point out at the start of their (2012) defense of sweatshops, Ian Maitland could still comfortably title his (1996) article “The Great Non-Debate Over International Sweatshops.” Since then, numerous economists and philosophers have offered defenses of sweatshops in both the
academic literature and popular media, largely focusing on the idea that sweatshop critics fail to understand the economics behind sweatshops (e.g., Kristof 2009; Kristof and WuDunn 2000; Krugman 1997; Maitland 1996; Powell 2014; Zwolinski 2007; for economic critiques of sweatshops see, e.g., Arnold 2010; Arnold and Hartman 2005; Coakley and Kates 2013; Kates 2015; Miller 2003; Pollin, Burns, and Heintz 2004).

In ethics, economic arguments can only take us so far. For example, they can give us reason to believe that sweatshop workers are benefitted (at least, financially) by the presence of sweatshops. They cannot show us that such benefits are all that matter morally, or even matter enough to outweigh other moral considerations, such as fairness. But suppose sweatshop work is both consensual and beneficial for sweatshop workers. And suppose that any steps third parties might take to improve sweatshop conditions would lead sweatshops to close. If moral objections to sweatshops are supposed to stem from concern for sweatshop workers, trying to improve their conditions starts to look morally counterproductive. If third parties care about people like Employee, perhaps they would do better allowing such workers to be exploited.

This line of thought echoes Matt Zwolinski’s (2007) seminal defense of sweatshops. There, Zwolinski focuses explicitly on third-party obligations:

My position is that there is a large gulf between concluding that the activities of sweatshops are morally evil and concluding that sweatshop labor ought to be legally prohibited, boycotted, regulated, or prohibited by moral norms. To the extent that sweatshops do evil to their workers, they do so in the context of providing their workers with a financial benefit, and workers’ eager readiness to consent to the conditions of sweatshop labor shows that they view this benefit as considerable. This fact leads to the ultimate practical conclusion of this paper, which is that there is a strong moral reason for third parties such as consumers and host and home country governments to refrain from acting in ways which are likely to deprive sweatshop workers of their jobs (Zwolinski 2007, 690).

Zwolinski’s argument against third-party interference may or may not be sound. Perhaps sweatshop workers are coerced into working in sweatshops or aren’t really benefitted (e.g., Arnold 2001; Kates 2015; Miller 2003). But even if his argument is sound, on its face it seems consistent with the claim that sweatshop owners wrongfully exploit their workers. This would be in keeping with the Reasonable View. If any relevant third-party intervention would lead Owner to rescind the offer—which, for the sake of argument, I assume to be the case throughout—perhaps third parties should refrain from interfering. What Owner does can still be wrong.

Yet, strikingly, as time as passed, the literature has increasingly ignored the Reasonable View, or even suggested that there is no theoretical space for it. The progression of Zwolinski’s work provides a perfect example. In the article just discussed, Zwolinski (following Wertheimer) argues along the following lines. Most of us interact with people like Employee in ways that are not beneficial (e.g., don’t interact with them at all). Owner, by contrast, benefits Employee consensually. So, if Owner is doing something wrong, this must be in virtue of the truth of some interaction principle on which “one has special responsibilities to those with whom
one interacts beneficially that one would not have if one had chosen not to interact with them” (Wertheimer 2006). If no interaction principle is defensible, it seems we should accept the nonworseness claim (again, that a consensual, mutually beneficial transaction cannot be morally worse than its absence). Zwolinski raises objections to some particular interaction principles, but is ultimately concessive:

There are no doubt other arguments that could be made in defense of the interaction principle. All I have tried to do here is to show that some of the more obvious ones are not successful. The burden of argument is thus on those who wish to criticize sweatshop wage agreements to provide a coherent defense of the interaction principle, and thereby show how sweatshops’ marginal benefit to the poor of the developing world is worse than the complete lack of benefit that most of us provide (Zwolinski 2007, 710).

Just one year later in an article on price gouging, however, Zwolinski claims that on “a consequentialist moral framework, the nonworseness claim seems obviously true.” He then offers a general argument against the plausibility of interaction principles, maintaining that the nonworseness claim “is meant to have traction for deontological theories as well” (Zwolinski 2008, 357).

Theories which subject A to a moral obligation to rescue B on condition of their interaction seem to suffer from both a defect of unfairness and a serious internal tension. A’s interaction with B is supererogatory, done with B’s consent, violates no independent moral constraints, and benefits B. Why, then, should the interaction itself place A under new moral obligations toward B, beyond those to which A and B mutually agree? On the face of it, it seems unfair to burden A with this extra requirement given that he is already doing more than is morally required of him (Zwolinski 2008, 359).

By the time we get to Zwolinski’s (2012) article with Benjamin Powell, we find him arguing that all “anti-sweatshop arguments” fail on theory-neutral grounds:

The argument in this article . . . does not seek to refute the case against sweatshops from the perspective of a single narrow and controversial moral theory. Instead, it seeks to show that anti-sweatshop arguments fail in one of two ways: Either they fail internally, by running afoul of the moral criteria to which they themselves proclaim allegiance, or they fail in a way that is external but uncontroversial, by succumbing to objections that any reasonable moral theory ought to view as legitimate concerns (Powell and Zwolinski 2012, 450).

