

The Morality of Blackmail

James R. Shaw

Draft, please don't cite

Blackmail raises a pair of parallel legal and moral problems referred to as the “paradox of blackmail”.¹ It can be morally permissible and legal merely to threaten an individual with an action one could legally and morally carry out (e.g., reveal their marital indiscretions to their spouse). It can also be legal and permissible merely to request money from them. How then could it be illegal or morally impermissible do them jointly?

I address the moral version of this paradox by subsuming blackmail under a general account of wrongful coercion. According to this account, and contrary to the intuitions giving rise to the paradoxes, a blackmailer’s threat to release harmful information is permissible unless it is impermissible to execute that threat. The tenability of the account turns on identifying a subtle wrong that can arise in the relevant harmful releases of information and on developing tools for understanding how a range of distinct moral issues interacts with that wrong.

The account makes no pronouncement on the corresponding legal versions of the paradox.² But it does show that the core moral case for the illegality of blackmail is precisely the same moral case for the illegality of the would-be threatened behavior that is presently legal. This suggests the presence of a genuine tension in existing blackmail law which may merit legal revisions.

1 The Paradoxes of Blackmail

1.1 Coercion

I use the word “coercion” to cover both permissible and impermissible threats of contingently enforced sanctions aimed at shaping an agent’s behavior. I define blackmail as an impermissible coercive threat to disclose harmful information. One can use “blackmail” more broadly to include permissible threats of information release or acts of extortion that don’t involve information. One can also

¹Williams (1954) is an early discussion of the legal problem, and primary focus of the literature, though without mention of ‘paradox’. Clark (1994) notes the label might be sensational as some pairs of separately legal and permissible actions are jointly illegal or impermissible—e.g., drinking and driving. The force of the puzzles draws on the asymmetries between blackmail and such cases.

²See Lindgren (1984), Wertheimer (1987), and Berman (2011) for an overview of many positions in the literature.

use “coercion” more narrowly, so that all types of coercion are impermissible. In the context of this paper, these are terminological choices.

Since blackmail is just one kind of wrongful coercion, a natural way to understand blackmail involves integrating it into an general account of impermissible coercion. Coercive announcements aim to influence deliberation by attaching sanctions to an agent’s available actions. When the mugger announces “your money or your life”, he tries to get you to hand over your money by persuading you that walking away with it now costs your life. When the mother tells her child “no dessert unless you finish your meal” she tries to convince him to eat his brussels sprouts by persuading him that failure to do so precludes his having ice cream later.

Since coercive announcements are designed to affect deliberation, the following account of wrongful coercive announcements suggests itself.

Simple Account: A coercive announcement is impermissible if it attaches a sanction to an option in an agent’s deliberation that the agent is entitled to deliberate without.³

Two clarifications about what *Simple Account* is not doing: First, I haven’t discussed what counts as a coercive announcement to begin with. What makes an announcement attach a restrictive sanction in deliberation characteristic of coercive announcements (as opposed to offers, which open up new deliberative options)? The question runs orthogonal to my arguments here, though I will presume some account of the relevant distinctions is available.⁴

Second, as regards explanations of wrongful coercion, *Simple Account* overtly passes the buck: it requires an independent account of when an agent is entitled to deliberate about an available action without sanction. What are our rights to deliberation? The simple account doesn’t provide guidance in these matters, and I won’t supply general answers to this question here. The reason for this is that the simple account combines with a broad principle governing entitlements to deliberation which has obvious intuitive merits and will suffice to cope with all cases I treat in the course of this paper.

Permissible Deliberation: If one is permitted to deliberate about performing action *A* at all, one is entitled to deliberate as if one could perform *A* free of any sanction on its performance that would constitute a wrong.

One is entitled, definitionally, not to be wronged. Accordingly, if one has rights to deliberate at all, we should expect an equal entitlement to deliberate *as if* one would not be wronged. Put another way, anyone who owes you, on pain of wronging you, not to do *X*, likewise owes you the right to deliberate about a permissible choice as if they will not do *X* pursuant to it. A full account of deliberative rights may ultimately be complex, but the partial account given by

³Perhaps an announcement is only wrongful if it has some actual effect on deliberation—otherwise the agent suffers no ‘deliberative harms’. I’ll ignore this complication.

⁴See Nozick (1969), Gorr (1986), Berman (2002) for elaborations and discussions of the common ‘baseline’ approaches to distinguishing threats from offers.

Permissible Deliberation seems justified on intuitive grounds without delving into additional detail.

Simple Account and *Permissible Deliberation* allow us to classify a wide range of coercive announcements as impermissible. For example, suppose that *B* can permissibly deliberate as to whether or not to do *Y*. Then we have:

Basic Coercion: If it is impermissible for *A* to *X* while *B* does *Y*, or for *A* not to *X* while *B* doesn't *Y*, then it is impermissible for *A* to communicate to *B* that *A* will *X* just in case *B* does *Y*.

Basic Coercion tells us that, for an important range of cases—those in which an agent was permitted to perform one of several actions—announcements attaching sanctions to actions inherit their impermissibility from the impermissibility of carrying out the sanctions. This is why the mugger's announcement is impermissible: it is impermissible for the mugger to shoot you if you walk away, cash in hand. If *Basic Coercion* is also the primary source of wrongful coercion, it will give an explanation of why parents are permitted to threaten their children with no dessert: it is permissible for parents to deny their children dessert (especially if they haven't finished their dinner). So it is likewise permissible for them to threaten to do so.^{5,6}

The account of coercion developed so far faces a key challenge. *Basic Coercion* threatens to leave key instances of wrongful coercion unexplained, including my targeted example of blackmail. If the paradoxes of blackmail start from correct assumptions, it is sometimes impermissible to threaten to do something that would otherwise be entirely permissible. My aim is to argue that these assumptions are suspect, and that most intuitively impermissible cases of blackmail can be accounted for solely with *Basic Coercion*. This is meant to form part of a 'mutually reinforcing' case for adopting both the simple account of wrongful coercion and my proposed explanation of the particular wrongs of blackmail together.

1.2 Simple Cases: Impermissible Disregard

Both the appeal and the unique challenges of integrating a treatment of blackmail into a theory of coercion have not gone unnoticed. Feinberg (1988), Gorr (1992), and Berman (1998) are examples of increasingly sophisticated theories which treat blackmail as a form of wrongful coercion and appeal to something like *Basic Coercion*. My view has important similarities to that endorsed by Berman who, as I will shortly do, pinpoints elements of motivational structure as key tools in resolving the paradoxes. The ensuing account can helpfully be thought of as deepening, extending, and occasionally amending Berman's treatment by doing two things: focusing more acutely on the moral foundations of

⁵I am non-committal as to whether *Basic Coercion* is the only principle needed to explain instances of impermissible coercion, but I will proceed here as if it is the primary principle of its kind.

⁶For related accounts of coercion, which also endorse *Basic Coercion*, see Wertheimer (1987) and Pallikkathayil (2011).

coercive wrongs, and giving priority of place to the moral version of the paradox of blackmail. These focuses lead to different problem cases to treat,⁷ different insights into familiar problem cases,⁸ and occasionally different verdicts.⁹

But before we get to these complexities, let's begin with the core issue. It seems easy to imagine cases where *A* becomes aware of *B*'s marital infidelity so that it is permissible for *A* to inform *B*'s spouse, though it is impermissible for *A* to use this information to extort money from *B*. But I want to claim that matters here are slightly more complex. To show this, I need to bring into relief a subtle form of moral turpitude which I'll call *impermissible disregard* (or for simplicity *disregard*).¹⁰ I'll then argue that in 'paradoxical' cases of blackmail, carrying out one's threat tends to involve this very moral failure.

