Sexual Harassment NDAs: Privacy, Complicity, and the Paradox of Blackmail

Scott Altman

*saltman@law.usc.edu

This working paper is hosted by The Berkeley Electronic Press (bepress) and may not be commercially reproduced without the permission of the copyright holder.

https://law.bepress.com/usclwps-lss/305

Copyright ©2019 by the author.
Sexual Harassment NDAs: Privacy, Complicity, and the Paradox of Blackmail

Scott Altman

Abstract

Harvey Weinstein and the #MeToo movement focused public attention on both sexual predation and the non-disclosure agreements (NDAs) that help it to persist. NDAs help repeat perpetrators avoid detection and punishment, increasing the risk of harm to future victims. At the same time, NDAs are thought to have benefits. They protect informational privacy interests of both perpetrators and victims, facilitate dispute settlement, and provide victims with larger settlement awards.

This article offers moral arguments against the supposed virtues of NDAs. Guilty perpetrators are not entitled to informational privacy about their wrongs. It might be thought that NDAs protect perpetrators from excessive social punishment in ways comparable to felony expungement statutes. But this analogy is not persuasive. The privacy interests of people who are falsely accused of wrongdoing are somewhat stronger. But NDAs are an excessively broad way to protect this interest.

While victims have an interest in not being identified, this does not justify taking money in exchange for keeping the perpetrator’s secret. Victims who agree not to disclose in exchange for money differ from victims who remain silent without payment because the former are complicit in the harassers’ wrongdoing and because they sometimes offer to violate a moral duty in exchange for payment. In advancing these claims, I draw on theories about the immorality of blackmail, which seek to explain why it is sometimes wrong to offer silence in exchange for payment. I also provide reasons to doubt that NDAs are important for facilitating dispute resolution and guaranteeing adequate victim compensation.
Sexual Harassment NDAs: Privacy, Complicity, and the Paradox of Blackmail

Scott Altman*
Sexual Harassment NDAs: Privacy, Complicity, and the Paradox of Blackmail

Harvey Weinstein and the #MeToo movement focused public attention on both sexual predation and the non-disclosure agreements (NDAs) that help it to persist. As a result, several states have made sexual harassment NDAs unenforceable, while advocates tried to limit other tactics that facilitate secrecy about harassment, including sealed litigation records and confidential personnel files.

The arguments against sexual harassment NDAs (as well as NDAs for other wrongdoers) are well known. Wrongdoer NDAs impose significant costs on third parties. They help repeat perpetrators avoid detection and punishment, increasing the risk of harm to future victims. Had the initial victim spoken openly, later victims might have been warned, and employers might have taken precautions. NDAs also harm later (or prior) victims by hiding information that might have helped them to recognize wrongdoing or to bring effective lawsuits. This deprives victims of compensation and undermines the specific deterrence function of damage awards. NDAs may also reduce the general deterrence effect of sexual harassment law. If sexual harassers are not publicly penalized for their behavior, other potential harassers may not be deterred. As well, NDAs might exacerbate public harms.

Contracts for secrecy are often useful for purposes unrelated to wrongdoing, for example when businesses negotiate to share trade secrets with suppliers, employees, or potential merger partners. The current controversy over sexual harassment NDAs is hardly the first time the issue has received public attention. See, e.g., Alan Garfield, Promises of Silence: Contract Law and Freedom of Speech, 83 Cornell L. Rev. 261 (1998).


See, e.g., Aditi Bagchi, Other People’s Contracts, 32 Yale J. Reg. 212 (2015). For application of this idea to NDAs, see Hoffman & Lampmann, Hushing Contracts, (working paper 2019).

