

**Replaceable lawyers and guilty defendants**

(forthcoming in *The Journal of Moral Philosophy*; this is the penultimate draft)

*Abstract:* Many criminal lawyers should expect that, were they to not defend a certain client, someone no less capable would do so. It is morally wrong for such attorneys to defend defendants who should be punished. This is true even if we grant that the defendant's right to be defended outweighs any rights that might be infringed by the defense and that the benefits of defending are greater than the harm. Nor does this argument depend on any particular view of punishment. The fact that the attorney expects to be replaced by someone equally capable has an asymmetric effect on the reasons for and against defending. The reasons that justify defending become extremely attenuated by this expectation, no matter what they are, while the reasons against defending are much less affected, no matter what they are.

*Keywords:* criminal defense; punishment; law; legal ethics; reasons; overdetermination; applied ethics

This paper considers whether it is morally permissible for criminal lawyers to capably defend clients whose punishment would be morally justified. By "capably" I mean, "In a way that increase the client's chance of avoiding morally justified punishment from what it would be had the client defended himself." Some capable defenses increase the chance of acquittal, but others only increase the chance of a significantly reduced sentence. The discussion of criminal defense in the philosophical and legal literature largely takes for granted that some amount of capable defense is permissible, and focuses instead on what tactics may be used – discussing, for

example, the use of deceptive evidence, or abuses of witnesses on the witness stand.<sup>1</sup> I will argue, however, that any capable defense of a guilty client, even a defense that uses no questionable tactics, will be impermissible in certain kinds of ordinary cases.

The most obvious ways to show that the defense of the guilty is wrong involve arguing that the guilty have no right to be defended, that defense violates some person's rights and these outweigh the rights of the guilty, or that the harms of defending outweigh the benefits.<sup>2</sup> I will not argue for any of those claims, however. In fact, I will assume for the sake of argument that all are false: I will assume that the guilty do have a right to be defended, that this right is more important than any other right which might be violated by the defense, and that the benefits of defending outweigh the harms. That is, I will assume the moral facts least congenial to my thesis. I will also be agnostic about what justifies punishment, as long as some relatively standard view is correct. My argument will instead focus on a fact under-appreciated in the literature on criminal defense. This is that a significant percentage of lawyers should not expect that the rights which might be protected or infringed by defending a client, or the harms that might be brought about

---

<sup>1</sup> Some advocate something close to maximal zeal in defense, well known examples being Abbe Smith, 'Defending Defending: The Case for Unmitigated Zeal on Behalf of People Who Do Terrible Things', *Hofstra Law Review* 28 (2000) 925-961, Abbe Smith, 'The difference in criminal defense and the difference it makes', *Washington University Journal of Law and Policy* 11 (2003) pp. 83-140, and Monroe H. Freedman, *Lawyers' Ethics in an Adversary System*. (Indianapolis: Bobbs-Merrill 1975). Others argue that certain overzealous or aggressive defensive tactics are wrong, but hold that defenses that do not use these tactics, and which are still capable in my sense, are permissible (for example e.g. William H. Simon, 'The ethics of criminal defense', *Michigan Law Review* 91 (1993) pp. 1703-1728). Even those quite critical of defense in civil cases explicitly exempt criminal defense from these criticisms (e.g. Richard Wasserstrom, 'Lawyers as professionals: some moral issues', *Human Rights* 5 (1975) pp. 1-24, and David Luban, 'Are criminal defenders different?' *Michigan Law Review* 91 (1993) pp. 1729-1766). One possible exception to this generalization is Harry I. Subin, 'The Criminal Lawyer's Different Mission: Reflections on the Right to Present a False Case', *Georgetown Journal of Legal Ethics* 1 (1987) pp. 125-153. Subin seems to argue that capable defense might be impermissible, if the only way to capably defend is to deceive.

<sup>2</sup> See discussions in e.g. Luban, 'Are criminal defenders different?' and Subin, 'The criminal lawyer's different mission.'

or avoided, are any more likely to be protected or infringed, brought about or avoided, if they defend than if they do not. This is mostly because, were they not to defend, someone else would, and there are not good reasons to expect this person to be less capable than they. This completely changes how we weigh the reasons for and against defending. Given that a lawyer expects to be replaced in this way, and that the defendant's punishment would be justified, the reasons to defend are categorically outweighed by the reasons against defending, no matter how strong the reasons to defend would be otherwise.<sup>3</sup>

### **Outlining the argument**

If I am to argue that the reasons to defend are categorically outweighed by the reasons against defending, I need to show that the reasons against defending belong in one category, and those to defend belong in some other(s). But before I can show this, it will be helpful to say something about the sorts of reasons that might plausibly be given for and against defending those who should be punished.

There are a number of moral considerations typically cited in favor of criminal defense. Some have to do with the protection of the specific defendant's rights or interests. For example, it is often thought that people in general have a right to be defended or to have their side of an issue heard.<sup>4</sup> Defendants also seem to have a right to be protected from potential improper prosecutorial tactics or excessive punishment.<sup>5</sup> And some writing on criminal defense have

---

<sup>3</sup> My argument should also apply to cases in which the defendant might not be justifiably punished, as long as it is possible that they will be. To make the argument easier to follow, however, for most of the paper I will focus on cases in which the punishment would be justified. I will explain how the argument can be expanded at the very end of the paper.

<sup>4</sup> Freedman, *Lawyers' Ethics in an Adversary System*, Smith, 'Defending defending,' Simon 'The ethics of criminal defense,' Luban, 'Are criminal defenders different?'

<sup>5</sup> Babcock, 'Defending the guilty,' Wasserstrom, 'Lawyers as professionals,' Simon 'The ethics of criminal defense,' David Luban, 'Lawyers as upholders of human dignity (when they aren't busy assaulting it)', *University of Illinois Law Review* (2005) pp. 815-845, Smith, 'Defending defending.'

pointed out that defendants benefit emotionally or mentally from being defended.<sup>6</sup> One might also give reasons to defend that have to do with larger social issues, which go beyond the effects on the particular defendant. Criminal defense is necessary to the continued functioning of an adversarial justice system.<sup>7</sup> Defending clients from wrongfully gathered evidence might help to discourage police misconduct. And failure to defend guilty clients might harm the lawyer/client relationship, since if lawyers regularly abandoned guilty clients, then innocent clients would have reasons to hide inculcating evidence from their attorneys.<sup>8</sup> It does not matter for my argument what the specific reasons that count in favor of defending are, as long as they arise either from the rights/interests of the specific defendant or from these sorts of larger social concerns. Let's assume for the time being that reasons to defend do arise from just these two sources; at the end of the paper, when I consider objections to my view, I will consider some other possible types of reasons to defend.