And while he and Powell do not reject interaction principles out of hand here, they take them to be largely irrelevant, on the grounds that sweatshop workers can simply be taken to have waived any relevant obligations:

We do not deny that entering into a relationship can create new obligations. We simply hold that it is implausible to hold that those new obligations are not waivable, even when one party regards the other’s waiving of the obligation to be a necessary precondition for entering into the relationship, and the other party strongly prefers the relationship without the obligation to no relationship at all (Powell and Zwolinski 2012, 469).
Importantly, this trend exists not only amongst sweatshop defenders; theorists on both sides have afforded insufficient attention to the theoretical space for the Reasonable View. Consider, for example, Jeremy Snyder’s (2010) overview of the recent literature on sweatshops and exploitation. Snyder discusses numerous accounts meant to ground distinctive obligations for sweatshop owners—accounts stemming from concerns about local unfairness (Meyers 2004, 2007; Valdman 2008, 2009; Wertheimer 1999); about obligations regarding structural injustice (Sample 2003; Young 2004); about obligations of corporate citizenship (Crane and Matten 2008; Néron and Norman 2008); about appropriate responses to globalization (Scherer and Palazzo 2007; Scherer, Palazzo, and Baumann 2006); and about the Kantian injunction against using others as mere means (Arnold and Bowie 2003; Bowie 1999; Popper 2006; Radin and Calkins 2006; Snyder 2008). Any one of these might be thought to wholly undermine the nonworseness claim. Yet while Snyder does question it, he doesn’t end up very far away himself, concluding that “a relationship is more likely to be exploitative but morally permissible when it is voluntary and mutually beneficial . . .” (Snyder 2010, 208). He does so, much like the later Zwolinski, because he runs together questions about Owner’s obligations with questions about third parties’.

[A] relationship might be said to be exploitative but morally permissible. This moral permissibility is typically extended in light of sufficiently great benefits to the exploitee or sufficient hardship if the relationship were prevented. The question of moral permissibility may be limited to whether third parties, such as state regulators, are justified in interfering with an exploitative relationship (Snyder 2010, 208).

Thus, we find theorists on both sides of the sweatshop debate ignoring the Reasonable View, if not ruling it out as indefensible on supposedly theory-neutral grounds, by way of the nonworseness claim. In the next section, I argue that defenses of the nonworseness claim, and the attendant failure to recognize the theoretical space available for the Reasonable View, rest on three related theoretical errors: failure to appropriately distinguish the evaluative from the deontic; failure to adequately account for the potential moral significance of different options’ being available to different agents; and conflation of value monism with value pluralism.

2. THE NONWORSENESS CLAIM AND THE REASONABLENESS OF THE REASONABLE VIEW

Here is Wertheimer’s original statement of the nonworseness claim:

Suppose that we criticize A for engaging in an unfair mutually advantageous transaction with B. A defends himself by appealing to what I shall call the nonworseness claim: “Look, you grant that I have a right not to transact with B. If B is better off if I engage in an unfair transaction with B than if I do not transact with B at all, it certainly can’t be morally worse for me to engage in such a transaction with B. Given that I have a right not to transact with B and that transacting with B is not worse than not transacting with B, it can’t be seriously wrong for me to engage in an unfair transaction with B” (Wertheimer 1999, 289).
Following this line of thought, the champion of the Reasonable View charges that Owner’s hiring of Employee is wrongful because it is unfair. Owner replies that since Employee is better off, and Owner didn’t have to benefit Employee at all, her benefitting him can’t be wrongful.

On its face, this is a rather strange reply. The obvious rejoinder is that the presence of a benefit isn’t what’s at issue; unfairness is. Zwolinski’s contention that the nonworseness claim seems obviously true for consequentialists further highlights this oddness. After all, Wertheimer just told us what moral value Owner’s exploiting Employee might negatively impact: fairness. Nothing prevents consequentialists from acknowledging the value of fairness. And so nothing prevents consequentialists from claiming that even though “transacting with B is not worse for B than not transacting with B at all,” it is nevertheless worse all things considered because of its unfairness (Zwolinski 2008, 357, bolding mine).

As an analogy, suppose someone were to propose a variant of the nonworseness claim that lacks the consent condition—i.e., claim that a morally optional interaction cannot be wrongful just so long as all parties are made better off, even if this is achieved through coercion. It would be baldly question-begging for someone to invoke this variant in response to a complaint that some employees don’t consent to their employment. And it would be quite odd to claim that this variant is obviously true for consequentialists. It is only obviously true for consequentialists who deny the value of anything other than welfare—i.e., for welfarists. Likewise, the real nonworseness claim is only obviously true for consequentialists who deny the value of anything other than welfare and autonomy (or whatever undergirds the importance of consent). And there are plenty outside of that camp.

If the nonworseness claim isn’t question-begging, something else must be going on here. I submit the following explanation. Both Wertheimer and Zwolinski are banking on our sharing a certain evaluative intuition: welfare is morally important; indeed, it is important enough that gains in welfare matter more than losses in fairness. Thus, the state of affairs in which Employee is unfairly benefitted seems more good than the state in which he is fairly left in poverty.

Now suppose you are a consequentialist. For consequentialists, more goodness means more rightness. This makes it natural to infer a first-party deontic claim from the evaluative one (though I’ll challenge this in section 3): Owner’s treating Employee unfairly can’t be more wrong than leaving Employee in poverty, given that the result is more valuable. This provides a better reading of Zwolinski: he doesn’t really mean that consequentialists have to endorse the nonworseness claim; he means that any consequentialist who accepts the above evaluative claim also has to accept the deontic one. And again he (not unreasonably, I think) expects most people to embrace the former.