See, Andrew Daughety & Jennifer Reinganum, Secrecy and Safety, Amer. Econ. Rev. 1074 (2005); Scott Moss, Illuminating Secrecy: A New Economic Analysis of Confidential Settlements, 105 Mich. L.Rev. 867 (2007). Efficiency arguments have sometimes been advanced for wrongdoer NDAs. For example, in a traditional tort suit, if victims of wrongdoing can accurately estimate the value of secrecy to the wrongdoer, they can extract a large payment, which might adequately deter future wrongdoing, even if it also makes future suits by those who are injured less likely (because future victims lack information about the tortfeasor that would have facilitated their lawsuits). Because such deterrence comes without the need for multiple litigations, it might be efficient. See Levmore & Fagan, supra at 315-18. Because sexual harassment victims are in a poor position to evaluate how many other victims a particular perpetrator has harmed, this argument does not apply to most sexual harassment NDAs. Id. At 319-20.
enforcement of important legal rights, secret settlements prevent the public from seeing that justice is done. When used repeatedly by individual companies, NDAs may allow those companies to violate rights with impunity, avoiding the scrutiny that accompanies enforcement of rights in full view.

Wrongdoer NDAs have some benefits. NDAs protect informational privacy interests of both perpetrators and victims. As well, NDAs are thought to facilitate dispute settlement and to provide victims with larger settlement awards.

Proposals to combat the perceived ills of NDAs vary. Outside the context of sexual harassment, settlement secrecy has been opposed, especially for cases involving public safety. For example, many states mandate that medical malpractice settlements be made available in publicly searchable databases that name the doctor involved, describe the underlying allegations, and provide details on settlement terms.

Limits on sexual harassment NDAs have taken several forms. Generally, they declare unenforceable terms that prohibit victims from disclosing underlying allegations. Some permit terms that protect victim anonymity or that maintain secrecy about settlement amounts. A new federal law denies business-expense tax deductions for when sexual harassment settlements that include NDAs. A few proposals would forbid lawyers from drafting NDAs or require lawyers to alert authorities

---

7 For example, the Florida Sunshine in Litigation act forbids any contract that conceals a public hazard. Florida Statutes 69.081(4).
8 Eric Helland & Gia Lee, Bargaining in the Shadow of the Website: Disclosure’s Impact on Medical Malpractice Litigation, 12 Amer. L. and Econ. Rev. 462 (2010)
9 See Christopher Drahozal and Laura Hines, Secret Settlement Restrictions and Unintended Consequences, 54 Kansas. L.Rev. 1457 (2006). Efforts to limit NDAs in settlements involving discrimination other than sex discrimination have not been as widespread. For example, California makes NDAs unenforceable as to sexual harassment, sexual assault, sex discrimination, and various sex offenses. Cal. Code of Civil Proc. §§1001, 1002. For other sorts of discrimination, it permits employers to settle cases with an NDA as long as the agreement settles a dispute and is freely negotiated. Cal. Gov. Code §12964.5.
11 Secrecy for settlement amounts but not for the nature of underlying claims has been called semi-confidential, or translucent. Some scholars regard this as a good compromise for many forms of settlement. See, Levmore & Fagan, Semi-Confidential Settlements and Civil, Criminal, and Sexual Assault Cases, 103 Cornell L.Rev. 311 (2018).
12 Sec 162(q) makes non-deductible any payment related to sexual harassment (including attorney’s fees) if the settlement of the claim includes an NDA. This provision was added to the code in the Tax Cuts and Jobs Act of 2017. A related suggestion was made by Levmore and Fagan, supra at 343.
if one is drafted.\textsuperscript{13} All these restrictions have been declared pointless or even counter-productive by some commentators because NDA bans can easily be circumvented.\textsuperscript{14}

Limits on wrongdoer NDAs are traditionally viewed as involving a difficult tradeoff, forcing the settling parties to sacrifice important interests in order to protect the public.\textsuperscript{15} I am not persuaded by this description. Instead, I suggest that parties to wrongdoer NDAs lack significant, legitimate, privacy interests in making secrecy agreements. For this reason, we should not regard banning them requiring a hard choice that asks private parties to make sacrifices for the public good. Instead, they are, in many cases, regulations to prevent private parties from imposing public harms without good reason.