There are a number of potential reasons that could count against capably defending. If we put aside those arising from questionable defensive tactics, the remaining reasons have to do with the possibility that the client will avoid some amount of justified punishment.<sup>9</sup> Exactly what these reasons are depends on what it is that justifies punishment. The standard views are that punishment is justified by retribution (considerations of justice or desert), by the need for specific deterrence or incapacitation (to prevent the defendant from commission of future crimes), by the need for general deterrence (deterrence of other possible criminals), by the need for restitution, or

---

<sup>6</sup> Babcock, 'Defending the guilty', Smith, 'The difference in criminal defense and the difference it makes.'

<sup>7</sup> Wasserstrom, 'Lawyers as professionals.'

<sup>8</sup> Smith, 'The difference in criminal defense and the difference it makes', Freedman, *Lawyers' Ethics in an Adversarial System*.

<sup>9</sup> See, e.g. Luban, 'Are criminal defenders different' or Simon, 'The ethics of criminal defense.'

by the need for rehabilitation.<sup>10</sup> Because capable defenses increase the chance that the guilty will escape a significant amount of justified punishment, they increase the risk that the guilty will not pay full restitution or get what they deserve, or that punishment will have a weakened or non-existent rehabilitative, deterrent, or incapacitative effect. My argument does not depend on which view of what justifies punishment is correct, with one exception (which I will say more about later). My argument will not apply when the *only* justification for a specific punishment is that the widespread practice of punishment is a general deterrent; that is, when the specific punishment is not itself a deterrent, but is the carrying out of a deterring threat.<sup>11</sup> This should not be an issue, because significant punishment cannot be justified on such grounds: the defendant is not himself a threat, nor does justice demand his punishment, nor will the punishment of the specific defendant actually have a deterrent effect on anyone.<sup>12</sup> However, if I am mistaken and the deterrent effect of the general practice of punishment can, by itself, justify

---

<sup>10</sup> For example, the Model Penal Code seems to endorse incapacitation, retribution, and deterrence as justifications for punishment, see American Law Institute, *Model Penal Code* (1962); retribution is advocated by Igor Primoratz, *Justifying Legal Punishment* (London: Humanities Press International, 1990); general deterrence is advocated by Norval Morris, *The Future of Imprisonment* (Chicago: University of Chicago Press, 1974); rehabilitation is endorsed by Egardo Rotman, *Beyond Punishment* (Westport: Greenwood Press, 1990); and restitution is advocated by David Boonin, *The Problem of Punishment* (Cambridge University Press, 2008).

Another potential view of the justification of punishment is the “rights forfeiture” view: criminals lose their right to not be punished when they commit crimes. It is not terribly plausible that rights forfeiture by itself justifies punishment; some other justification is needed as well.

Otherwise the view allows that any punishment, no matter how harsh, can be justified as long as the possibility of it occurring is publicized. See, however, Christopher Wellman, ‘The rights forfeiture theory of punishment,’ *Ethics* 122 (2012), pp. 371-393 for a defense of rights forfeiture as, by itself, justifying punishment.

<sup>11</sup> For some defense of a view close to this, see Zachary Hoskins, ‘Deterrent punishment and respect for persons,’ *Ohio State Journal of Criminal Law* 8 (2010), pp. 369-384.

<sup>12</sup> Note that, in “magistrate and mob” type cases, it is clearly wrong to punish a person who does not deserve it even when the actual punishment will definitely deter a huge amount of crime (see Philippa Foot, ‘The problem of abortion and the doctrine of double effect,’ *Oxford Review* 5 (1967), pp. 5-15). It should then clearly be wrong to punish someone who does not deserve it when that actual punishment will do nothing. It may be permissible to punish such a person with fines or tickets (parking tickets are a good example). But general deterrence does not seem to justify the sorts of punishments that defense attorneys defend one from.

specific acts of punishment, my thesis should be slightly restricted: it would then say that, when a lawyer expects her choice to defend to not make a difference, it will be wrong to defend a client whose punishment would be justified unless the punishment is justified solely by the deterrence effect of the practice of punishment.

Why is it irrelevant to my argument which of these putative reasons in favor of or against defending are actual reasons? This is because, when an attorney expects her choice to defend or not defend a client to make no difference (in the way I will soon specify more precisely), the reasons against defending categorically outweigh the reasons in favor of defending. This is true no matter what those reasons are, or how strong they would be if the attorney did expect to make a difference. That is, we can assume that any of the reasons to defend would be much more morally important than all of the reasons against defending if the defense attorney expected to make a difference; even so, it will be wrong to defend clients who should be punished when the defense is expected to make no difference. Before we see why this is, let's first discuss how, why, and when an attorney should expect her decision to defend to make no difference.

There are many sorts of odd or rare conditions in which an attorney should expect her decision to make no difference, but I am only interested in two very common conditions, which obtain for a significant percentage of attorneys a significant percentage of the time. The first occurs when an attorney is typical. A typical attorney, as I will use the term, need not be average in ability, but rather must not have good reason to think that they are better than average or worse than average. I think this is close enough to common usage of "typical." Of course, a typical attorney can predict that they are probably not exactly average, but the typical attorney lacks good reason to think it more likely that they are worse than average than that they are better, and *vice versa*. Being average is relative to some comparison class, and the comparison class that is relevant here is the class of attorneys who work in the same area on the same sorts of

cases. So a high priced defense attorney might perhaps know that she is better than most public defenders while still not being in a position to think she is better or worse than average for those who serve the same clientele. This would make her typical in the sense that I am interested in. Many, perhaps most, attorneys fit this description. When an attorney is typical in this way, then when a client approaches her, she should expect that, were she to not defend this client, someone else would. Let's call the potential replacement the "backup attorney" and the attorney faced with the choice to defend or not the "original attorney." If the original attorney is typical, then she has no reason to expect that the backup attorney will do any worse than she, or any better. Of course, the backup attorney might be worse, or might be better, and so the decision to defend or not defend might make a difference. But the typical attorney has to no reason to expect either of these possibilities over the other. This is what I mean when I say she should not expect her decision to defend to make a difference.