The wrong of disregard needn't involve coercive threats, nor rights to information. Consider Fred, a landlord renting his apartment on a monthly lease to Lucy. Fred has warned Lucy that he may terminate their arrangement any month, on short notice. In October, Fred realizes that terminating Lucy's contract will make it very hard on Lucy. This month alone she will be unlikely to find a new apartment in time to move. Also, this would exacerbate her finances which are temporarily troubled through no fault of her own. Suppose Fred nonetheless decides to terminate Lucy's contract, knowing he can find another renter to take her place. Fred's grounds for the change? He's grown a tired of having Lucy in the building, and feels it might be nice to see a change of faces.

Fred may be within his legal rights to do this, but his action is unconscionable. Even though the apartment is Fred's property, and even though he gave Lucy ample warning, his actions show an impermissible disregard for Lucy's well-being in the following sense: he knowingly creates harms to her which are not appropriately offset by the value of his own furthered ends. I say the harms Fred causes Lucy are not 'offset' because the disregard—and impermissibility—of Fred's action may disappear if they promote his own sufficiently worthwhile goals. Suppose Fred has a renter who is willing to pay him twice what Lucy does, and Fred himself is in dire financial straights. Or suppose Fred needs the apartment to move into himself, since his own must be fumigated on short notice. These are considerations which make Fred's actions *less* disregardful, and may ultimately make them permissible. Importantly, the harm to Lucy in these cases is not lessened. Fred may nonetheless be permitted to cause this harm because he has his own rightful goals to pursue. By contrast, Fred is not actually permitted to terminate Lucy's contract because the good of seeing new faces in his building doesn't justify bringing serious, easily avoidable harms to another.

⁷For example, punitive threats and permissible bluffs, treated in §§2.2–2.3, are only manifest as pressing problems once we probe the foundations of coercive wrongs.

⁸In particular, threats to do obligatory acts, discussed in §2.1, require added commentary when we focus on moral, and not legal, permissibility.

⁹In particular, I will be led to a different treatment of the most problematic cases for each of Feinberg, Gorr, and Berman—those involving what appear to be morally discretionary information releases to do moral good, treated at the end of §2.1.

¹⁰I'll remain neutral on the foundational sources of this wrong but it has clear connections, for example, to something like a Kantian duty of beneficence.

Saying precisely how harms and furthered-ends are weighed in assessing an act for disregard, and what counts as a legitimate kind of end to offset a harm, are both delicate matters that I won't pursue. My arguments only presuppose that disregard exists as a form of moral impropriety and that one ascertains whether an action exhibits disregard by attending to the harm knowingly done and the worthwhile ends furthered, and weighing them against each other somehow.

Some crucial facts about disregard arise from corresponding facts about value: the same action can be performed for different ends, and it is a condition of the value of some ends that we pursue or value them. What this means is that the value that is furthered by a particular action can depend on the motives and values of the agent. And this in turn means that whether an agent has exhibited disregard can be sensitive merely to the motives and values of that agent. Consider: Fred is the last collector of rare petrological-fungal formations. He is the only person who can appreciate their beauty, so the formations have no value to anyone else. Fred's most prized formation, however, is presently held by someone leaving the country, who knows that to Fred it is absolutely priceless. The only way Fred can acquire it, and fulfill his lifelong dream, is to evict Lucy and get added rent before the formation leaves the country with its owner. Certainly, with the details of Fred's passion and Lucy's plight adjusted in the right way, this might make Fred's doing so intelligible and permissible. Fred may have set the terms of his lease with this very thing in mind. But change the scenario *simply* by altering Fred's values and this may no longer be so. If Fred continued to collect his formations but over time lost interest, now seeing the completion of his collection as about as worthwhile as finishing a crossword puzzle, evicting Lucy may be unforgivably cruel. This shows one way in which the presence of disregard can be sensitive merely to whether and how an agent values the ends they pursue.

The thesis I want to develop is this: that in typical 'paradoxical' cases, blackmail is impermissible because the release of the relevant information exhibits impermissible disregard (provided it is not impermissible for other, more obvious reasons).¹¹ The cases where the release of the information *seems* permissible tend to be ones where the blackmailer has a default entitlement to release the information (e.g. legally), which is trumped by the disregard the release would involve, much as Fred has a default entitlement to switch tenants in his apartment that can sometimes be morally trumped by the harms that would result.

Let me motivate this claim with a pair of cases.

¹¹Berman (1998, 2011) comes to a very similar conclusion about core paradoxical cases of blackmail. One difference between our accounts concerns the nature of the 'weighing' process involved in ascertaining an act for permissibility. Berman often states, and relies on, the claim that only moral reasons can justify action causing foreseeable harm. For example: "an actor *wrongs* another if he knowingly causes him harm without reasonably believing that producing the harm is consistent with the balance of undefeated moral reasons..." (Berman (2011) p.68). The example I've recently given is supposed to clarify that moral reasons aren't always needed to supply adequate justification for doing harm. The point is not entirely incidental, since it may slightly complicate Berman's evidentiary grounds for keeping blackmail illegal.

Case 1A. Alva, a photographer, takes a picture in a public place whose background depicts Bea’s car in the parking lot of a store of ill repute. Bea works at a job where public image is important to company profits. Though Bea’s private life is irrelevant to her job performance, if information about her behavior were made public she would be fired to placate consumers. When Alva considers which photos to include in an upcoming book collection, he finds the photo of no special interest and is largely indifferent as to whether to include it or several others. Alva is about to mark the photo to be included in the collection when he recognizes Bea’s car and realizes that printing the photo would destroy Bea’s career. Alva shows the photo to Bea, explains the situation, and announces “I will print this photo in my upcoming collection, unless I’m paid \$5,000.”

Intuitions about this case and the others I present may depend on details which I haven’t spelled out. But on reasonable ways of filling those in, I take this to be a case where Alva acts impermissibly. Trying to use the photo for profit in this case constitutes an objectionable attempt to manipulate Bea for personal gain. The case is also very much like those said to give rise to the puzzles about blackmail I mean to address. The picture Alva took was made in a public place. We can suppose the store Bea entered was not, for example, a doctor’s office or other establishment associated with protected confidentiality rights. The picture and its reproduction, as a matter of property would then likely belong to Alva.¹² Isn’t it accordingly permissible for Alva to print his photo in his collection? Before I answer this question, consider a contrasting case.

Case 1B As in 1A, except Alva reacts differently to the photo. He sees in the lighting and composition a perfection that is a unique product of chance and idiosyncratic choices he brings to his photography. He thinks this is one of the best photos he has taken—the kind that he always hoped to produce as a photographer, and representative of what he hoped to achieve with his work. He is about to mark the photo to be included in the collection when he recognizes Bea in the shot. He hesitates, recognizing the harm that would come to Bea. Leaving out the picture could be foregoing an opportunity that comes but a few times in one’s life. Alva is willing to make the sacrifice on Bea’s account, but is hoping for *some* compensation. Alva eventually shows the photo to Bea, explains the situation, and announces “I will print this photo in my upcoming collection, unless I’m paid \$5,000”.

On reasonable ways of filling in details, I take this to be a case where Alva behaves permissibly. Indeed, Alva’s behavior on slight elaboration could be

¹²Whether the reproduction of the photograph is legally permissible is, of course, subject to privacy laws which may differ from country to country. The force of the case really shouldn’t turn on a particular set of such laws, and the details of the case can be adjusted to ensure a legal release while preserving my intended point.

considered self-sacrificial and *supererogatory* if he considers the \$5,000 small compensation for his artistic losses. Consider that it is compatible with what I've said that Alva hopes that Bea decides *not* to pay so that he may publish his photo with peace of mind having given Bea a chance to compensate him. If so, a question arises: Alva has made the same announcement, to the same individual, to the release the same information, with equally harmful effects. So why is that announcement impermissible in 1A but permissible, and possibly even supererogatory, in 1B?