This also provides a clearer reading of the Wertheimer passage. A’s defense is that given that B consents and is made better off, the results of their morally optional interaction are more valuable than the results of its absence, regardless of how fair it is. This is why third parties shouldn’t do anything to prevent or eliminate that interaction. A expects everyone to accept this premise. A then infers the relevant deontic claim: “it can’t be seriously wrong for me to engage in an unfair transaction with B.”
It is worth pausing to notice how ambiguous language contributes to confusion here. Both Wertheimer and Zwolinski talk about what is “morally worse,” which is ambiguous between the evaluative and the deontic (among other things). This causes particular trouble in Zwolinski’s claim, quoted in section 1, that champions of the Reasonable View need to “show how sweatshops’ marginal benefit to the poor of the developing world is worse than the complete lack of benefit that most of us provide.” This is only true if read as a deontic claim: what sweatshop managers do must be more wrong than what most of us do. But Zwolinski’s phrasing invites an evaluative reading—suggesting, mistakenly, that champions of the Reasonable View need to show the results of sweatshop labor to be less good than the results of third-party inaction.

If all parties to this debate were consequentialists, this failure to distinguish the evaluative from the deontic might not seem so striking or problematic. But many are not; indeed, Snyder’s (2008) views are explicitly Kantian. Yet as we saw, Snyder doesn’t leave much room for the Reasonable View either. This, I believe, is because of a related theoretical misstep.

To see it, consider a new kind of objection to the Reasonable View. Benjamin Ferguson argues that the Reasonable View should be rejected because it faces the “paradox of deontology.” The Reasonable View suggests that one is sometimes required to bring about less of what’s good. This is paradoxical, Ferguson charges, for “[p]rinciples that prohibit what they are designed to promote suffer from a form of inconsistency” (Ferguson 2016, 964).10

Ferguson’s argument can be dismissed fairly quickly: it is question-begging. He seems absolutely right that the correct deontic principles should reflect what matters morally. But he simply assumes that features of states of affairs exhaust what matters morally. And that is just to assume that consequentialism is true. Suppose, instead, that acting fairly itself matters morally. It may be that this cannot be, or is not best,11 understood as a reflection of the value of some feature of states of affairs—e.g., the idea is not that one should maximize fair actions. That does not reveal some deep problem with theories that require acting fairly even when this would produce a worse state of affairs; it reveals only that they are not forms of consequentialism.12

Nevertheless, it is worth considering what makes Ferguson’s objection tempting. For while Ferguson is not the first to claim that deontology is paradoxical, this is far from the most common reaction to deontological claims on which more rightness goes hand-in-hand with less goodness. Consider Thomson’s (1985) classic case of the surgeon who considers sacrificing one patient to save five others, often deployed as a counterexample to consequentialism. The deontologist’s position is that killing the one to save the five is impermissible, despite the fact that it would generate an all-things-considered more valuable outcome than (permissibly) allowing the five to die.

The typical consequentialist response is not to call deontology paradoxical, but rather to resolve the apparent tension between the evaluative and the deontic, to either bite the bullet and insist that killing the one is right, or accept that killing the one is wrong but insist that the outcome of doing so wouldn’t be better after all (e.g., by arguing that doctors’ doing such things would discourage people from seeking medical attention).
I submit that part of the explanation for this difference in reaction is that most judge not only that it would be wrong for the surgeon to kill the one to save the five, but also that third parties should prevent the surgeon from doing so. By contrast, the intuition that third parties should allow Owner to exploit Employee often survives reflection, alongside the intuition that Owner is doing something wrong. I believe this apparent tension between first- and third-party obligations has exacerbated the appearance of tension between the Reasonable View’s evaluative and deontic commitments. This may make it more tempting to raise principled objections like Ferguson’s, rather than to attempt to alleviate the tension, as most do in the surgeon case.

In addition to the quotation from Snyder in section 1, more ambiguous (and in some cases downright bizarre) language in the literature provides dialectical evidence for this explanation. Wertheimer raises the question of whether actions like Owner’s could be “seriously wrong” without telling us what distinguishes wrongness from “serious” wrongness. The implicit question seems to be: if third parties shouldn’t do anything to stop it, just how wrong could it be? In the same vein, Robert Mayer goes so far as to claim that sweatshop owners might “be right to do wrong” (Mayer 2007, 614).

But the apparent tension between first- and third-party obligations here is illusory. The asymmetry can be readily explained by the fact that different options are available to different agents, and the relationship between this fact and certain deontological constraints. Consider a deontological theory, comprising the following claims:

D1 It is supererogatory to promote welfare.
D2 It is impermissible to distribute gains in welfare unfairly, unless fair distribution is impossible or otherwise impermissible.
D3 It is impermissible to force someone to do something she is morally permitted not to do.

By D1, Owner is not obligated to hire Employee. But since Owner could distribute the benefits of her interaction with Employee fairly, by D2 her failure to do so is impermissible. If Owner hires Employee, she is required to compensate Employee fairly (this is our interaction principle). Third parties find themselves in an importantly different position. By D3, they cannot force Owner to hire Employee and offer fair compensation. They therefore have no way to permissibly distribute gains in welfare fairly, as we are assuming their only other options are to do nothing, in which case the gains in welfare will be distributed unfairly, or to take actions that would prevent or eliminate Owner’s offer, in which case there will be no gains to distribute. This is a perfectly coherent set of claims, one consistent with the Reasonable View.