How does this relate to the reasons for and against defending? Each of the potential reasons for or against defending arises out of some reason-generating consideration, which is a right to be respected or a harm to be reduced or avoided. Let's call these reason-generating considerations *ends*. For cases involving typical attorneys, the achievement of most of these ends is overdetermined in a particular way. If our typical attorney does not defend the defendant, then a backup attorney will, and the typical attorney should not expect that her defending, rather than the backup, will make a difference to the achievement of the end. Let's call ends that are overdetermined in this way *individually overdetermined ends*. An end is an individually overdetermined end for agent A in some situation iff there is some act A can do which would be conducive to the end in this situation, but if A does not perform the act then some backup will, and A should not expect the end to be any more or less likely to come about if she performs the act than if a backup does. How can A's act be conducive to an end if the end will occur with or

without her act? It can still be A's act that actually does bring the end about even if it someone else would have brought it about otherwise. For example, protection of a particular client's rights is an individually overdetermined end for typical attorneys: if typical attorney A defends some client, then it is by her action that the client's rights are protected, even though, had she not acted, someone else would have protected those rights just as well. The increased risk of the client committing future crimes is also an individually overdetermined end for typical attorneys. As far as the original attorney can reasonably expect, the backup attorney is just as likely as she to help the client go free to commit crimes, but if the original attorney does defend, then it is through her defense that the risk is actually increased.

Some of the ends relevant to the choice to defend are overdetermined in a different way. For these ends, whether the end is achieved is the result of the actions of many attorneys in many cases. No one decision about defending should be expected to make any appreciable difference to these ends; no one act is conducive to the end in the sense we have just discussed. Let's call these *collectively overdetermined ends*. An end is collectively overdetermined for agent A iff the end can only be achieved by a great many acts on the part of a great many people, and A should not expect her act or omission to make an appreciable difference to the achievement of the end. Reducing police misconduct and protecting lawyer/client privilege are collectively overdetermined ends. The choice of a single typical lawyer to defend or not defend a single person is not going to make an appreciable difference to whether police gather evidence inappropriately, or whether clients generally trust their lawyers.

The distinction I have introduced between individually and collectively overdetermined ends cut across the division between reasons for and against defending; some of the reasons for defending are individually overdetermined, and so are some of the reasons against defending. Because I want to argue that the reasons to not defend are of a type that categorically outweighs

the reasons for defending, I need to introduce another distinction. Here it is: some of these overdetermined ends are *positive* and some *negative*. I will use these as technical terms. Negative ends are those that are *prima facie* wrong to bring about in cases not involving overdetermination – in situations with no overdetermination, actions that bring these ends about must be justified to be permissible. For example, it is *prima facie* wrong to increase the risk that innocents will be harmed, at least when the risk is not overdetermined, so this increase in risk is a negative end. Positive ends are not ordinarily wrong to bring about in non-overdetermination cases. The literature on criminal defense typically assumes that it is permissible to protect the rights of a defendant, at least when this is not overdetermined, and so the protection of these rights seems to be a positive end. I will say more about both of these examples in the next section. We can combine the positive/negative distinction with the individually/collectively overdetermined distinction and get four categories of ends: individually overdetermined negative ends, individually overdetermined positive ends, collectively overdetermined negative ends, and collectively overdetermined positive ends.<sup>13</sup>

I can now more precisely articulate my argument. As we have just seen, for typical attorneys, all of the ends that generate reasons for and against capably defending the guilty are overdetermined. In the next section, I will argue for two claims about these ends. First, if punishment of a defendant would be justified, and the attorney is typical, then the ends that generate reasons against capably defending that defendant are individually overdetermined negative ends. This is true regardless of what justifies the punishment (unless, as noted above, it is only justified by the need for general deterrence). Second, none of the ends that generate reasons in favor of defending are individually overdetermined negative ends: some are individually overdetermined *positive* ends, and some are collectively overdetermined. I will then

---

<sup>13</sup> We will discuss examples of each of these in the next section.

go on to show that reasons to avoid bringing about individually overdetermined negative ends categorically outweigh the reasons generated by other types of overdetermined ends. Thus, the weight of reasons is against defending the guilty, at least for typical attorneys when the defendant can be justifiably punished. Further, the reasons against defending the guilty are the sorts of reasons that make actions wrong. So it is wrong for typical attorneys to defend a guilty client when the client can be justifiably punished.

**Positive and negative ends**

It is easy to see that the reasons to not defend guilty clients arise out of negative ends. Again, a negative end is one that there is a *prima facie* duty to not bring about in cases that do not involve overdetermination. If the punishment of the defendant can be justified, then it is justified by incapacitation, deterrence, restitution, retribution, or rehabilitation. If it is justified by incapacitation or deterrence, then capable defense increases the risk of future harms to innocents: a defendant who receives less punishment than is called for (on these accounts of justification) is a defendant who is released while still a threat. We ordinarily have a *prima facie* duty to not increase the risk that innocents will be harmed. And so, on this account of the justification of punishment, defending a client whose punishment is justified is a negative end. If punishment is justified by restitution or retribution, then defending the defendant risks that the defendant will avoid justice, either entirely if an acquittal is secured, or partially if the sentence is reduced below what justice demands. If justice really matters enough to justify punishment, then there is a *prima facie* duty to not help someone try to avoid justice. To illustrate, if morality calls for bringing aged and unthreatening war criminals to trial, then there is a wrongful injustice in helping these criminals to hide. Finally, if punishment is justified by the need for rehabilitation, there must be a significant duty to rehabilitate the defendant. This is because there is clearly a *prima facie* duty to not punish people, and so punishment will be justified only when there is an overriding *prima*

*facie* duty. This duty to rehabilitate would then be the sort of duty that we have a *prima facie* duty to not interfere with. For example, assuming that the state had a duty to rehabilitate prisoners, were some citizen to work prevent that – for example by smuggling contraband into prison that undermines this rehabilitation – their action would need to be justified to be permissible. So, for defendants who may be justifiably punished, all of the ends that generate reasons to not defend the guilty are negative ends.

These negative ends are individually overdetermined for typical attorneys. An end is individually overdetermined for an agent iff the agent can do something conducive to the end, but should not expect the end to be any more or less likely to be achieved with or without this act. If a typical attorney did not defend the client, she should expect that someone equally capable would do so. But she can still do something which is conducive to increasing the risk to innocents, or conducive to undermining rehabilitation, or conducive to helping the client escape justice: she can capably defend the client, and if she does, then it will be her action that brought the end about. So, each of the reasons for typical attorneys to not defend clients who should be punished arise from individually overdetermined negative ends.