The only difference between the cases, by design, is Alva's relationship to the photograph and its release. A tempting way to state how this affects permissibility is by saying that Alva makes a *threat* in 1A, but an *offer* in 1B. This may be correct, but it is uninformative. What makes the former a threat and the latter an offer? And why is the offer, but not the threat, permissible? As already noted, it is a challenging task to distinguish threats from offers. But more importantly, even if we have such an account in hand, we'll still be faced with the challenge of saying how this lines up with questions of permissibility and impermissibility. After all, offers can be impermissible, and threats permissible.

We should set aside the issue of threats and offers, since pinpointing the morally relevant difference between 1A and 1B is simple. This pair of cases is directly analogous to those of Fred *qua* collector. In both pairs of cases, judgments of the permissibility and impermissibility of certain actions track facts about an agent's relationship to the value of the ends they pursue. In Fred's case it's relatively clear how these facts make a difference to his obligations. I want to claim that essentially the same story holds in 1A/1B, and that we can see this by adopting the simple account of wrongful coercion given in *Basic Coercion*.

To appreciate this, let's revisit the question I raised about 1A: It was permissible for Alva to take the picture, and he has ownership over it and its reproduction. So isn't it permissible for Alva to print it? Alva certainly has a legal right to do so. But it is still impermissible for Alva to publish the photograph *once Alva has seen the harm it will do to Bea*. To make this clear, let's set aside the issue of threats by supposing that it never occurs to Alva to use the photograph to extort money from Bea. Recall that in 1A Alva is indifferent as to whether or not to include the photo in his collection. It was stipulated that he needed only the slightest reason to include one of several other photos instead. Imagine in these circumstances Alva discovers the photo could destroy Bea's career. He has no grudge against Bea, no reason to think that she deserves to lose her job, or that the information that leads to that consequence ought to be revealed—no *good* comes of that release. If we keep these facts in view, and Alva nonetheless does not make the minimal effort to substitute a separate photo, it's clear that Alva has done something impermissible. The action here is thus an instance of what I have been calling "impermissible disregard": Alva knowingly causes serious harm to Bea without the harm being offset by the value of the ends that Alva furthers in so doing. The harm, in combination with Alva's relative indifference, trumps antecedent rights Alva had to publish the photo at his discretion.

I noted earlier that assessing an action for disregard involves attending to an agent's motives and values, since these can shape the value of the ends the agent promotes in acting. The value of these ends is weighed against harms in testing if an action exhibits impermissible disregard. Switching from case 1A to 1B gives another example of a change in the value of ends owing to a change in an agent's values. In 1B, Alva attaches a great value to distributing the photograph which happens to picture Bea, and takes distributing it to be part of a worthwhile lifelong goal of creating and sharing objects with that value. As such, the disregard in releasing the information vanishes, offset by the value the distribution of the photo acquires for Alva.

What this teaches us is that intuitions about the permissibility of a threat are directly tracking facts about the disregard of the action threatened. And this means that intuitions about the permissibility of a threat in these cases are, after all, tracking intuitions about the permissibility of carrying out that threat. It is easy to overlook this because disregard is a subtle kind of wrong: it can arise in cases where one has a default discretionary authority to pursue certain actions, and an action may exhibit disregard because of facts about the values of the agent performing that action. This subtlety, I claim, is precisely what accounts for the illusory appeal of the paradox.

So we can explain the wrong involved in 1A—why it constitutes a case of blackmail—by appeal to no more than *Basic Coercion*. Blackmail threatens wrong a victim by impermissibly interfering with their deliberation about whether to proceed without compensating the blackmailer. The impermissible interference comes by burdening that course of action with sanctions that the blackmail victim is entitled to deliberate without.

Many common 'paradoxical' cases of blackmail can be treated in the same way. When it seems permissible for someone to release harmful information, we need to ask what they would gain by doing so, and whether this gain offsets the harm caused by the information's release. When we do this, I claim, we repeatedly find that judgments about the permissibility of threats tracks our judgments about the permissibility of carrying the threat out.

The appeal to disregard does not exhaust my account of blackmail. Sometimes complexities arise from an interaction between disregard and other moral issues. What might seem like the simplest case of blackmail—that involving threats to disclose marital infidelity—turns out to involve such interactions. But before I get to these complex cases in §2, I want to show how my view deals with another challenge facing any account of blackmail.

1.3 The Second Paradox of Blackmail

DeLong (1993) has drawn attention to a second puzzle about of blackmail: sometimes it is illegal (impermissible) to request goods as compensation for withholding information, but it is legal (permissible) to accept an independent offer to withhold the very same information. Why? For example, why might it matter if, in 1A, Bea independently searches out Alva, and offers him \$5,000 not to publish the embarrassing photo?

The account of §§1.1–1.2 addresses the moral version of this puzzle, and the key work is again done by the simple account of wrongful coercion. Recall that according to this account, impermissible coercive announcements are impermissible because of the inappropriate influence they have on deliberation. In making an impermissibly coercive announcement, one attaches sanctions to an option in an agent’s deliberations, when the agent was entitled to deliberate free of such sanctions. This means that the particular kind of wrong involved in blackmail cannot be present without an announced sanction.

This might be easier to see if we first contrast a case where a wrong like that involved in impermissible coercion is made more tangible. Cid is traveling on foot to a neighboring town, and is deciding which road to take: the longer scenic path, or the quicker path. Dan, for no reason but to interfere with Cid, tries to ensure Cid chooses the longer path. There are many ways Dan could do this. Perhaps Dan places an obstacle visibly at the start of the shorter path. Perhaps he only creates the illusion this is so. If Cid is influenced by any of this when deciding what to do, then barring special circumstances Dan will have wronged Cid, probably in a number of ways. How?

The wrong I’m interested in isn’t the wrong that Dan perpetrates by forcing Cid to traverse the obstacle. Cid is wronged even if he doesn’t confront the obstacle because he chooses the longer path. That wrong isn’t just the added costs to Cid’s journey owing to the presence of the real, or illusory, obstacle. If Cid takes the longer path, which turns out to be the better one for unforeseen reasons, Dan’s actions were still impermissibly manipulative. Similarly, the wrong is still present if Cid takes the shorter path only to find he needn’t traverse any obstacle, because it was illusory. In this latter case, the wrong isn’t just the wrong of deception: the same wrong is present whether or not the obstacle really is there. The problem is an undue influence of Dan’s actions on Cid’s deliberations. Moreover if Dan places an obstacle where Cid *can’t see it*, and it doesn’t affect his choices, then the special kind of wrongful effects present in all the cases I’ve just given disappear.

The problem is that Cid has a right to pass on both paths unhindered by Dan and, accordingly, has an equal right to choose as if he will not be so hindered. The simple account says that all wrongful coercion operates by infringing on essentially this right: by influencing deliberation in creating the impression that obstacles attach to acts which they should be free of. This wrong can be perpetrated whether or not the person influenced knows who is influencing them, whether they know the influence is inappropriate, and even whether they know *that* they are being influenced at all.

I bring up this case to stress the following point: there is no way for Dan to recreate this wrong to Cid without interfering with Cid’s thought process. If Dan refrains from such interference, Cid may end up choosing the path that everyone, including Cid, acknowledges is manifestly worse for him. That wouldn’t necessarily place any blame on Dan—and if it did, it would be a different kind of blame than in the cases where Dan directly meddles with Cid’s deliberative process. In the case of impermissible coercion, it’s the announcement (prototypically) which constitutes the way of tampering with another’s deliberation.