The failure to recognize the importance of different options’ being available to different agents is also reflected in Powell and Zwolinski’s claim that sweatshop workers have waived employers’ obligations to them. Much like third parties, Employee has no permissible means of ensuring that he is hired and compensated fairly; it’s one or the other. Given that, of course Employee will consent to being exploited. But, clearly, the fact that Employee agrees to be exploited simply because
Wage Exploitation and the Nonworseness Claim

he has no way to get Owner to give him his due does not mean that Owner’s obligation is waived. If I owe you $100 but offer to pay you only $50, the fact that you agree because you have no way to get me to pay you the total amount does not mean you have waived your right to the other $50, at least not in any sense that implies I don’t still owe you that $50.

Hopefully, the reasonableness of the Reasonable View is becoming more apparent. But to complete its defense, one further issue must be addressed: Zwolinski’s objection to interaction principles discussed in section 1, his contention that “it is puzzling how it could be worse by any moral standard . . . for A to provide some help than it is for him to provide none” (Zwolinski 2008, 357).

To make this concrete, consider an interaction principle that requires you to donate 10 percent of your income to charity, but only if you choose to donate anything at all. This does seem unfair; indeed, it seems downright bizarre. Presumably, something like the moral significance of welfare undergirds the fact (if it is a fact) that donating to charity is supererogatory. But if the moral significance of welfare merely suggests donating, why would it require donating more once you begin? This seems no less bizarre as a deontological side constraint than it would as an implication of a consequentialist calculus.14

That bizarreness, however, is not a feature of interaction principles per se; it is a feature of monistic interaction principles on which a single value requires more of you as you respond properly to it. Crucially, this is not the sort of interaction principle at work here. The intuition behind the claim that Owner acts wrongly is not that Owner has to benefit Employee more because Employee has benefitted him some. Rather, it is that Owner is keeping an unfair portion of the benefits they create together; a greater portion of those profits is rightfully Employee’s. Owner’s hiring Employee generates new obligations not because the moral significance of welfare requires more of her now that she has helped Employee, but because the independent moral significance of fairness has become salient within their interaction.15 Once we understand this, the worry that interaction principles place unfair burdens on the generous vanishes. Owner runs afoul of her interaction obligation because Owner seeks to take for herself too much of what she and Employee create together.

This is the crucial feature of the Reasonable View that Zwolinski’s criticisms miss. It is misleading to say that Owner is “providing help.” If that’s all Owner were doing, then of course it would be better than doing nothing. But Owner’s interaction with Employee is mutually beneficial, and fairness may well have something to say about how those benefits are to be distributed. This is simply a reflection of the fact that some values may generate new obligations within—perhaps only within—interactions.16

It should now be clear that the intuitions used to motivate the nonworseness claim pose no principled threat to the Reasonable View, at least for deontologists. Again, this is not to say that the nonworseness claim is false, but only that it admits of no theory-neutral defense. If it is true, it is true because, say, only welfare and autonomy matter morally. But that is precisely what champions of the Reasonable View can and should deny.
In the next section, I argue that consequentialists, too, can embrace the Reasonable View. In addition to its obvious import for consequentialists, this is useful as a further response to Ferguson’s “paradox of deontology” objection and because, like many “consequentializing” exercises, it serves to further clarify certain structural features of the view in question.

3. REASONABLE CONSEQUENTIALISM

In this section, I develop an example theory, “Reasonable Consequentialism,” that entails the Reasonable View, given two further stipulations about Owner and Employee’s case. The first is that Owner’s exploiting Employee would decrease net fairness; the second is that forcing Owner to hire Employee would decrease net autonomy. Reasonable Consequentialism is consequentialist in that all of its deontic commitments are entailed by its evaluative ones, where value supervenes on sets of states of affairs. It is also, therefore, consistent with the idea that more goodness entails more rightness.

Because the Reasonable View involves supererogation—Owner isn’t required to hire Employee, though it would be good for Owner to do so—Reasonable Consequentialism is a form of satisficing consequentialism (e.g., Slote 1984). For simplicity’s sake, the satisficing level is set at the status quo: actions that generate greater expected value than inaction are supererogatory; actions that generate less expected value than inaction are impermissible.

Now, consider the following states of affairs:

- **EXPLOITED**: The result of Owner’s choosing to exploit Employee.
- **NO HIRE**: The result of Owner’s choosing not to hire Employee.
- **FAIR**: The result of Owner’s choosing to hire Employee and compensate him fairly.
- **FORCED**: The result of Owner’s being forced to hire Employee and compensate him fairly.

The view that third parties produce more good by allowing Owner to exploit Employee suggests that the relevant gain in welfare is more important than the loss in fairness (EXPLOITED > NO HIRE). But if Owner acts wrongly in exploiting Employee, then given consequentialism the result of exploiting Employee must be less good than the result of not hiring Employee (NO HIRE > EXPLOITED). That looks straightforwardly inconsistent. And that’s why the Reasonable View seems to be incompatible with consequentialism.

But consequentialism is more flexible than this suggests. For it is perfectly possible for consequentialism to weight values differently in different contexts. Here, there is at least one relevant shift in context, one already discussed: the states of affairs accessible to third parties are different from those accessible to Owner. Everyone can access EXPLOITED and NO HIRE. Owner can make the relevant choices; third parties have no way to incentivize Owner to willingly offer Employee fair compensation. And only third parties can access FORCED, because...
Owner can’t force herself to do something. So, Owner can access exploited, no hire, and fair, while third parties can access exploited, no hire, and forced. What’s more, there is a principled difference between these sets: only the latter involves value conflict—i.e., lacks a member that weakly dominates with respect to all values.