One might wonder if what I have just said only applies when a defendant has broken a law that prohibits an inherently immoral act, such the law against murder.<sup>14</sup> It should not. There often are strong moral reasons to punish people for breaking so-called *mala prohibita* laws, laws which prohibit conduct which would be permissible if not illegal. For example, those who drive on the left side of the street in the United States do something dangerous, and perhaps immorally reckless, even though what they do might be morally permissible if it were not illegal. The punishment of such a person can potentially be justified by considerations of specific deterrence, incapacitation, or retribution, and my argument would apply when it is. However,

---

<sup>14</sup> My thanks to an anonymous referee for bringing this concern to my attention.

punishment of some particular defendants for violating some *mala prohibita* laws might not be justifiable in the above ways. The specific punishment of the specific defendant would not be deserved, nor would it deter, rehabilitate, or properly incapacitate. Rather, the punishment could only be justified on the grounds that the social practice of punishing such acts generally deters their commission. Parking tickets may be an example of this sort of thing. As I mentioned earlier in the paper, my argument does not apply to such punishments even if they are morally justifiable (although I do not think they are). Now I can explain why: in such cases, there is no individually overdetermined negative end implicated in defending. Since it is the practice of punishment that deters, but not any specific punishment, defending would not risk harming or wronging anyone.

None of the ends that generate reasons *to* defend the guilty are individually overdetermined negative ends. As noted above, these ends fall into two categories: those which have to do with the interests or rights of the specific defendant, and those which arise out of larger social considerations. In the ordinary world, the latter ends are collectively overdetermined. The ends arising from the defendant's rights or interests, on the other hand, are individually overdetermined *positive* ends. As noted in the introduction to the paper, I have assumed that the defendant's rights or interests are morally significant enough to justify defense in non-overdetermination cases. Obviously, if this assumption were false then my thesis would straightforwardly follow. Given this assumption, these ends are by definition positive – it is not wrong to bring them about in the absence of overdetermination.

One might worry that I have overlooked some negative ends that generate reasons to defend. Choosing to not defend a client is failing to protect their rights, and there arguably is a *prima facie* duty to protect rights. This seems to suggest that there are reasons to defend that arise out of negative ends: it is not merely the promotion of the positive end of rights protection that

counts in favor of defending, but also the wrongness of failing to defend rights that counts against not defending. However, in cases of overdetermination a defendant's rights are not threatened by the decision to not defend. So in the cases we are talking about there is no relevant negative end of failing to protect a defendant's rights (we will see this illustrated at length in the next section). One might still worry that not being the person who protects this defendant's rights is a negative end. This end is brought about by the decision to not defend a defendant, even if their rights are still protected, and it is somewhat intuitive that it is impermissible to bring about. However, the question of whether we may or may not bring this end about is exactly what is at issue in this paper. We thus cannot take it to be a negative end without first working through my argument.

One might object that the ends promoted by defending the guilty are not really negative when attorneys bring them about. One might say that, even if it is ordinarily wrong to help criminals go free and escape justice, because attorneys have a special social role this changes the moral status of what they do. However, merely playing a role does not automatically make acting out that role permissible. We can see this with two examples. First, the actions of secret policemen in fascist states are not permissible, despite the special social role these secret policemen play. Of course, this may be because their social role is not at all morally valuable. So let's consider a second example. Assume that police in an ordinarily decent state are generally justified in virtue of their social role in using the threat of coercion to compel obedience with the law. Even so, it would still be morally wrong for a police officer to knowingly force conformity with the law on someone speeding their pregnant wife to the hospital in the middle of the night (when no other cars are around), whether or not the officer's social role required that they do

so.<sup>15</sup> What these examples show is that the reasons one has to act in accord with some social role, if any, arise from the morally valuable ends that roles serve.<sup>16</sup> Secret police do not have real moral reasons to perform their roles because their roles serve no valuable ends at all, and the police officer does not have good reason to enforce the law in the given case because in this situation that enforcement does not serve the ends which motivate the law. This view about what reasons one has to act in accord with one's social role seem to be accepted by those who discuss criminal defense. The reasons given in this literature to explain the permissibility of criminal defense are reasons that are present in all real-world cases of criminal defense: a generally-held right to be defended, or the possibility of excessive punishment, or the potential for prosecutorial misconduct. Appeal to these sorts of reasons would not be necessary if defense were justified just because it typically served a valuable social role.

Perhaps one can have reasons to act in accord with a social role arising from the commitment one has made to play that role, even when these acts do not serve the ends that make the role valuable. It is widely acknowledged, however, that commitments to act immorally – even commitments to play otherwise valuable social roles – are non-binding.<sup>17</sup> In cases of overdetermination, an attorney's action in accordance with their social role does not serve any ends, since these ends would be promoted either way. If, as I will show, the defense attorney's acting in accord with their role is immoral in cases of overdetermination, their commitment to act in accord with this role will be non-binding.

---

<sup>15</sup> It is possible that a decent state would have laws requiring people to conform to speed limits in such cases. This is because it is hard to write effective laws that still make exceptions for every appropriate case; the laws become too difficult to understand, predict, or enforce. So there will be cases in which even the best-written law will clash with morality.

<sup>16</sup> See, for example, David B. Wilkins, 'In Defense of Law and Morality: Why Lawyers Should Have a Prima Facie Duty to Obey the Law', *William & Mary Law Review* 38 (1996) pp. 269-295.

<sup>17</sup> See e.g. Jeffery Brand-Ballard, *Limits of Legality: the Ethics of Lawless Judging*. (Oxford University Press 2010).

Have I overlooked the possibility that attorneys are free riding when they fail to defend, or in some sense undermining the justice system, and this gives additional reasons to defend? While I think that the issue of undermining is covered by what I have said – the end of not undermining the justice system is collectively overdetermined – I will return to these concerns at the end of the paper when I consider objections.

To summarize: all the reasons against typical attorneys defending clients who would be justifiably punished arise from individually overdetermined negative ends. All of the reasons in favor of defending arise either from individually overdetermined positive ends, or from collectively overdetermined ends. The next step in the argument is to show that reasons to avoid bringing about individually overdetermined negative ends categorically override those given by the other sorts of overdetermined ends. I will start by considering individually overdetermined negative ends versus individually overdetermined positive ends.

**Reasons and individually overdetermined ends**

I will start with some examples and then explain how we can generalize from them. Let's start with a clear case involving an individually overdetermined negative end:

*Gunstore:* Lona runs a gun shop. One day a customer comes in and says that he wants to buy a gun with which to murder his wife. Unfortunately, Lona's phone is broken, and the security measures intended to protect Lona from her customers make it impossible for her to stop the customer. The customer tells Lona that if she does not sell him the gun, he will just buy it from another shop, which is obviously true. Realizing that she cannot prevent the wife's murder, Lona figures she might as well make some money, and sells the man a gun.