So without the announcement, the special wrongs associated with coercion disappear.

I want to be careful to stress that it is the wrong of impermissible *coercion* that disappears, because wrongs of a different kind may surface. For example, when an offer is made to compensate another for withholding information, it may be made by someone who misapprehends their situation. This may place one under an obligation to rectify the misapprehension. We need to understand these potential obligations if we want to understand the variant of 1A where Bea independently offers Alva money to not publish the photo. Is it really permissible for Alva to accept the money? I think our answer here depends on further details. Why is it that Bea is making this offer? What does Alva think Bea is thinking?

To see why these questions matter, revisit Cid and Dan's more concrete case. Suppose Cid is about to choose the shorter path until he sees Dan observing him. Cid eyes Dan suspiciously and then starts off on the longer path. Suppose this gives Dan good reason to think that Cid is making this choice only because he believes Dan has tried to place obstacles in his way on the shorter path. I think it's not unreasonable to hold that Dan may sometimes have an obligation to communicate to Cid that this is not actually so. If not, I think some things might make it more likely that Dan acquires such obligations—for example if he's done things in the past to foster Cid's mistaken belief. If in such a case Dan allows Cid to go through with his choice, he hasn't impermissibly interfered with Cid's deliberation, but he may have perpetrated a similar wrong: allowing Cid to labor under the misapprehension that his options were impermissibly constrained.

Similar things can be said about Alva and Bea. Suppose Bea believes that Alva plans to publish the photo expressly, and maliciously, for the purpose of defaming her. If Alva knows this, and has no such intention, and no special reason to publish the photo, he might be under an obligation to reveal this to Bea before accepting any money. This seems especially likely if Alva has done anything to foster Bea's misapprehension.

Bea could have other kinds of mistaken belief as well. Perhaps Bea mistakenly believes that Alva is in a situation more like that in 1B, and is aiming to compensate him for the gains to be had by releasing the photo. This would be akin to a case where Cid mistakenly thinks that Dan has some legitimate reason for creating an obstruction on the shorter path. Again, on some elaborations of this scenario, it's not unthinkable that Dan acquires obligations to correct Cid's misapprehension. Similarly, Alva may be under an obligation to divulge that he doesn't stand to gain from publishing the photo, and would refrain for less.¹³

I have no special commitments as to when the kinds of obligations I've just

¹³These last obligations may be very easily defeated. They involve the kind of information that it seems permissible to hide in bargaining transactions. If we see Alva's acquisition of money in this light, then Alva may be free to demand more money (given the offer) than the photo is actually worth to him, by fostering the illusion that the photo is of independent value. The topic of what kinds of information we owe to one another in bargaining situations is a vexed topic that I'll have to leave aside.

been discussing actually arise or are defeated. I only want to note that the presence of such wrongs is compatible with my primary point: on the account I've given about the impermissibility of blackmail, we understand why a sense of grievous wrong disappears when goods that would have been received after a threat are instead exchanged after an independent offer. In such cases there is no instance of wrongful coercion. This leaves open that there may sometimes be subtler, and generally less severe, wrongs associated with failures to properly disclose pertinent information. And this seems true to the relevant cases.

A final comment on a third puzzle: why is blackmail wrong even if sometimes people *prefer* to be blackmailed.¹⁴ In our original case, Bea may sometimes wish that Alva comes to her asking for compensation when he does not. How is Bea harmed by an outcome that she longs for? Why is there a moral proscription against her wish being fulfilled?

The account of §1.2 shows us what's confusing about this question. In cases like that of Alva and Bea, one of two situations holds. Either Alva's informing won't exhibit disregard, in which case Bea isn't being harmed, but helped by the request for compensation. Accordingly, there is no mystery about why Bea can wish Alva gives her the opportunity to compensate him. But if, by contrast, Alva's informing does exhibit impermissible disregard, what Bea doubtless wishes for is that Alva behave a particular way *given* that he is already disposed to wrong her. It's no surprise that one could wish for something that would be a wrong to oneself in this kind of way. You might, for example, wish of the serial killer that he merely tortures, but does not kill you. One makes this wish only while ceasing to hope that the serial killer mends his ways entirely. Similarly if Bea wishes Alva to come to her for compensation, knowing he has no justifiable reason to release the information, it is because she is worried that he will wrongfully release the information regardless. As such, she hopes he brings about the lesser harm of extorting money from her. Of course if she let her hopes range more freely, she'd wish for him to destroy the photo at no cost.

In some cases we may not know whether we are dealing with someone prone to exhibit impermissible disregard for our well-being or not. We simply worry about the harmful information getting out and want confirmation that it won't. But there's no reason to think that when we hope that we are given the opportunity pay for this confirmation we aren't merely hoping for less than we were already owed.

2 Complicating Factors

Though the morality of many cases of blackmail can be treated using the distinctions drawn in §1, others are more complex. In this section I want to consider three groups of such cases:

- (i) threats to release information one is obligated to disclose,

¹⁴The point is often stressed by libertarians who favor legalizing blackmail. See, e.g., Block et al. (2000) p.595.

- (ii) threats to release information as punishment, and
- (iii) threats that constitute permissible ‘bluffs’.

2.1 Threats to do One’s Duty

Until now I’ve focused on cases where (bracketing impermissible disregard), threats or offers were issued concerning information that it was *merely* permissible to release (i.e., it seemed equally permissible to withhold it). This leaves out two extreme cases: those where it is obligatory to release the information, and those where it is impermissible to do so. The latter are easy to treat. If I steal information that belongs to you and threaten to release it to your detriment, the simple account captures perfectly well why this would be wrong. But what of the former cases? The simple account might seem to predict that it is never impermissible to threaten to release information that one has antecedent obligation to release—and this may seem incorrect.

Case 2A. In investigating the crime scene of a theft Earl, working for the police, comes across a finger print which he identifies as that of Finn. Instead of bringing these facts to light, Earl confronts Finn with the information, and demands that Finn pay him in exchange for staying quiet.

Many have the intuition that this case involves impermissible coercion or a similar wrong. Can the account I’ve developed answer such intuitions?

Yes, though seeing why requires careful work. After all, the case seemingly involves several distinct wrongs we need to separate out. First, there is a wrong connected with the Earl’s claim that, given compensation, he won’t release information he should. We see this brought out more clearly in the following variant.

Case 2B. In investigating a crime scene Earl, working for the police, comes across a finger print which he identifies as that of Finn—a close friend. Earl decides to cover up the information for Finn’s sake. The finger-print is already recorded in the police’s database, but Earl can erase that information in a traceless way using expensive electronic equipment. Earl confronts Finn, asking him to shoulder the cost of the equipment needed to clear him of suspicion.

On reasonable ways of filling in the details of 2B, what Earl does is impermissible, but there is no sense that he coerces Finn. What Earl seems to be doing is *expanding* rather than limiting Finn’s options. Finn’s actions seem problematic because, and only because, it is wrong to countenance withholding the relevant information that ought to be released. This wrong is no wrong *to Finn*.

2B reveals that if we want to account for the intuitions in 2A, we need to separate out wrongs to Finn from other wrongs. The case thus draws attention to a tempting move on behalf of a defender of *Basic Coercion* that I think should

be avoided. That move is simply to say that all intuitions about wrongdoing in 2A are supplied by *Basic Coercion*. The problem isn't that *Basic Coercion* fails to apply to the case. In fact reasoning very similar to that generating *Basic Coercion* may apply in every case of this kind: if someone announces they will do something obligatory only failing compensation, they thereby announce that they will do something impermissible given compensation. If so, a principle like very much like *Basic Coercion* can engage. This might explain both why the threat in 2A and the offer in 2B are impermissible. Intuitively, though, there is something problematic about Earl's relationship with Finn in 2A which isn't manifested in 2B, and this requires saying more.¹⁵

We can, as always, get clearer on the issue by simply removing the distracting issue of threats and offers altogether. Consider the contrast between 3A and 3B.