To see this, start with the states accessible to third parties: exploited, no hire, and forced. By doing nothing, third parties bring about exploited. Third parties can instead bring about no hire by doing something (expected) to incentivize Owner not to hire Employee. Or, by forcing Owner to hire Employee, third parties can bring about forced. As seen in Table 1, there is value conflict within this set: no member has at least as much welfare, fairness, and autonomy as the others. Something must be sacrificed.

Now consider the states accessible to Owner: exploited, no hire, and fair. By doing nothing, Owner remains in no hire. By exploiting Employee, Owner can bring about exploited. Or by choosing to hire Employee and compensate him fairly, Owner can bring about fair. Here, there is no conflict, as seen in Table 2. Owner can promote welfare without sacrificing anything of value, by choosing to hire Employee and compensate him fairly. Fair weakly dominates.

Reasonable Consequentialism achieves this by calculating overall value with respect to sets of states of affairs, holding that (a) autonomy’s value is lexically prior to both welfare’s and fairness’; (b) welfare’s value is lexically prior to fairness’ conditional on the presence of value conflict within the relevant set; and (c) fairness’s value is lexically prior to welfare’s conditional on the absence of value conflict within the relevant set.

On Reasonable Consequentialism, promoting welfare is merely supererogatory. Therefore, (1) it is permissible for Owner not to hire Employee.

If Owner exploits Employee, this would bring about greater welfare, but less fairness, as seen in the left-hand column in Table 2. Since there is no value conflict in

Table 1: Expected Value Gains/Losses for Third Parties*

<table>
<thead>
<tr>
<th></th>
<th>EXPLOITED</th>
<th>NO HIRE</th>
<th>FORCED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Welfare</td>
<td>-</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Fairness</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Autonomy</td>
<td></td>
<td></td>
<td>-</td>
</tr>
</tbody>
</table>

* Shaded cells are the available maximum for a row (value). A shaded column (state of affairs) weakly dominates; here none do.
† Gains/losses indexed to status quo (result of inaction).
the set of states accessible to Owner (FAIR weakly dominates), the loss in fairness outweighs the gain in welfare, and so exploiting Employee brings about less overall value. Therefore, (2) it is impermissible for Owner to exploit Employee.

Third parties’ preventing or eliminating Owner’s exploitative offer would increase fairness but lower welfare, as seen in the center column in Table 1. Since there is value conflict in the set of states accessible to third parties, the gain in welfare outweighs the loss in fairness, and so preventing or eliminating the offer leads to lower expected value. Therefore, (3) it is impermissible for third parties to do anything (expected) to prevent or eliminate Owner’s offer.

Finally, since autonomy always has priority, forcing Owner to hire Employee and compensate fairly would also lead to lower expected value, as per the right-hand column in Table 1. Therefore, (4) it is impermissible for third parties to force Owner to hire Employee.23

Of course, this conditional weighting scheme is ad hoc. For our purposes, though, that’s unobjectionable: its single motivation is exactly what we’d hope; it serves to capture, within a consequentialist framework, an intuition about how the value of fairness constrains us—an intuition naturally associated with deontology. When one must choose between welfare and fairness, one should choose welfare. This is why, where welfare and fairness conflict, the former has priority. But one is required to distribute welfare gains as fairly as one (permissibly) can. This is why, where welfare and fairness do not conflict, it is impermissible to promote the former without maintaining the latter.

Reasonable Consequentialism entails and therefore serves as a possibility proof for consequentialist endorsement of the Reasonable View. It also avoids Ferguson’s complaints. Because the view’s deontic commitments are a direct consequence of its evaluative ones, it never recommends bringing about less of what it deems valuable; more goodness always means more rightness.

4. IMPLICATIONS

Both deontologists and consequentialists can endorse the Reasonable View. They can accept that the consensual, beneficial nature of certain employment relations (among others) should make us reluctant to do anything that would eliminate them. They can, at the same time, accept that the employers in question act wrongly in offering unfair compensation.

---

Table 2: Expected Value Gains/Losses for Owner*

<table>
<thead>
<tr>
<th></th>
<th>Exploited</th>
<th>No Hire†</th>
<th>Fair</th>
</tr>
</thead>
<tbody>
<tr>
<td>Welfare</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fairness</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Autonomy</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Shaded cells are the available maximum for a row (value). A shaded column (state of affairs) weakly dominates; here FAIR does.
† Gains/losses indexed to status quo (result of inaction)
To see the dialectical implications of this conclusion, consider the following progression of argument in the exploitation section of Powell and Zwolinski’s (2012) defense of sweatshops. They begin with intuitive motivation for the possibility of consensual, beneficial, wrongful exploitation: a case in which “A offers to rescue B from drowning by selling B a spot on A’s boat for $10,000” (Powell and Zwolinski 2012, 466). They then discuss potential differences between this case and sweatshops that might vindicate the idea that what A does in the boat case is wrong, without condemning sweatshop employers. For instance, rescuing B seems to impose little cost on A, but if Powell and Zwolinski are right that wage determination is made within a fairly narrow margin, the same may not be true of sweatshop employers. Next, they acknowledge that sweatshop wages may be lower than morally appropriate, but argue that this will often be due to background circumstances largely or entirely outside employers’ control. Thus, the thought goes, sweatshops may be bad, but it does not follow that employers are shirking their moral responsibilities. Finally, they consider Snyder’s (2008, 2010) view that employers owe a duty of beneficence to their employees. They respond by appealing to the intuitions undergirding the nonworseness claim, and respond briefly to Snyder’s objections to that claim.