In advancing this claim, I will consider separately the privacy interests of guilty perpetrators, wrongly accused perpetrators, and victims. Guilty perpetrators lack a strong interest in informational privacy. It might be thought that NDAs protect perpetrators from excessive social punishment in ways comparable to felony expungement statutes. But this analogy, I will argue, is not persuasive. The privacy interests of people who are falsely accused of wrongdoing are somewhat stronger. But NDAs are an excessively broad way to protect this interest.

As to victim privacy interests, I will argue that while victims have a significant interest in not being identified as a victim (or a victim who brought a lawsuit), this does not justify taking money in exchange for promising to keep the perpetrator’s secret. Indeed, some victims who accept money for secrecy lack adequate reasons for keeping silent in circumstances that might endanger others. Victims are often wrong to exchange silence for money, but not necessarily wrong to remain silent for other reasons. Victims who exchange silence for money differ from victims who remain silent without payment because they are complicit in the harassers’ wrongdoing and because they sometimes offer to violate a moral duty in exchange for payment. In advancing these claims, I draw on arguments made about the well-known paradox of blackmail, which asks why it is wrong to seek payment for silence but not wrong to reveal a secret or to be silent without payment. I also provide some reasons to doubt that


\textsuperscript{14} Drahozal and Hines, supra.

NDAs are important for facilitating dispute resolution and guaranteeing adequate victim compensation. Unfortunately, these reasons also suggest that NDA restrictions may be ineffective in expanding publicity about sexual harassers.

I. Interests of Alleged Wrongdoers

Do wrongdoers have a legitimate interest in keeping their wrongs secret—legitimate in the sense that it is morally acceptable for them to take steps to prevent discovery, or that others have a duty to respect or facilitate those efforts, or not to disclose facts about wrongdoing that they happen to know? What about alleged wrongdoers who are falsely accused? For both groups (though in different ways), the most plausible interest they have is in informational privacy.

Various theories explain why informational privacy is valuable. Some of these tie privacy to aspects of human flourishing, such as autonomy or intimacy. But these theories have little application to wrongdoer NDAs. Other theories of informational privacy rely on preventing the inappropriate use of information. If information might be used wrongly by others, then those with secrets reasonably seek privacy in order to protect themselves. This is an increasingly prominent justification for privacy protections. For example, we protect the privacy of genetic information in part because it might be used unjustly as the basis for insurance pricing or employment discrimination. This injustice-prevention account of privacy is largely prophylactic. People are not necessarily harmed by the mere fact that others know their genetic information—so they have no direct interest in informational privacy. Their interest derives from the practical difficulty of preventing information from being misused.

A related theory, sometimes called a right to identity, focuses not on how information is used, but on how information affects the way others view us. Sometimes a single fact about a person will disproportionately affect their reputation, displacing other important aspects of their lives until they are associated with just one fact. For example, even if genetic information never led anyone to discriminate against me, if public disclosure led people to think of me primarily as "that person disposed to

---

16 Ruth Gavison, Privacy and the Limits of Law, 89 Yale L. J. 421 (1980) (Adopting or considering unorthodox views or practices will be inhibited if it must be done in public, especially if social punishment for unorthodox views or behaviors is severe. Privacy thus promotes autonomy by facilitating experimentation and deviation from restrictive social norms.)

17 Julie Innes, Privacy, Intimacy and Isolation (1992) (We treat some aspects of our lives as confidential so that we can reserve them to share exclusively with intimates).

18 For a discussion of this version of privacy, as well as many others, see Daniel Solove, A Taxonomy of Privacy, U. Penn. L.Rev. 477 (2006).
dementia,” I would be harmed by the undue focus on this single fact. Similar concerns underlie the movement for a right to be forgotten.19

Preventing information misuse and identity distortion might be thought to justify NDAs in sexual harassment cases. The arguments for this position vary depending on whether the alleged perpetrator of wrongdoing is in fact guilty.