Case 3A. Earl is a crooked cop, utterly disdainful of the law, with ties to the crime family of which Finn is a member, and comes across the latter's prints at a crime scene. Earl usually works to spare members of the crime family from incrimination, hoping to get favors in return. Even if he didn't want to take up this opportunity, he would typically fail to report the prints out of laziness. But Earl has never liked Finn, so he expressly divulges the information to his superior, taking delight in Finn's eventual arrest and incarceration.

Case 3B. Earl is an upstanding man of the law, finds Finn's prints at a crime scene, and dutifully reports the information to his superior.

Whether Earl is a crooked or lazy cop or not doesn't affect what Earl *ought* to do in 3A and 3B. In both cases he ought to bring the information to his superiors. But there's something insidious about Earl's actions in 3A. They don't seem particularly praiseworthy, and they also seem like a kind of affront to Finn. Finn seems to have loose grounds for resenting Earl in 3A that he lacks in 3B.

It's easy to see why this is the case. In 3A, though Earl has good reasons to report the information to his superiors, he doesn't *treat* them as good reasons, and doesn't *act* on them. Instead, he treats the fact that his action will make Finn suffer as his reason to report the information. This isn't a reason for Earl

¹⁵The temptation to say that 2A is exhaustively explained as a case involving an impermissible offer is heightened if one focuses on the legal, rather than the moral, dimensions of blackmail. The illegality of Earl's actions in 2A plausibly depend on antecedent legal obligations to perform his job. See for example the discussion of "misprison" in Feinberg (1988) pp. 241–245. The potential problems with limiting one's treatment of the case in this way are appreciated by Wertheimer (1987) p.91. Gorr and Berman, who don't make the moral version their primary focus, seem either unable to locate the sense of coercion or unable to properly relate it to the grounds of the threat's impermissibility. Even Berman, who recognizes that part of what is at issue is deplorable motives, still effectively treats the case as one involving an impermissible offer (Berman (1998) pp.861–2). This weakens Berman's case for imposing harsh legal sanctions against ordinary citizens who threaten to report crimes unless compensated.

to report the information. That's the sense in which Finn has loose grounds for resentment: Earl's actions show disregard for Finn *given* Earl's antecedent stance towards issues of legal justice.

Ordinary intuitions track some sense in which the relationship between Earl and Finn is problematic in 3A. This is important because any account of those intuitions will extend to 2A. After all, in 2A Earl is not giving sufficient weight to the moral and legal reasons for divulging the information about Finn. If he were to release the information, it would be for the wrong reasons—effectively to punish Finn. But Earl has no special right to punish Finn (even given Finn's crime). The kind of announcement that Earl makes signals this kind of disregard for Finn's well-being. This, I submit, is the sensed 'wrong' Finn does to Earl.

There are several ways of explaining the distorted relationship of Earl and Finn in both 2A and 3A. One is to develop a 'non-ideal' theory involving relativized notions of permissibility and obligation that track what one (morally) ought to do bracketing certain, perhaps very strong, moral reasons. There's some suspicion that we need something like this to account for ordinary language use of conditional "oughts" (e.g., "What you ought to do is stand alongside Jane and actively help. Given that you're not, you ought at least to steer clear of her altogether"). Alternatively we can introduce a distinction between impermissibility and blame as done by Scanlon (2008), saying that 3A presents us with a case where Earl's actions are permissible (since in the sphere of legitimate action) but blameworthy (because done for objectionable reasons). In the case of 2A, on this view, Earl's actions are just as impermissible as in 2B, but additionally blameworthy because of their grounds.

There might be other ways still. The important point is that accounting for intuitions about 3A will extend to give an account of the 'added wrongs' sensed in 2A, and in threats to release information that ought to be released in general. The end result will be that coercive threats to do what one is obligated to do may be both impermissible and show a morally objectionable disregard for the person threatened. It's just that the moral disregard is not in these cases the *ground* of the threat's impermissibility. This seems to give precisely the correct diagnosis of such cases, especially when we contrast alternatives, like 2B and 3A, where each of the two individual 'wrongs' is factored out independently.

Now that we've gained an understanding of threats to do one's duty, we can finally treat a case of blackmail whose discussion I've postponed—that of the cheating spouse. It almost always seems wrong to ask for compensation for not revealing information about marital infidelity. Why? The answer turns out to be different, depending on the details of the case. There are clear examples where one may have a duty to inform a spouse of marital infidelity, for example when one has a close connection to the wronged party. There may be admittedly rarer cases, depending on prevailing cultural norms, where one may have a duty not to interfere in strangers' business. Demanding compensation for retaining information about infidelities will be impermissible in both cases, but receive different explanations.

Finally there is a third group of cases where it may be discretionary whether to interfere and inform a spouse about infidelity.¹⁶ In these cases, almost always, the only value in informing is connected with providing assistance to the wronged spouse, which is a morally significant value. This feature of the case is important, because to the extent one actually intends on informing, the value of the choice is not one which is appropriately offset by personal monetary gains from the perpetrator of the relevant wrong. In other words, even if promoting a moral end is discretionary, forming an intention (even a conditional one) to pursue that end can place one under additional obligations not to be deterred in certain ways.

To gain an appreciation for this point consider a case where, as always, threats are no longer at issue. Suppose a hunting organization is bribing government officials to hunt endangered species. You may certainly take up this cause, but there is too much injustice in the world to say that you have an obligation to do so. You may at the very least permissibly devote your time to other causes. It may even be permissible to base your decision on financial grounds: it might be prohibitively expensive to deal with the hunters, for example. But suppose you decide to get involved by disseminating information about the hunters nonetheless. A member of the hunting group takes notice of your intentions and offers you money in the hopes of stopping you. It is clear, I think, that accepting this money cannot permissibly be your grounds to relent. Now that you are involved in the issue, and you have taken it up as your concern, only certain considerations count as reasons against it—and payment from the perpetrator of the wrong is ruled out. The problem is not that financial concerns must be regarded as irrelevant to your decision, as we just saw. Nor is it that the money is ‘dirty’ (as in a case where it was stolen). If this individual had come to you beforehand, and offered the money in payment for services or as a gift, you could permissibly accept. It’s the relationship between the money and what it is offered for that is at issue. This is what makes it a bribe that is impermissible to make and to accept.¹⁷

The structure of the case is analogous when one is choosing whether to become involved in someone else’s marriage. If a threatener in our third set of cases isn’t lying, he has an intention (perhaps a conditional one) to inform a wronged spouse. But given this intention, it is impermissible to seek out an offer from the wronging spouse to relent. An announcement to tell unless paid constitutes an impermissible offer to take an impermissible bribe. This

¹⁶These cases have proven the most intractable for accounts of blackmail as a form of coercive wrong. To skirt the difficulties in accounting for them Gorr and Berman try to argue that these third cases effectively never arise, since one is generally either obligated (for Berman: on reflection) to divulge the information or withhold it. Gorr’s case for this surprising conclusion is largely unargued, and part of Berman’s case relies on the claim that only moral reasons can be weighed in assessing permissible harmful information release (Berman (2011) p.71). I’ve already noted the suspect nature of this assumption in n.11.