In my view, the final step takes things in an unfortunate direction. What A does in the boat case does seem wrongfully exploitative, and not merely on the assumption that A violates a duty to rescue. Yet unless the wrongness of what A does is grounded in violation of a duty to rescue (i.e., unless “no interaction” is wrongful), the nonworseness claim suggests that what A does isn’t wrong after all. Even Powell and Zwolinski seem to find this counterintuitive. And if it seems at all tempting, I submit this is yet again because of a failure to distinguish first- from third-party obligations. If B would drown without A’s help, and third parties have no way to rescue B themselves, nor to get A to rescue B without exploitation, surely they should permit A to exploit B, rather than let B drown. We should not make the mistake of concluding from this that A does nothing wrong.

The purpose of Owner and Employee’s case is to provide the same intuitive result, but within the context of an employment relation. Consider the dialectical situation if Powell and Zwolinski were to acknowledge the intuitive possibility of such cases. They might proceed much as they did with the boat case, asking what makes these cases seem wrongful, then considering the extent to which such wrong-making-conditions are met by actual sweatshops.

For example, I stipulated that “Employee’s compensation represents only a small fraction of the profits his work generates for Owner.” This claim was meant to elicit the intuition that this “small fraction” is unfairly small. Of course, sweatshop defenders sometimes suggest that market-clearing wages are fair wages, so they might simply reject this intuition. But this is far from necessary, unless their view is that market-clearing wages are ipso facto fair. A more common view, I suspect, is that market-clearing wages are fair because market forces (tend to) ensure that market-clearing wages are fair—i.e., tend to prevent workers’ wages from representing “only a small fraction of the profits” their work produces—and thus that Owner and Employee’s case is importantly different from actual sweatshops. Indeed, Powell and Zwolinski say things arguably meant to make precisely this point:
The often cited fact that a sweatshop worker who produces, say, a pair of Reebok shoes is paid only one US dollar to make a shoe that sells for around $100 does not mean that Nike is walking away from the exchange with $99 and the worker with only $1. Most of the $100 goes to paying for advertising, retailer markup, raw materials, transportation costs, and so on. The amount that actually accrues to Reebok as profit is generally no greater as a percentage of their investment than the profits in any other competitive industry (Powell and Zwolinski 2012, 466–67).

If accurate and representative, this claim might allow Powell and Zwolinski to accept the wrongful unfairness of what Owner does, but draw an important moral distinction between that case and actual sweatshops. If sweatshops are defensible on these grounds, if they do not meet relevant conditions of unfairness, there is simply no need for such moves like the appeal to the nonworseness claim—especially when such moves threaten intuitions shared by defenders (e.g., about the boat case). We do better to focus our attention on development of relevant theories of fairness and other values, and questions about the extent to which actual employment relations (fail to) meet those theories’ conditions. This is not merely theoretically superior; acknowledging an intuitive wrong-making feature but showing that it does not obtain is surely more rhetorically forceful than simply rejecting your opponents’ intuitions about what makes things wrong, especially when you yourself have similar intuitions about similar cases. It also allows us to more carefully distinguish various available positions—e.g., here, two different views about why market-clearing wages are fair.

Of course, this is far from the end of this debate. One need not accept that market-clearing wages are fair even if they are “equitable.” Perhaps they must also be living wages, or meet some other standard—e.g., a counterfactual market standard, like those championed by Valdman (2009) or Wertheimer (1999). Nor is this the only relevant point of potential disagreement about how realistic Owner and Employee’s case is. For example, some sweatshop critics will likely argue that the stipulation that third parties are limited to inaction or elimination of Owner’s offer is unrealistic, that there are third-party interventions that can improve actual sweatshop workers’ conditions without threatening their employment (e.g., Dawkins 2012; Malmqvist 2013). Perhaps such interventions are morally permissible or even required. Indeed, this might be the case even if we accept that sweatshop employers do nothing wrong. Nothing prevents us from, say, championing redistribution on prioritarian grounds.

In addition to these dialectical points, the arguments herein have important theoretical and practical implications. One of my tasks has been to expose some instances of theoretical oversimplification, which understandably abound in applied ethics. Knock-down theoretical arguments are attractive for their simplicity and elegance, as well as their ability to shield us from the need to do difficult empirical work. But we should not overestimate the prevalence of such “magic bullets.” Moral theory is incredibly flexible, and we should regard with suspicion arguments purporting to show that our intuitions in various cases are “paradoxical” or in “tension” with one another, merely because they do not fit easily into a simple theoretical structure.
Finally, at the practical level, the problems with the nonworseness claim serve as a cautionary tale of a different sort. Many readers have likely had the experience of trying to keep students from moving too quickly from claims about what is morally wrong to claims about what we should act to prevent—often: what should be illegal. In these discussions, we tend to appeal to extra-moral considerations, such as epistemological worries or principles of liberalism. One of my goals in this article has been to reinforce the idea that there can be purely moral reasons to distinguish first-from third-party obligations. Even if a particular relation is wrongfully exploitative, it may be that the rest of us have moral reasons to allow this wrongdoing, because doing so will bring about more of what’s good.