A. Falsely Accused Defendants

NDAs might protect from unjust social punishment people who are falsely accused. False accusations include both intentionally fabricated claims, and also claims based on misunderstandings, failures of memory, or incorrect interpretations of what constitutes sexual harassment. Additionally, some people who commit serious wrongs are not guilty of everything for which they are accused.

The law rarely uses secrecy to protect people who might be wrongly accused of a crime. We usually do not keep secret the names of people arrested for crimes before they are charged or convicted, and we open criminal trials to the public. Even in civil cases, transparency is the norm. Court filings in civil cases are public records, which are only rarely sealed for adult litigants.20 Admittedly, some recent trends do facilitate secrecy in civil disputes. For example, mandatory employment arbitrations are not open to the public, and most arbitration documents are not public records.21 Similarly, laws sometimes protect the privacy of employment records, including by creating incentives for employers not to give reasons for termination in public.22 On the other hand, there seem to be contrary trends: a few large companies and some law firms have abandoned mandatory arbitration for

21 Many arbitration awards can be found on Lexis, but these mostly redact party names. For class action arbitrations, published versions include party names. See American Arbitration Association, Supplementary Rules for Class Arbitrations 7 (2003). https://www.adr.org/sites/default/files/Supplementary_Rules_for_Class_Arbitrations.pdf. 
22 Payton v. City of Santa Clara, 132 Cal. App. 3d 154 (1982) (finding a prima facie case for privacy violation when employer posted the reasons for plaintiff’s employment termination on a bulletin board). Most employer reasons not to explain employment termination to outsiders seem to derive from concerns about defamation, rather than privacy. But even here, there are fairly strong protections for employer disclosure – including to other employees – under the common interest privilege. E.g. King v. UPS 152 Cal.App.4th 426 (2007). Despite efforts to protect employers who disclose information about employee misbehavior, most employers do not disclose such information, fearing lawsuits either by former employees or by those who hire former employees.
sexual harassment cases. One Catholic archdiocese has released all past sexual abuse victims who signed NDAs from their duty of silence. Some states have enacted laws discouraging employer silence about employee sexual harassment.

Rather than using secrecy to protect wrongly accused people from unjust social punishment, the law uses other strategies. Sometimes it regulates the use of information, rather than facilitating secrecy. For example, some states forbid employers from asking potential employees about arrest records. In other cases, the law addresses misuse through deterrence and compensation. Defamation law aims to protect people against false accusations. But it does not accomplish this task by silencing people directly.

Perhaps we should protect innocent people’s reputations more aggressively, for example by keeping secret the identity of people arrested for crimes if they are not later charged. But allowing accused sexual harassers to enforce NDAs does far more than protect the innocent. It shelters justly accused people as well. Unlike protecting the identity of arrestees who are not charged, which would tend to protect innocent arrestees more often than guilty ones, NDAs protect guilty and innocent alike. Furthermore, unlike arrestees, whose identity can later be revealed if charges are brought, or if they are convicted, NDAs protect the identity of alleged wrongdoers without any means of terminating that protection based on further investigation.

---


25 For example, in California, the privilege against defamation for employers who reveal sexual harassment allegations against their employees was recently strengthened. Cal. Civ. Code 47(C.). As well, the California Public Employees’ Retirement system is currently weighing a policy that would urge companies to disclose settlement payments made to victims of sexual harassment on the behalf of executives and directors. CalPERS weighs push for sexual-harassment corporate disclosure https://www.ocregister.com/2018/04/11/calpers-weighs-push-for-sexual-harassment-corporate-disclosure/. Corporate and securities law might be another avenue for encouraging greater transparency. Daniel Hemel & Dorothy Lund, Sexual Harassment and Corporate Law, 118 Columbia L.Rev. (2018).