¹⁷Note that in this case, even a (permissible) conditional intention to bring the hunters to justice unless you acquired certain funds, or something unexpected happened, would still place you under the obligation not to allow the hunters’ money in particular to be the reason you failed to follow through.

is why it is much harder to find a ‘cheating spouse’ case like my 1B, where requesting compensation is permissible, even if one is doing something of value in speaking to the spouse. So in this third kind of case of infidelity we have an explanation of why requesting money for retaining information is impermissible, but that explanation is again different, and more subtle, than those supplied for the previous two kinds. Indeed, in these third cases we don’t really have an instance of wrongful coercion at all.

But isn’t the wrong perpetrated by the announcement an offense to the cheater and not merely to their spouse, in these third cases? And isn’t there a sense that the wrong is coercive? The present theory will account for precisely such a sense of a coercive wrong to the cheater, provided the announcement is made by someone who involves themselves in the marital strife but fails to recognize or respond to the new obligations their involvement generates. (Doubtless this is the most common instance of this kind of announcement.) Such a character is behaving just like Earl in case 2A. Accordingly their threat again exhibits wrongful disregard to the cheater characteristic of coercive threats, but as in 2A this disregard is not the immediate ground of the announcement’s impermissibility. Instead, it is the threatener’s failure to respond to the moral reasons for providing assistance to the wronged spouse, generated by their decision to become involved in that spouse’s affairs.

On this view, it is extremely dangerous to take the case of marital infidelity as the ‘standard’ case of blackmail, even though it may be one of the most commonly occurring instances. The commonality belies a great deal of complexity which requires teasing out the relationship between coercive wrongs and several distinct moral concerns.

2.2 Punishment and Circularity

I’ve now commented on the complete range of statuses that an act of information release (bracketing disregard) could have: merely permissible, impermissible, and obligatory. But challenging issues remain. Consider the following cases.

Case 4A. Gia steals money from Hana. Hana is unable to prove the wrongdoing, but happens to have sensitive information about Gia. Hana announces “Give back the \$100 you stole from me or I’ll tell everyone your secret”.

Case 4B. Gia hasn’t stolen money from or otherwise wronged Hana, though Hana still happens to have the sensitive information. Hana announces “Give me \$100 or I’ll tell everyone your secret”.

It’s easy to fill out the details of 4B to make Hana’s threat impermissible. But some feel the threat in 4A is permissible, or at least a lesser wrong. In general, it seems less underhanded to use sensitive information to right wrongs in general, regardless of whether one is the victim of the original wrongs or not. Can an account which relies on *Basic Coercion* make sense of the idea that blackmail threats could permissibly be used to right wrongs?

Trying to understand this phenomenon touches on an interesting issue for my account of wrongful coercion. An instance of my core principle, *Basic Coercion*, can be derived from *Simple Account* and *Permissible Deliberation* only with an added assumption: that the addressee of a threat is morally permitted to deliberate about the actions to which a threat attaches sanctions. We've so far considered cases where that assumption is uncontested. But in 4A we enter on more controversial territory. Is Gia permitted to deliberate about keeping Hana's money? How my account is applied here will depend on our answer to this question, and accordingly on the details of an independent account of permissible deliberation.

I won't take a stand on these issues in this paper, since a key virtue of my account is its ability to integrate with different accounts of deliberative permissions. For now, let me explore the possibility that Gia is entitled to deliberate with her impermissible choice of retaining the stolen money, since this will lead to the most interesting and challenging discussion of 4A. In the next section, we'll return to the question of whether this assumption is plausible, and how my account interacts with questions of deliberative permissions.

If Gia is permitted to deliberate with the option of retaining the money, we will be able to derive an instance of *Basic Coercion* and the question will arise as to how, if ever, this could be consistent with the permissibility of Hana's threat. At first accounting for this possibility seems easy. We can continue to understand this phenomenon within the simple account of wrongful coercion, by claiming that those who see the threat in 4A as permissible are tracking the value connected with enforcing a just punishment. After all, if we suppose that Gia *doesn't* return the money in 4A, then to the extent it seemed permissible for Hana to threaten the release of information, it likewise seems permissible for Hana to carry out that threat as a form of retaliation. In 4B, where the threat seems impermissible, enforcing the threat likewise seems inappropriate as a retaliatory measure. Viewing the release of information as a form of punishment helps understand why *Basic Coercion* could yield different verdicts in 4A and 4B: in 4A Hana may have a legitimate aim of punishing Gia absent in 4B. And without the aim of enforcing a just punishment, Hana's telling the secret would exhibit impermissible disregard for Gia's well being.

So far so good. But there is a serious further complication. We can't fully subsume these cases under *Basic Coercion* until we account for a crucial detail: part of what seems to make the threatened release of information appropriate is that the threat was made in the first place. It's less clear that Hana can release information about Gia without any announcement, out of the blue, as a form of 'punishment'. The permissibility of punishment often depends on an awareness that the punishment would be enforced.

This immediately raises a circularity worry. The permissibility of the threat, according to *Basic Coercion* depends on the permissibility of what is threatened. But if the permissibility of what is threatened, *qua* punishment, turns on the permissibility of the threat, there seems to be no non-circular way of explaining the permissibility or impermissibility of threats to punish at all. So can we get along only with *Basic Coercion*?

This worry poses a serious challenge to the simple account because the key conditional that threatens to generate circularity is correct: *if* the permissibility of retaliation turns on the permissibility of the threat, then vicious circularity will arise in *Basic Coercion*'s application. Consequently, to save the simple account we must deny the antecedent of this conditional. We've seen is that the permissibility of retaliation turns straightforwardly on a threat *being made*. But note the troublesome antecedent claims more: that the permissibility of the threat depends not only on the *existence* of the threat, but its *permissibility*.

Ideally, to defend *Basic Coercion* we would give a direct counterexample to the claim that the permissibility of executing threats can only ever depend on the existence of a threat but not its permissibility. To do that, we would have to find a case of a threat that is impermissible, but whose execution is permissible. Unfortunately, if *Basic Coercion* is the main way to test for impermissible coercive announcements, this is precisely the kind of case that should be very difficult to find. Instead of going that route, though, we can give an 'indirect' counterexample by examining cases where an impermissible threat to do one action can actually render permissible a *different* act of retaliation. Consider the following case:

Case 5. Ida is a child old enough to go out on her own unless, of course, her parents say otherwise. Ida wants to go out tonight with Jan, who her parents don't approve of. Consider two possible things her parents could announce:

- (A) "If you go out with Jan while we're away, you'll be grounded."
- (B) "If you go out with Jan while we're away, we'll burn every one of your most cherished possessions in front of you *and* (to top it off) you'll be grounded."

Let's suppose that grounding is a fair punishment for mild disobedience, but burning cherished possessions is over the top. Since Ida is entitled to go out before her parent's announcement, the simple account of wrongful coercion predicts correctly that the announcement of 5B is impermissible and that (barring other special explanations of wrongful coercion besides *Basic Coercion*) that of 5A is permissible. But suppose that in 5B Ida throws caution to the wind and goes out with Jan anyway. Suppose further that her parents come to their senses and realize that they can't follow through on their punishment as stated. Instead, they merely ground Ida on her return. Is this permissible? I think so (though the parents probably also owe Ida an apology for their reckless announcement).

This case is of interest because, by stipulation, Ida couldn't be punished in any way for going out with Jan without *some* announcement. But if what I say about case 5B is right, then she can be punished even after a manifestly impermissible announcement. Why? The case brings out, I think, that threats play the role of legitimating retaliation by giving 'fair warning' that retaliation of a certain *kind* is on the way—perhaps a certain action type or retaliation of a certain severity. To give a warning of this kind just requires that the person

threatened (i) realizes that this kind of retaliation may occur and (ii) is given an opportunity to avoid that retaliation. To play these two roles, nothing requires a threat to be *permissible*. Impermissible threats—especially those that threaten more than would be permissible—can still fulfill roles (i) and (ii). That’s why in 5B Ida’s parents can still punish Ida despite having gone overboard with their threat.