CONCLUSION

Defenders of wage exploitation offer powerful economic and moral arguments that the welfare- and autonomy-promoting power of certain exploitative employment relations should make us reluctant to prevent or eliminate those relations. Some have taken this to further show that such exploitation is not wrongful. I have argued that this does not follow; the Reasonable View is a live option for both deontologists and, more strikingly, consequentialists. In light of this, I submit that debate over wage exploitation should be framed more directly in terms of the conditions that elicit intuitions of unfairness (or other values) and the extent to which those conditions obtain within actual cases of apparent wage exploitation. Whatever we conclude, we should be more open to the theoretical and practical possibility that there are moral reasons to allow wrongdoing. Sometimes, allowing wrongdoing may do the most good.

ACKNOWLEDGEMENTS

Thanks to Michael Kates for work and discussions that inspired this article. Thanks to Tristram McPherson and Daniel Wodak for helpful comments. Also thanks to John Hasnas, Peter Jaworski, Govind Persad, Abe Singer, and other participants in the Georgetown Institute for the Study of Markets and Ethics Faculty Workshop Series. Finally, thanks to editor Denis Arnold and three anonymous reviewers at Business Ethics Quarterly.

NOTES

1. My argument does not depend on the assumption that the wrongness here stems from unfairness; other values might do the same work, several of which will be discussed in section 1. However, the assumption of unfairness is dialectically useful because fairness is something many take to be valuable independently of its impact on autonomy and welfare, the values on which my main targets focus. It is also the value Wertheimer appeals to in his original discussion of the nonworseness claim. (Also see note 5.)

2. If “exploitation” is a thick moral term (with a negative valence), then saying that Owner acts wrongfully is redundant. I will talk as though exploitation is morally neutral simply so that I can refer to this as a case of exploitation without taking sides. This is merely for convenience; I take no official position on whether “exploitation” is morally neutral. For arguments that it is neutral, see Valdman (2009). It is worth mentioning that even some defenders of wage exploitation seem comfortable with such neutrality.
For example, Kristof (2009, A35) writes that “while it shocks Americans to hear it, the central challenge in the poorest countries is not that sweatshops exploit too many people, but that they don’t exploit enough.”

3. For present purposes, “deontology” is just non-consequentialism.

4. Much of my argument can be read as an expansion upon Wertheimer’s own reasons for rejecting the nonworseness claim. More recent critics of the nonworseness claim include Arneson (2013), Bailey (2010), Barnes (2013), Malmqvist (2017), and Preiss (2014).

5. This is another advantage of focusing on fairness (see note 1): most arguments against the unfairness of wage exploitation are not principled, but rather depend on empirical claims about how actual cases (e.g., sweatshops) differ from Owner and Employee’s. I discuss this further in section 4.

6. “[T]o the extent that there is something morally objectionable with [sweatshop conditions] the most natural explanation for this is that these conditions are bad for the persons who suffer them” (Powell and Zwolinski 2012, 451).

7. This is not to be confused with the claim that workers are coerced within their employment relations, rather than coerced to enter into them. See, e.g., Arnold and Bowie (2003); for responses, see Sollars and Engander (2007) and Zwolinski (2007).

8. Snyder does allow that permissible exploitation “may create a form of moral residue, where the exploiter may have a duty to make up for the element of exploitation or to take steps to reduce the structural causes of exploitation” (Snyder 2010, 208).

9. Or at least appears to do so. Given other things he says, it may be possible to read Snyder as using “morally permissible” in a non-standard sense, meaning something like “moral for others to permit.” But even if this is Snyder’s intention, the phrasing is unfortunate given the standard meaning of “morally permissible” and the pattern of conflation recorded here.

10. Ferguson’s solution is to accept a more limited version of the nonworseness claim: “If one or more strictly Pareto improving and consensual transactions are possible, then for at least one of these transactions, transacting is morally better than not transacting” (Ferguson 2016, 965). I think this fails to sufficiently capture the intuitiveness of the evaluative claim that an unfair benefit is more valuable than fair poverty, as well as its plausibility as grounds for accepting that third parties should allow Owner to exploit Employee.

11. This may all be moot if every moral theory can be “consequentialized.” Campbell Brown (2011) argues that they cannot. (He references Dreier [1993], Louise [2004] and Portmore [2007] as theorists who endorse the thesis that they can, “or something close to it” [Brown 2011, 750, fn.1].) However, Brown defines consequentialism in terms of “maximizing the good.” This is too narrow for our purposes, given that all relevant parties accept that Owner is not obligated to hire Employee even if doing so would maximize the good. (The one exception of which I am aware is Bailey [2010], who argues that consequentialists should reject the nonworseness claim because it violates consequentialism’s commitment to maximizing the good; but this suggests to me only that Bailey has an overly narrow conception of consequentialism, since few contemporary consequentialists in- or outside the wage exploitation literature deny the existence of the supererogatory.) In section 3, I take consequentialism to include any theory whose deontic commitments are entailed by its evaluative commitments, where value supervenes on features of (sets of) states of affairs. I offer no argument here that all views can be consequentialized on this understanding of consequentialism. My suspicion is that they can, but that many—including, arguably, the Reasonable View—are not best understood as forms of consequentialism.