It is wrong to sell someone a gun knowing they will use it to murder. It is wrong even if the sale is individually overdetermined. This shows that it can sometimes be wrong to be the person who

brings about an individually overdetermined negative ends.<sup>18</sup> But one might wonder if this is only true when the end is extremely serious, as it is in the *Gunstore* case. To address that, let's consider an example with a less extreme negative end:

*Smashing pumpkin:* A young child has carved a jack-o-lantern of which he feels proud.

Trent and Marilyn happen to be standing nearby, holding hammers. Marilyn is just about to smash the pumpkin, which will make the child sad. But then she says to Trent, "If you want to smash this, go ahead. Otherwise I am going to." Trent ordinarily would not do this, but he realizes he cannot stop Marilyn, so he smashes the pumpkin. This makes the child sad.

Trent expects to make no difference to whether the child is sad, and the sadness does not counterfactually depend on his action, yet he is still responsible for it. It is wrong for him to smash the pumpkin. We see from this that it can be wrong to bring about individually overdetermined negative ends, even when those ends involve only minor wrongs.

The above two cases are ones in which the overdetermination involved certainty: the killing or the sadness would definitely occur no matter what. Defending the guilty has no certain outcomes, only risks. Let's consider a case involving overdetermined risks:

*Pram:* Lucy is curious if an empty baby carriage shoved into a roadway will roll safely across. The carriage belongs to a young child who bought it with his allowance money.

If the carriage is destroyed, the child will be very sad, but otherwise will not be. Because the street is not very busy, the chance of the baby carriage being hit is only moderate.

Lucy offers \$10 to the first person to give it a try. Jack says he will try, and Jill knows that

---

<sup>18</sup> Remember that an end is only individually overdetermined for an agent if the act can act in a way *conducive* to the end; by acting in this way, the agent can be the one who brought the end about, even though it would have happened without their act.

he is sincere. She cannot stop him, but she knows that, if she asks, he will let her push it instead. She figures she might as well be the person who gets the \$10, and gives it a push. Jill does something wrong by pushing the carriage into the street, even though Jack would have done it otherwise, and even though she is only risking the child's sadness. She is responsible for the risk, even though the risk does not counterfactually depend on her actions. Notice here that the overdetermination in this case is just like the individual overdetermination relevant to typical attorneys defending the guilty. There is a chance that, were Jack to push the carriage, it would be less risky than were Jill to push it – perhaps he would time the push better, or push it faster. And there is a chance that, were Jack to push the carriage, the risk would be higher. Jill has no information either way, and so she should not expect to make a difference to the relevant negative ends in exactly the sense I talk about earlier in the paper. Her action is wrong, and this shows that there can be a *prima facie* duty to not bring about individually overdetermined negative ends, even when the ends involve only risks of relatively minor harms.

Let's now consider a case involving individually overdetermined positive ends:

*Bathing baby 1:* There is a drowning baby in a pool. Leslie could easily save the baby.

Howard could also easily save the baby. Both know that the other will save the baby if he or she chooses not to. Leslie stands back, letting Howard be the rescuer.

If Leslie had saved the baby, she would be responsible for the rescue occurring, although the rescue would not depend on her action. But she does nothing wrong by not saving the baby. This is true even though she would be obligated to save the baby were Howard not around. In this case, Leslie knows that the end will be achieved with or without her. We get a similar lack of obligation when there is just an expectation that one's action makes no difference:

*Bathing baby 2:* There is a child drowning in the ocean. The only way to save it is to

throw out a life preserver, but there is only one to throw. Leslie knows that if she does not

throw the preserver, then Howard will. The only concern is about whether the preserver will reach the child. Neither has any information about who has the stronger or more accurate throwing arm. Leslie lets Howard throw the preserver.<sup>19</sup>

Here again it is permissible for Leslie to let Howard act to save the child, because she has no reason to think she would do better than he would. We can thus conclude that there is no obligation – not even a *prima facie* one – to bring about individually overdetermined positive ends, even when those ends are extremely morally significant.

While it is wrong to merely risk making a child sad, even when that risk is individually overdetermined, it is not wrong to refrain from saving a life when the saving is individually overdetermined. Because it is wrong to bring about individually overdetermined negative ends, even when they are of relatively little moral significance (as in *Pram* and *Smashing pumpkin*) and it is permissible to not bring about an individually overdetermined positive end, even when that end is very important, we can reasonably generalize to cases involving more significant negative ends and less significant positive ends. It is *prima facie* wrong to bring about negative ends, whether or not they are individually overdetermined. This seems to be because bringing them about makes one responsible for their obtaining, despite the lack of counterfactual dependence. It is generally permissible to not bring about individually overdetermined positive ends, even when it would be obligatory to bring about the positive end were it not overdetermined.

These claims are not just supported by intuitions about these cases. It is a familiar truism that it can be wrong for *me* to kill one innocent, even if that prevents someone else from committing more killings.<sup>20</sup> This coheres nicely with intuitions about individually overdetermined negative ends. Morality does not seem agent-centered in this way with regards

---

<sup>19</sup> My thanks to [removed for blind review] for her help with this example.

<sup>20</sup> This truism is a widely discussed; for just one example, see Stephen L. Darwall, 'Agent-centered restrictions from the inside out,' *Philosophical Studies* 50 (1986) pp. 291-319.

to positive ends. An agent cannot force another into wrongdoing by fulfilling that other agent's obligations (at least, not in usual cases).

Can we conclude from this that it would be wrong to bring about an individually overdetermined negative end even if doing so is the only way to bring about an individually overdetermined positive end? Not yet. Sometimes it is permissible to violate a *prima facie* obligation in order to do something supererogatory. For example, imagine that one wanted to donate a kidney to save a stranger, which is not obligatory. However, doing so is only possible if one steals a car – let's say the donation is extremely time sensitive, and the theft is necessary to make it to hospital before it is too late. Even though stealing is *prima facie* wrong, it is permissible in this situation. So, in order to determine the moral status of bringing about overdetermined negative ends by defending the guilty, we have to consider whether the reasons to defend might be weightier than the reason to not defend, even if the reasons to defend do not create obligations.