Examining case 5 dispels the circularity worry recently broached. The legitimating power of threats doesn’t turn on their permissibility. This makes it possible to give the following description of the permissibility of the threat in 4A: Hana’s releasing the information in 4A would be permissible only because Hana antecedently threatened it. But Hana’s threat to release the information is permissible only because it was permissible for Hana to release the information *with the threat to release it having been given*. There is a loose circularity here, but it never leads to a circular account of the permissibility of Hana’s actions.¹⁸

2.3 Permissible Bluffs

In the previous subsection I noted that my account so far is neutral on the question of how to treat threats aiming to influence wrongful choices. There are different, incompatible, extensions of my account that pronounce on these questions—how should we choose? The following case of a coercive “bluff” helps reveal what is at stake in our choice.

Case 6. Kat catches Lya vandalizing her property, and knows that the authorities will arrive too late to save the property or to have evidence to hold Lya responsible. Kat knows that if she tells Lya’s abusive husband of the event he will believe Kat, and injure Lya to punish her. Kat judges it impermissible to do this. Nonetheless, she tries to ‘bluff’ Lya into stopping, by announcing that she will tell Lya’s husband if she doesn’t stop (with no intention of carrying the threat out).

Some think that this kind of bluffing is ‘fair game’. Others find bluffing morally objectionable. What claims would be required, in conjunction with my current view, to supply each of these verdicts?

One natural way of extending my account is by accepting additional principles governing impermissible deliberation. For example, the following principle doesn’t seem implausible.

Impermissible Deliberation: No person is entitled to deliberate about the option of performing any action *A* that would involve wronging someone.

¹⁸My view presupposes the possibility of giving a theory of morally permissible punishment independently of an account of the morality of coercive threats. Given the avoidance of the circularity challenge, this will be satisfied by the majority of retributive and deterrence theories of punishment. But the presupposition is not entirely uncontroversial. Quinn (1985), who attempts to derive rights to punish from antecedent rights to threaten as a form of self-defense, is a noteworthy detractor.

The basic motivation for this principle is the fact that we're not entitled to wrong others and we shouldn't be entitled to deliberate about options that we are not entitled to actually perform. Accordingly we're not entitled to deliberate about options that involve wronging others.¹⁹

If we accept this principle, and *Basic Coercion* is the only principle governing the existence of wrongful coercion, then my account will predict that bluffs *never* constitute cases of impermissible coercion provided they only attach sanctions to impermissible behavior. To see why consider case 6 again. Kat is not entitled to deliberate about any actions on which Kat intentionally wrongs Lya. This includes any option, regardless of the outcome, on which Kat vandalizes Lya's property. When Lya makes her announcement she interferes with Kat's deliberation. But, given *Impermissible Deliberation*, she is not doing so impermissibly: she's attaching a sanction to an action of Kat's which Kat is not entitled to deliberate about *at all*. Of course this doesn't mean that Lya is allowed to follow through on her threat! Not only is Lya is not allowed to do so, but by *Impermissible Deliberation* Lya isn't even entitled to *countenance* doing so. The claim is only that Lya is not wrongfully coercing Kat if all Lya does is make an impermissible option in deliberation seem less attractive, however Lya manages to do this.

If we accept *Impermissible Deliberation* and the exhaustivity of *Basic Coercion*, then all bluffs to prevent wrongdoing can be treated this way. But even adopting these principles doesn't necessarily ensure such bluffing never involves any wrongdoing, only that it doesn't involve coercive wrongs. After all, bluffing by its nature involves deception. Perhaps only certain kinds of deception are permitted to right wrongs, so that some bluffs, but not others, are exempted from proscriptions against deceptive announcements. Also, deception comes with moral risks. An announced intention invokes trust, in the way that all assertions do, and if an addressee acts on that trust in detrimental ways, the person making the announcement might be held responsible. So there are many reasons to think that some bluffs to prevent wrongs might themselves be wrong, even if none are wrongfully coercive.

If one isn't comfortable with this outcome, this is no problem. The simple account can again be extended in a different direction to treat threats like Kat's as wrongfully coercive. Instead of placing added constraints on what we are free to deliberate about, we can expand the range of threats which count as impermissibly coercive:

Simple Account⁺: A coercive announcement is wrongful if it aims to influence an agent's deliberation over whether or not to do *X* by threatening to do something impermissible depending on whether or not the agent

¹⁹Deliberation about wrongdoing might be important to the significance of certain good choices, and may even be an epistemic prerequisite to recognizing choices as wrong. To avoid making such deliberation a form of wrongdoing, *Impermissible Deliberation* should probably be read to claim that we have obligations to stop treating an action as a live option as soon as we ascertain that it would be morally impermissible. This would still lead to the permissibility of coercive bluffs, while allowing that one could provisionally entertain an impermissible option (perhaps for some time), before ruling it out, without thereby having engaged in wrongdoing.

does X .

Simple Account⁺ involves a strengthening of *Simple Account* and *Permissible Deliberation*. The latter two principles leave open that perhaps one can attach impermissible sanctions to impermissible actions permissibly, as just discussed. By contrast, *Simple Account*⁺ bypasses consideration of entitlements to deliberation (something which would be appropriate if entitlements to deliberation were universal). Accordingly, it dictates that one can never attach impermissible sanctions to any actions. This would ensure that bluffs of the kind that Lya makes are in fact impermissibly *coercive*.

The neutrality of the simple account is one of its virtues. It can be coherently extended in different ways to cope with corresponding intuitions about coercive bluffs. Moreover we only need the bare structure of the simple account to cope with a vast and diverse range of coercive threats of information release.

3 Broader Implications

I've argued that attention to moral disregard, and its interaction with other moral principles, can successfully resolve the moral paradoxes of blackmail. I now want to turn from defending my account to examine two of its consequences: first, for theories of coercion more generally, and second for the legal paradoxes of blackmail.

3.1 Blackmail and the Simple Account of Impermissible Coercion

Blackmail involves (i) a threat to release information in (ii) an act of coercion. Accounts of blackmail accordingly focus on (i) or (ii) to try to explain blackmail's unique features. It is tempting to fixate on the involvement of sensitive information, and to try to use this in unraveling the difficulties that prohibitions of blackmail present. For example, Murphy (1980) argues that blackmail is illegal is to protect privacy, as the legality of blackmail would ultimately create a harmful market for private information. Owens (1988) argues that blackmail is impermissible in part because it involves an 'unrenderable' service: blackmailers offer to keep certain information from getting out, but prototypically they cannot ensure this, leaving the victim at the mercy of further blackmailers offering the same 'good'.

In contrast, the account I've given focuses entirely on the involvement of coercive threats. On this account there is nothing about information (beyond the fact that its release can be harmful) that figures in an account of what makes blackmail wrong. Such an account has two virtues. First, it avoids the problem of explaining what about information makes it governed by more special duties than other property. Second, subsuming an account of blackmail under a more general account of wrongful coercion strengthens both by giving the latter an attractive simplicity and generality. Moreover, the tools we develop in coping

with blackmail can then be reused in accounting other instances of wrongful coercion.

After all the *moral* paradoxes of blackmail aren't obviously particular to blackmail.

Case 7. Moe, a real estate developer, is buying up land to make a golf course. He needs Nora's property to do so, but she has no interest in selling. A standing law dictates that if one of two neighbors wants to build a fence between neighboring properties, the first neighbor may force the second to share the costs. The fencing around Nora's land would be expensive, and a nuisance since it would block her beautiful views. Moe announces "If you don't sell to me, I'll buy up all the property around you and sue to ensure you pay half the costs of the fencing. Does that make selling worth your while?"