27 Sadiq Reza, Privacy and the Criminal Arrestee or Suspect: In Search of A Right, in Need of a Rule, 64 Md. L.Rev. 755 (2005). In some states it is possible to have arrest records sealed to protect the privacy of those who are not charged. See, e.g., Cal. Penal Code 851.91.
We should hesitate to protect wrongly accused defendants if doing so also shelters guilty defendants, especially if the guilty outnumber the wrongly accused. NDAs would be an overbroad solution, akin to preventing defamation by forbidding all derogatory comments, even if true.

Although the percentage of false allegations in sexual harassment cases cannot be known for certain, several considerations suggest that it is not very high. Harassment is significantly under-reported, in part because victims rightly fear retaliation, and in part because reporting can itself be traumatic.\(^{28}\) Given the costs of complaining, one would not expect many fabricated allegations (though to be fair, fabricating an allegation might sometimes be advantageous, such as when an employee fears being fired for some unrelated reason).\(^{29}\) As well, the limited empirical data on false sexual harassment complaints suggests that it occurs in relatively few cases.\(^{30}\) For example, in a survey of Canadian HR directors about sexual harassment allegations in their companies, on average, about 8% of claims were thought to be unfounded.\(^{31}\)

Perhaps these estimates understate the frequency of false claims. Many sexual harassment allegations are found to be inconclusive. Although some inconclusive cases are no-doubt valid claims that cannot be proven, others may be false claims. False claims certainly might occur in more than the 8% estimate noted above. But there is little support for the idea that false claims make up a majority of


\(^{29}\) Popular perceptions are deeply divided along partisan lines about whether false allegations of sexual harassment are a significant problem. See Pew Research Center, Sexual Harassment at Work in the Era of #MeToo, 2018 https://www.pewsocialtrends.org/2018/04/04/sexual-harassment-at-work-in-the-era-of-metoo/.

\(^{30}\) Data on false allegations is limited. Most scholars believe that false allegations of sexual violence occur in between 2% and 10% of cases. See Philip Rumney & Kieran McCartan, Purported False Allegations of Rape, Child Abuse and Non-Sexual Violence: Nature, Characteristics and Implications, 81 J. Crim. L. 497 (2017); Ferguson & Malouff, Assessing Police Classifications of Sexual Assault Reports: A Meta-Analysis of False Reporting Rates, 45 Archives of Sexual Behavior 1185 (2016). For sexual harassment, there is even less research. Although not directly probative of false complaints, data on substantiated (compared to unproven) complaints supports the idea that a large percentage of sexual harassment complaints are well founded. The most recent report on sexual harassment complaints in the US military substantiated 62% of sexual harassment complaints. https://sapr.mil/public/docs/reports/FY17_Annual/Appendix_F_Sexual_Harassment_Data.pdf. There is also a body of scholarship on settling low-value lawsuits (which include, but are not limited to, false allegations). For a review of this literature, see Lucian Bebchuk & Alon Klement, Negative Expected Value Suits in Procedural Law and Economics (Chris Sanchirico, Ed 2012). There is also a literature on the NDAs and low value suits. See Moss, supra; Hush Money, supra.

sexual harassment allegations. For example, one recent survey of employment lawyers found that only 6.5% of respondents thought that most sexual harassment claims were fabricated.\textsuperscript{32}

Of course, false claims—those with no basis in fact—are not the only legitimate concern that might be addressed by NDAs. Even if very few claims are fabricated, and even if most alleged perpetrators engaged in some inappropriate behavior, we might still worry that alleged perpetrators are innocent of some allegations against them. Sometimes there are mitigating facts or mistaken allegations mixed in with correct allegations. NDAs might be thought necessary to protect those guilty of some things from being punished for additional wrongs they did not commit.

Even allowing for reasonable disagreement about how many allegations are false, NDAs probably more often protect guilty perpetrators—at least perpetrators guilty of some wrongs—than the wrongly accused. Does this mean that innocent people accused of sexual harassment must suffer unjust social punishment so that we can more easily detect actual perpetrators? Like those who are arrested for crimes they did not commit, or people who are defamed in other ways, they can proclaim and perhaps try to prove their innocence. But of course, this may only further publicize allegations of wrongdoing.