How do we compare the weight of the reasons for and against bringing about individually overdetermined positive and negative ends? The most obvious way unfortunately is not trustworthy. The most obvious way is to consider examples in which an agent must bring about some individually overdetermined negative end in order to bring about some individually overdetermined positive end. Unfortunately, these examples will be either too ordinary or too bizarre. The bizarre cases I have been able to come up with are so bizarre that I find few strong intuitions about them. The ordinary cases look just like the subject of this paper: they tend to involve people (like attorneys) doing their jobs. My worry is that our intuitions about the permissibility of an individual doing such a job tend to conflate the permissibility of doing one's job when the outcome is overdetermined with the permissibility of doing the job absent overdetermination, or with the permissibility of the job existing at all. This worry arises from my

discussions of these cases with other philosophers and with people outside of philosophy. People report being gripped by the fact that it would be harmful if *no one* did such jobs, even if it makes no difference that some particular agent does the job or not. Typically, this sort of grip is a sign that we have discovered an important morally relevant feature of a case. But this does not make sense in this context. It's clear that Mary's failure to act in the two *Bathing baby* cases is permissible, even though it would be terrible if no one were to save the baby; these intuitions are also widely shared. But these intuitions are not coherent with the idea that, in ordinary "doing one's job" cases, it matters that it would be harmful were no one to do these jobs. Since people do seem aware of and attentive to the overdetermination in the *Bathing baby* cases, and seem to be hung up on facts that have nothing to do with the overdetermination in the "doing one's job" cases, the best explanation of this incoherence is that the intuitions in the job cases for some reason are not properly sensitive to the overdetermination. We cannot trust them, and so we need some other way of comparing the weight of the reasons to (not) bring about individually overdetermined positive and negative ends.

To make this comparison, let's use a variant of the first *Bathing baby* case. Imagine that Leslie knows Howard is planning to jump in to save the baby, and that the only way she can be the one to save it is through committing some sort of ordinarily immoral act to prevent his diving in. She also knows that Howard will successfully save the baby if he dives in, and that she would successfully save the baby if Howard does not dive in. By determining what otherwise immoral behavior Leslie would be justified in committing in order to be the one who saves the baby, we will see just how strong the reasons are to bring about individually overdetermined positive ends; they are strong enough to justify that amount of wrongdoing. However, there is *no* immoral act that Leslie would be justified in doing to prevent Howard from diving in. Consider some very minor wrongs Leslie might do: in order to distract him, she might slap Howard, throw his phone

in the bushes, insult his mother, or tell a harmless (but immoral) lie. Leslie would clearly be justified in doing none of these in order to save the baby, and it seems that this must be because the baby would be saved without these acts. Since the overdetermined end in question – saving a baby’s life – is quite important, and these wrongs very minor, it is safe to generalize from this to the conclusion that one is never justified in committing any act that would otherwise be wrong in order to bring about an individually overdetermined positive end.

Let’s briefly consider an objection to this last argument. All I have shown is that one is not justified in committing certain very minor wrongs in order to bring about an individually overdetermined positive end. I want to draw the conclusion that one is not justified in bringing about individually overdetermined negative ends in order to also bring about individually overdetermined positive ends. But this conclusion only follows if the reasons to not bring about individually overdetermined negative ends are as strong or stronger than the reasons to not commit the minor wrongs I discussed in the previous paragraph. I have not yet shown this. In response, let’s consider variants of the examples we have already discussed. Imagine that Lona, in the *Gunstore* case, wants to not sell a gun to the man who plans on murdering his wife. For some reason, she can only avoid this by lying to the man (perhaps she lies that her cash register is broken so she cannot sell any guns). Such a lie would clearly be justified in this case, even though in other circumstances one may not lie to a customer. In *Smashing pumpkin*, Trent would be justified in committing a minor wrong if that was the only way to avoid smashing the child’s pumpkin. For example, if he could only avoid smashing the pumpkin by insulting Marilyn’s mother, this would be justified. So the reasons to avoid bringing about individually overdetermined negative ends are greater than the reasons to avoid committing the sorts of minor wrongs I discussed in the previous paragraph. And since commission of such minor wrongs is not justified by the need to bring about individually overdetermined positive ends, we

can conclude that one cannot justify bringing about individually overdetermined negative ends in order to bring about individually overdetermined positive ends.

What have we learned? Bringing about individually overdetermined negative ends is *prima facie* wrong. A typical attorney who defends a client whose punishment would be justified brings about some individually overdetermined negative end. One is not justified in committing even the most minor wrongdoing in order to bring about individually overdetermined positive ends. Thus, a typical attorney cannot justify defending a guilty defendant on the grounds that this will protect the defendant's rights or interests, as long as the defendant's punishment would be justified.

There is another sort of reasons in favor of defending that we have not yet discussed: reasons arising from collectively overdetermined ends. We will turn to these next.

**Reasons and collectively overdetermined ends**

Let's start again with some examples:

*Los Angeles:* Carla lives in Los Angeles. Her choice to drive, rather than taking the bus, would make no noticeable difference to traffic or pollution, but if a significant percentage of the driving population chose to take the bus, pollution and traffic would be greatly alleviated, saving thousands of lives. Carla drives her car.

*President:* It is the presidential election. One candidate (but not the other) will help the nation withdraw from a pointless war, saving thousands of lives. Ralph, who is a highly trained and knowledgeable statistician, knows that the preferred candidate will win. He would have voted for the better candidate had he voted, but chooses not to vote at all.

Philosophers are divided on whether Ralph and Carla have moral obligations that they are violating.<sup>21</sup> Let's sidestep that debate and assume for the sake of argument that both do (this is the assumption that helps my argument the least). The important question for our purposes is how the reasons which generate these obligations weigh against the reasons which make it wrong to bring about individually overdetermined negative ends. Intuitively, Carla's reasons to take the bus, or Ralph's reasons to vote, are easily overridden. If we ask what otherwise immoral act Carla or Ralph would be justified in committing to satisfy their obligations to vote or take the bus, the answer seems to be "none." If Carla can keep a minor (but not immoral) promise only by driving her car, breaking this promise would not be justified by the *prima facie* obligation to take the bus. If Ralph has to break a minor promise in order to vote, this would not be justified by its necessary role in allowing Ralph to vote.

We can draw some generalizations from these intuitions about these examples. This is first because the relevant overdetermined ends in these two cases are hugely more significant than those in any of the individually overdetermined cases discussed above, since they involve saving thousands of lives. Further, these examples illustrate two important categories of collectively overdetermined ends. In the *Los Angeles* example, the end that is overdetermined is a bad end – Los Angeles will have bad traffic no matter what Carla does. In the *President* example, the overdetermined end is good – the right president will be elected. Since Carla and Ralph's reasons to be part of the groups which work towards these ends are overridden by reasons to not commit even minor wrongs, despite the importance of the ends and regardless of whether the ends are good or bad, we can generalize that all reasons to be part of groups which work towards collectively overdetermined ends are overridden by reasons to not commit even minor wrongs.