12. This is perhaps easiest to see on—but crucially does not demand—a buck-passing view on which facts about value are a function of facts about reasons (e.g., Scanlon 1998), and (say) one always has decisive reason to distribute benefits fairly. Note that this is not just a problem for Ferguson, but for many (perhaps all) appeals to the “paradox of deontology.” For example, Alexander and Moore (2015) write that “[d]eontologists need their own, non-consequentialist model of rationality, one that is a viable alternative to the intuitively plausible, ‘act-to-produce-the-best-consequences’ model of [moral] rationality that motivates consequentialist theories. Until this is done, deontology will always be paradoxical.” But while deontologists do need a model of rationality, it is a mistake to claim that without one, their view is “paradoxical.” That is to question-beggingly judge deontological theories’ coherence against a background assumption that acting morally is about producing good outcomes. This is especially problematic given that deontologists are often working “bottom-up,” starting from substantive intuitions about what morality or rationality requires, ones the presumed consequentialist default can’t (easily) accommodate. All of this is symptomatic of the all-too-common, illegitimate assumption that to say that something matters
Wage Exploitation and the Nonworseness Claim

morally (is valuable) is to say that it merits promotion as a feature of states of affairs. For a relevant, useful distinction between “Moorean” and “Kantian” value concepts see Bradley (2006).

13. Or, perhaps, vice versa. It is unclear to me which of these mistakes is primary, whether people have the evaluative intuition and infer both first- and third-party deontic claims, or whether people have the third-party deontic intuition and infer the evaluative and first-party deontic claims. It may differ from person to person. Perhaps Zwolinski, a consequentialist, (implicitly) reasons the first way, while Snyder, a deontologist, reasons the second.

14. Though for an attempt to make this more palatable under deontology, see Snyder (2008).

15. This is assuming that any prior unfairness placed no obligations on Owner. This, of course, relates to the view that wage exploitation is wrongful because it involves taking advantages of past injustice (e.g., Malmqvist 2013; Sample 2003).

16. “Any transaction or relationship that creates a social surplus gives rise to a new moral feature—fairness or unfairness—that does not arise outside of that transaction or relationship” (Wertheimer 1999, 291).

17. The first stipulation is fairly natural. It could turn out that in exploiting someone, one makes things more fair overall. But given Employee’s poverty, it does not seem much of a stretch to suppose that the local unfairness of Owner’s exploitative offer would lead to a globally less fair outcome. The second stipulation might seem more problematic. But note that any consequentialist who accepts that third parties can’t force Owner to hire Employee and compensate fairly, despite the increase in welfare this would arguably bring—including Zwolinski and other defenders of wage exploitation—will presumably have to vindicate a fairly general prohibition on forcing people to do supererogatory things (or take this prohibition to be merely contingent on the details of employment practices, which seems counterintuitive).

18. Some, of course, endorse narrower definitions of consequentialism. But as mentioned in note 11, it is already doubtful that we can rest easy with a traditional “maximize the good” understanding of consequentialism, since it is a near point of agreement in the relevant dialectic that benefitting Employee is supererogatory. In any case, whether or not Reasonable Consequentialism appeals to consequentialists, it avoids structural objections like Ferguson’s. Thanks to Daniel Nolan for discussion on this point.

19. Neither inaction nor negatively defined actions such as not giving Employee a million dollars count as actions here. If inaction were an action, determining the status quo would require determining what action Owner would have taken if she took any action at all. If not giving Employee a million dollars were an action, Owner might be required to give Employee a million dollars, since not giving Employee a million dollars might well decrease welfare compared to the state of affairs resulting from not not giving Employee a million dollars (depending how wealthy Owner is).

20. For deontic claims, the relevant set contains the states accessible to the agent in question. For comparative evaluative claims, the relevant set contains the states being compared. Thus, for instance, because the set comprising exploited and no hire involves value conflict, Reasonable Consequentialism judges that exploited > no hire with respect to this set, in keeping with the evaluative claim used to motivate the nonworseness claim.

21. Or whatever underpins the impermissibility of forcing Owner to do what she is morally permitted not to do.

22. Holding autonomy fixed, this generates: Given a set of states of affairs (S₁ - Sₙ), Sₓ is all-things-considered better than Sₙ if (i) Sₓ contains more fairness but no less welfare than Sₙ or (ii) Sₓ contains more welfare but no less fairness than Sₙ or (iii) Sₓ contains more welfare but less fairness than Sₙ and there is no Sₙ with no less welfare than Sₓ and no less fairness than Sₓ or (iv) Sₓ contains more fairness but less welfare than Sₙ and there is some Sₓ with no less welfare than Sₓ and no less fairness than Sₓ.

23. Reasonable Consequentialism’s other implications are also in line with the Reasonable View: (5) By allowing Owner to exploit Employee, third parties create more welfare but less fairness, as seen by comparing the left-hand and center columns in Table 1. Given welfare’s priority under value conflict, this generates greater expected value, and is therefore supererogatory. (6) If Owner were to hire Employee and compensate fairly, this would bring about greater welfare and have no negative impact on fairness or autonomy, as per the right-hand column in Table 2. This generates greater expected value, and is therefore supererogatory.

24. Of course, they would still need a response to Snyder, since his view concerns a duty of beneficence, rather than fairness. But this is not hard to come by: they initially appeal directly to what they take to be the intuitive implausibility of “holding that failing to benefit needy workers at all is better than benefiting..."
them at a level which is (significantly) greater than zero but less than the morally required amount—even if workers themselves would strongly prefer and would like to choose the latter over the former,” only afterwards noting that this “draws on many of the same intuitions as those invoked in Zwolinski’s presentation of the ‘non-worseness claim’” (Powell and Zwolinski 2012, 469). They need not accept the general claim to raise these intuitive worries. And, indeed, I think these worries are more forceful against Snyder than against fairness-based objections, as framing things in terms of a duty of beneficence invites Zwolinski’s own framings of what employers do as “helping,” as well as his fairness-based objection to monistic interaction principles.

REFERENCES


Wage Exploitation and the Nonworseness Claim