Moe's threat is morally impermissible. But isn't it morally permissible for Moe to buy the land, and to build a fence around it if he wants? And if, before any threat, Nora comes to Moe to sell her land because the surrounding land is being bought up for a golf course, can't Moe permissibly accept? This case presents the same kinds of puzzles as the traditional *moral* paradoxes of blackmail. And because it presents them for the same reasons, all the work done within the framework afforded by the simple account of wrongful coercion resolves these puzzles as well.

Though these cases are identical from the moral perspective, from the legal perspective they are not. Moe's threat is typically sanctioned by the law. Why the legal asymmetry between coercive uses of property and information? This is a very good question.

3.2 The Legality of Blackmail

So far I've only focused on the moral versions of the paradoxes of blackmail and have had to forego a discussion of the vast legal literature on the topic. Without delving too deeply into the details of the legal debates, though, we can say something general about the importance that unraveling the moral paradoxes might have for the corresponding legal puzzles.

Broadly, resolutions of the legal paradoxes of blackmail fall into two categories: *internal* and *external*.²⁰ Internal theories look to something inherent to the act of blackmail itself—the threat, the exchange of goods—to explain the legal puzzle. External theories look to the harmful external consequences of allowing blackmail to transpire unfettered.

Two examples of external theories are given by Nozick and Epstein.²¹ Nozick claims that blackmail is prohibited because it is 'unproductive'. Normally

²⁰The terminology is from Wertheimer (1987) pp.92–103, but the broad form of classification is also used to great critical effect in Lindgren (1984) and Berman (2011). I closely follow these discussions here.

²¹Nozick (1974) pp. 84–86, Epstein (1983).

purchasers of goods are better off for the transaction, even if the prices are exploitative. By contrast the victim of blackmail would have been better off had the blackmailer not existed. The problem, at bottom, is a kind of inefficiency. Epstein claims the problem with blackmail is that legalizing it would lead to systematic cases of fraud against third parties—those who are owed the information which is withheld in a successful instance of blackmail. Again, there is nothing inherently wrong with the blackmail threat, just its ‘downstream’ effects.

External theories often result in inadequate extensional classifications of illegal blackmail (e.g. Epstein’s argument only applies when a third party is legally entitled to the relevant information). But they obviously suffer more systematically from an inability to explain why blackmail violates the *rights of the person blackmailed*. Why is the blackmailer the only party punished?

Internal theories sidestep these difficulties by looking to interpersonal violations of rights. But to the extent they try to make sense of the existing laws, they tend to fall prey to the paradoxes of blackmail quite directly. If one takes on board that releasing the information is protected under the law, it becomes extremely challenging to give a convincing explanation of why threats to release the information shouldn’t be protected as well.

The resolution to the moral paradoxes provides some insight into the problems here. The moral puzzles teach us that *if* the illegality of blackmail is grounded primarily in moral concerns, then the law is giving distinct treatments to two kinds of behavior which should on a legal par. The moral basis for banning certain threats to release information is the very same moral basis for banning the release of that information.

This leaves us with two options for addressing the legal puzzles if. First, we could concede that moral concerns ground the illegality of blackmail and go in for a form of legal revisionism: either the forms of information release that are presently legal should, in the special cases in which disregard is involved, be made illegal, or what is presently classified as illegal forms blackmail should be made legal. This option puts my view close to those espoused by libertarians like Block & Gordon (1986), but for very different reasons. It’s no libertarian stance which provides any impetus towards legalizing blackmail. Rather, if moral grounds for illegality are all that is at issue, settling on legality may simply be the most prudent way to give a requisite parallel treatment.

There is a second option. We can allow that other non-moral concerns are at work in the legal setting which make a differential treatment reasonable. An obvious concern here might be enforceability. Forbidding certain kinds of information release that are morally impermissible because they exhibit disregard will require tests for such disregard. As such they will require speculative inquiries into the values and intentions of individuals releasing sensitive information. Perhaps this problem is avoided in cases of blackmail because one’s values are made more transparent when one actively seeks compensation for retaining information.²² If so, perhaps we can capitalize on differences in the

²²This is the position advanced in the ‘evidentiary theory’ explored in Berman (1998, 2011).

availability of information about motives to draw a distinction between threats to release information and acts of information release themselves. I'm not confident this distinction can be drawn, or whether other distinctions might make for an appropriate differential legal treatment. I'll be content for now to set this aside as a separate problem in legal research.

References

- M. Berman (1998). 'The Evidentiary Theory of Blackmail: Taking Motives Seriously'. *U. Chi. L. Rev.* **65**:795–878.
- M. Berman (2002). 'The Normative Functions of Coercion Claims'. *Legal Theory* **8**(1):45–89.
- M. Berman (2011). 'Blackmail'. In J. Deigh & D. Dolinko (eds.), *The Oxford Handbook of Philosophy of Criminal Law*, pp. 37–106. Oxford University Press.
- W. Block & D. Gordon (1986). 'Blackmail, Extortion and Free Speech: A Reply to Posner, Epstein, Nozick and Lindgren'. *Loy. LAL Rev.* **19**:37–54.
- W. Block, et al. (2000). 'The Second Paradox of Blackmail'. *Business Ethics Quarterly* **10**(3):593–622.
- M. Clark (1994). 'There is no Paradox of Blackmail'. *Analysis* **54**(1):54–61.
- S. DeLong (1993). 'Blackmailers, Bribe Takers, and the Second Paradox'. *University of Pennsylvania Law Review* **141**(5):1663–1693.
- R. Epstein (1983). 'Blackmail, Inc.'. *The University of Chicago Law Review* pp. 553–566.
- H. Evans (1990). 'Why Blackmail Should be Banned'. *Philosophy* **65**:89–94.
- J. Feinberg (1988). *The Moral Limits of the Criminal Law: Harmless Wrongdoing*, vol. 4. Oxford University Press.
- D. Ginsburg & P. Shechtman (1992). 'Blackmail: An economic analysis of the law'. *U. Pa. L. Rev.* **141**:1849.
- M. Gorr (1986). 'Toward a theory of coercion'. *Canadian Journal of Philosophy* pp. 383–406.
- M. Gorr (1992). 'Liberalism and the Paradox of Blackmail'. *Philosophy & Public Affairs* **21**(1):43–66.
- J. Lindgren (1984). 'Unraveling the Paradox of Blackmail'. *Columbia Law Review* **84**(3):670–717.
- J. Lindgren (1993). 'Blackmail: An Afterword'. *University of Pennsylvania Law Review* **141**(5):1975–1989.

- J. Murphy (1980). 'Blackmail: a preliminary inquiry'. *The Monist* **63**(2):156–171.
- R. Nozick (1969). 'Coercion'. *Philosophy, Science, and Method: Essays in honor of Ernest Nagel* pp. 440–72.
- R. Nozick (1974). *Anarchy, state, and utopia*. Basic Books.
- D. Owens (1988). 'Should Blackmail be Banned?'. *Philosophy* **63**:501–514.
- J. Pallikkathayil (2011). 'The Possibility of Choice: Three Accounts of the Problem with Coercion'. *Philosophers Imprint* **11**(16).
- W. Quinn (1985). 'The Right to Threaten and the Right to Punish'. *Philosophy & Public Affairs* **14**(4):327–373.
- T. M. Scanlon (2008). *Moral dimensions: Permissibility, Meaning, Blame*. Belknap Press.
- A. Wertheimer (1987). *Coercion*. Princeton University Press.
- G. Williams (1954). 'Blackmail'. *Criminal Law Review* **79**:79–92, 162–72, 240–46.