---

<sup>21</sup> See, e.g. Derek Parfit, *Reasons and Persons* (Oxford: Clarendon 1984) versus Walter Sinnott-Armstrong, 'It's not my fault'. In Caney, Gardiner, Jamieson, Shue (eds) *Climate Ethics: Essential Readings* (New York: Oxford University Press 2010).

While these intuitions seem widely shared, some deny them and claim that we do have relatively strong obligations to work towards (some) collectively overdetermined ends.<sup>22</sup> Rather than trying to completely deny this claims, I will show that this does not make a significant difference to my thesis. First, note that most of the collectively overdetermined ends which generate reasons to defend are also individually overdetermined. That is, if one typical attorney does not defend the client, someone else will, and is likely to be equally capable. So the first attorney makes no expected difference to the amount of effort towards or away from the relevant end. The choice attorneys face is between being the person who makes an unnoticeable contribution or letting someone else make that identical contribution. This is different than in standard collective overdetermination cases: if Ralph votes, for example, this contributes something that would otherwise not be contributed towards the end, even though that contribution is more or less unnoticeable. So, even if collectively overdetermined ends could generate very weighty duties, these ends would mostly not generate very weighty duties to defend the guilty, since they are weakened by being individually overdetermined as well.

However, at least one of the ends which generates reasons to defend the guilty is not both individually and collectively overdetermined. This is the end of protecting the lawyer/client relationship. It is only collectively overdetermined, and not also individually overdetermined, because it is worked towards by each individual attorney not turning down guilty clients. Does this generate a weighty duty to defend? The most plausible cases in which collectively overdetermined ends generate strong obligations are those where the end is a harm that is occurring – the cases Parfit famously discusses involve collective torture, but there also seems to be strong reasons to avoid participating in collective oppression or destruction of the environment, for example. In such cases, it seems plausible that there are duties arising from

---

<sup>22</sup> E.g. Parfit, *Reasons and Persons*, Brand-Ballard, *Limits of Legality*.

collective responsibility for the harm, or perhaps duties arising out of the need to avoid complicity. These cannot be present when the collectively overdetermined end is some good thing which is or will be achieved, or when the end is some harm which is or will be avoided. It is much less plausible that there are strong duties to work towards these sorts of ends. (Some think that we do have reasons to contribute which have to do with fairness or free riding; I will discuss fairness and free riding in the next section) Even if collectively overdetermined harms could generate strong obligations, typical attorneys in our current legal system would not have strong obligations to defend the guilty. The end of protecting the lawyer/client relationship is currently being achieved, and so this is not a collectively overdetermined harm. If this end did generate significant duties, it would only be because many lawyers were systematically turning down guilty clients. If this were to obtain, and one of the controversial views claiming we have strong duties to avoid contributing to collectively overdetermined harms were true, then my thesis might no longer apply to typical attorneys. This conjunction is unlikely, but those who endorse something this sort of view can qualify my thesis as follows: it is wrong for typical attorneys to defend clients who should be punished, unless other attorneys systematically refuse to do so to the extent that this is harming the lawyer/client relationship.

So, what have we seen? It is controversial that people have obligations to work towards collectively overdetermined ends. But even if they do, most will agree that the reasons that generate these obligations are overridden by any other obligation, no matter how minor. This is true either in general, or at least in the world as it currently is. It is wrong for typical attorneys to defend clients whose punishment would be justified, and this wrong cannot be justified by the attorney's need to work towards collectively overdetermined ends. The wrong also cannot be justified by the attorney's need to be the one who brings about individually overdetermined

positive ends. These two points cover all the ends which I have claimed generate reasons to defend the guilty. In the next section, I will consider a number of objections to this argument.

### **Objections**

The main objections to my view fall into two categories. Some have to do with alleged bad consequences of my view (although one need not be a consequentialist to advance this sort of objection). Others have to do with concerns about fairness.

One might worry that what I say cannot be true because the justice system could not survive if attorneys adopted this view. In its weakest form, this objection starts from the claim that, were attorneys to adopt my view, they would stop defending those of their clients who should be punished. But then no one would defend these clients, who have a right to be defended. It cannot be the case that universal conformity with moral requirements leads to systematic rights violations. This version of the objection is easily responded to. Everything I have said in this paper is consistent with it being permissible to defend when defense is not overdetermined. Were all attorneys to adopt the view argued for in this paper, then for each defendant, some attorney would be able to predict that there is likely not a backup attorney to defend were she to reject the defendant. My view does not require this attorney to turn the client down.<sup>23</sup>

---

<sup>23</sup> We can imagine a different sort of odd scenario arising if everyone adopted my view. If all attorneys knew that other attorneys held my view, they might predict that, for any given client, there would not be a backup attorney. So they should refuse no one. But they would then know that other attorneys would predict this, and so they could predict that there *would* be a backup attorney, so they should refuse all punishable clients. But then they can predict that other attorneys would predict this ... and so forth.

I do not see this sort of instability as a problem for my view. Rather, it is a problem for any view about moral behavior in social conditions. There will always be odd coordination problems that arise in sufficiently abnormal circumstances (and universal adoption of the correct moral view is a very abnormal circumstance). I am not sure what the right thing for any attorney to do in this situation would be, but the answer to this will fall out of the correct moral decision theory.

There is a more sophisticated version of this objection. This version points out that, were attorneys to adopt this view, they would sometimes mistakenly refuse to defend. They could mistakenly think that a defendant's punishment would be justified, and so refuse to defend when that defense does not actually bring about negative ends. Or they could mistakenly think that other attorneys will defend the client, and so refuse to defend when the defense is not overdetermined. The first worry is not that serious, since some backup attorney will defend the client so no harm will be done. The second is more of a concern, because it would lead to the defendant's rights being violated. But this is not a serious objection to my view. People can be mistaken about what they should do, and sometimes acting when they are mistaken leads to violations of rights. That a moral theory allows for this to happen is no objection to it. Consider the claim that people have obligations to save drowning children. Sometimes people will be mistaken about who is drowning. They will pull children out of pools that those children have every right to be in. This is not an objection to the claim that we have an obligation to save drowning children.

It might be a problem if adoption of my view led to *systematic* violations of rights. But it need not lead to this. Rights would only be violated when a defense attorney mistakenly believes that defense is overdetermined and so refuses a client. Such mistakes will be rare, as long as attorneys are reasonably conscientious, and it is no flaw that a moral view leads to problems when applied unconscientiously. These mistakes will be rare because it will be rare that one is the last attorney in a position to defend a client without also having some information that this is the case. The presence of this information would eliminate the overdetermination, and so one who adopted by view could at that point defend.

Let's turn then to the other class of objections to my view: objections having to do with unfairness. If my view were true, then Attorney Z may be forced to defend a client because

attorneys A-Y refuse to do so. This would occur when Z is last attorney left to defend that defendant, so that A-Y all have backup attorneys when they refuse but Z does not. That seems unfair for a number of reasons. First, Z has an obligation that A-Y do not have, though there is no difference between Z and any of the other attorneys. Second, attorney Y in some sense forces this obligation on Z: if Y and Z are the last available attorneys to defend the client, once Y says “No,” Z acquires the obligation to say yes<sup>24</sup>. Third, A-Y are allowed to force Z to bear the costs of defending the client.<sup>25</sup> Finally, it seems that A-Y are free riding on Z’s willingness to defend the client: they can only say “No” because they expect Z to step in and say, “Yes.”<sup>26</sup> Many authors have claimed that morality should not make unfair demands of us, and so these varieties of unfairness pose a problem for my view.<sup>27</sup>

I will make some general points in response and then address the specific varieties of unfairness I have just listed. For the record, I do not think it makes any sense to apply the term “unfair” to putative moral obligations. Fairness is a moral evaluation, and morality evaluates actions, states of affairs, policies, or people. Putative moral obligations are none of these. But even if I am wrong about that, the requirements of my view are not unfair. It would be unfair if, on my view, two people in the same situation were held to different standards – if they faced different moral requirements. But that is not my view: every attorney is held to the same moral standard. A may refuse to defend when A expects to make no difference. That is true of B-Z as well. Were A and Z’s situations reversed, then Z could refuse to defend the client. This does not seem unfair to me.

---

<sup>24</sup> See M.B.E. Smith, ‘Is there a prima facie obligation to obey the law?’ *Yale Law Journal* (1973) pp. 950-976 for an argument that we cannot force obligations on others.

<sup>25</sup> Thanks to [removed for blind review] for bringing this version of the objection to my attention.

<sup>26</sup> See Brand-Ballard, *Limits of Legality*, for more this worry in a different legal context.

<sup>27</sup> David Copp, “‘Ought’ Implies ‘Can’, Blameworthiness, and the Principle of Alternate Possibilities’. In McKenna, M. and D. Widerker (eds.), *Moral Responsibility and Alternative Possibilities* (Aldershot: Ashgate Press 2003).

That is my general response. Let's now consider the specific claims about unfairness that I have raised as objections to my view. First, and weakest, one might say it is unfair for two attorneys who are not different in any relevant way to be faced with different moral obligations. That is not unfair, or if it is it is a sort of unfairness that is undeniably pervasive in morality. If my identical twin is near a drowning baby and I am not, my twin has an obligation I lack. Moral requirements vary with situations, and this is not a problematic sort of unfairness. Second, one might think it is unfair that attorneys A-Y must force an obligation to defend on Z. That is, on my view A-Y can be obligated to turn down a client, because they know that there are backup attorneys; since Z is the last attorney in line, A-Y are morally required to force an obligation to defend on Z. But this is not problematic either. This sort of thing can arise in the context of uncontroversial moral obligations. Here is an example. Barry and Mary see a drowning baby. For some reason, the baby cannot be saved in a straightforward way by either. But Barry has the ability to quickly construct a device that allows the baby to be saved. However, the device requires two operators. If Barry does not construct the device, Mary has no obligation to save the baby, as she is unable to do so. However, if Barry does construct the device, Mary is now obligated to help him operate it. Barry is clearly obligated to construct this device. It is thus hard to deny that we sometimes have obligations to force obligations on other people, and the fact that my view can lead to this is not a real objection to it.

The last two versions of the unfairness objection are that A-Y force Z to bear a cost they do not have to bear, and that they free ride on Z's actions. I do not see either as problematic, as long as each of A-Y would be required to bear the cost were they in Z's shoes. We have already seen that morality can require us to force costs on others – it can require us to force obligations on others, and acting on obligations almost always has (at least) opportunity costs. What I think really worries people about this sort of unfairness is that it seems selfish to force a cost on another

that one is unwilling to bear one's self. We do not want others to get away with doing nothing. But that is not what my view endorses. I have argued that we should be responsive to our moral reasons, but that these vary in surprising ways depending on our circumstances. When typical attorneys have no obligation to defend clients who should be punished, then they should not defend these clients. This is not "getting away" with anything. Instead, it is just being faced with different moral reasons than one would be faced with were the defense not overdetermined. It might perhaps be problematic if an attorney were manipulating situations to avoid ever being the backup attorney: this would be more like traditional free riding, in which one puts one's self interest over one's social or moral obligations. But this is not what I have advocated.

**Conclusion**

The typical attorney should expect her choice to capably defend a client to make no difference to morally relevant ends. Even so, capably defending a client whose punishment would be justified makes her responsible for bringing about individually overdetermined negative ends. This makes the defense *prima facie* wrong. The reasons typical attorneys have to bring about other types of overdetermined ends cannot justify even the smallest amount of wrongdoing. This is true even if the attorney would be justified in defending absent overdetermination. Thus, if the client's punishment would be justified, it is wrong for the typical attorney to defend that client.

I will conclude by briefly sketching how to apply my argument to other sorts of cases involving criminal punishment. One sort involves defense of defendants who risk being punished too much, and the other involves prosecution of those who might possibly be innocent.

Consider a defendant who should be punished to a certain extent, but where it is possible that they will instead be over-punished. This might occur when the state gives judges or juries significant discretion in sentencing. One reason to defend such a person is to bring about the end

of avoiding over-punishment. This is a positive end, and it is an individually overdetermined positive end for typical attorneys. On the other hand, as long as this defendant *might* receive proper punishment, defending capably brings about individually overdetermined negative ends: it increases the risk that the defendant will be acquitted or have his sentence reduced below one that would be justified, which increases the risk that the defendant will avoid justified punishment. So, in cases where the defendant might receive too much punishment, but also might receive just enough, some of the reasons against defending still arise from individually overdetermined negative ends, while none of the reasons in favor of defending do. And so my argument shows that it would still be wrong to defend in these conditions.

A similar argument can be made about prosecution in relatively common situations. If a prosecutor expects that she would be replaced by an equally capable prosecutor were she to step down from a case, and there is a chance that the defendant in the case is innocent, then the risk of incriminating an innocent person is an individually overdetermined negative end. The chance of convicting a guilty person in this case is an individually overdetermined positive end. So, for replaceable prosecutors and possibly innocent defendants, prosecution would be immoral even if the defendant were almost definitely guilty.<sup>28</sup>

---

<sup>28</sup> [Removed for blind review]