One strategy that could address this problem was recently proposed by Ian Ayers. He would make sexual harassment NDAs generally enforceable but would terminate them if their beneficiaries prove to be repeat offenders.\textsuperscript{33} His goal—protecting the privacy of defendants who pose no future danger—would benefit some guilty perpetrators, but would also protect those who are falsely accused. Ayers proposes that all sexual harassment NDAs be deposited into an information escrow (perhaps operated by the EEOC) along with details of the underlying allegations. If another complaint were submitted naming the same perpetrator, the first victim would be notified and the NDA would become unenforceable. A similar escrow device is already in use for sexual assault on university campuses.\textsuperscript{34}

Ayers’ suggestion is worth considering, even though it protects first-time guilty harassers in addition to protecting those who are falsely accused. But the idea is problematic. Unless the threat of


\textsuperscript{33} See Ian Ayers, Targeting Repeat Offender NDAs, 71 Stan. L. Rev. Online 76 (2018). Ayers’ article does not specify whether the initial NDA becomes unenforceable. His main suggestion is that receipt of a second complaint would release information from the first complaint for EEOC investigation.

\textsuperscript{34} Callisto allows college students who are sexually assaulted to conditionally report what happened to them. The student describes the event and names the perpetrator. But the report is not sent to the university unless some other student also reports assault by the same perpetrator.
broken escrow deters bad behavior, it would not protect the second victim, but would only become effective as a warning for the third victim. Moreover, because most sexual harassment is not reported at all, such a system might protect serial sexual harassers. There are also practical impediments to its effectiveness—among them that not all sexual harassment settlements are currently submitted to the EEOC, making any transition to a working escrow system difficult.

B. Guilty Defendants

Might NDAs be justified for guilty perpetrators if they were subject to excessive social punishment or if their wrongs inappropriately eclipsed other aspects of their reputations? Assuming a sexual harasser has compensated the victim, perhaps avoiding excessive social punishment or preventing identity distortion justifies protecting even a guilty perpetrator’s privacy.35

The interests in avoiding excessive social punishment or preserving reputation are not often legally protected. Criminal records are usually open to the public. Sex offender registries subject people to exclusion, discrimination, and harassment.36

On the other hand, our culture of social punishment might be thought excessive.37 Criminal record expungement is the best example of protecting wrongdoers’ privacy to prevent excessive social punishment and to allow former convicts to reclaim their identities.38 But criminal expungement differs in many ways from NDAs. These differences suggest that NDAs for sexual harassment are not a suitable means of avoiding unjust punishment. First, expungement does not occur until a criminal sentence has been completed, which implies that a just punishment has been imposed. Perhaps we can equate compensating sexual harassment victims with completing a criminal sentence. But because such

35 In addition to avoiding social punishment, if settling legitimate claims encourages strike suits (suits brought with no chance of succeeding with the hope of settling for less than the defendant’s litigation costs), then avoiding such suits could justify keeping settlements secret. This concern may less important in sexual harassment cases. Most restrictions on sexual harassment NDAs allow secrecy about settlement amount, but not about underlying allegations. Disclosing underlying behavior is thought less likely attract strike suits than is disclosing settlement amounts. Alison Lothes, Quality, not Quantity: An Analysis of Confidential Settlements and Litigants’ Economic Incentives, 154 Penn L.Rev. 433, 462 (2005).
payments are often made by employers, it is unclear that harassers who settle suits have been adequately punished. The fact that harassers often suffer no employment losses also calls into question the just-punishment conclusion.\(^{39}\)

Second, expungement is usually not available to repeat offenders, especially those who re-offend soon after the initial punishment. NDAs allow repeat offenders to hide their repeat-offender status. Unlike expungement, we cannot verify that the beneficiary of an NDA is a first-time offender. Third, the government retains records of expunged crimes, which can be used as the basis for later punishment if the person does re-offend. NDAs prevent anyone from monitoring for future re-offense. Of course, adopting Ayers’ suggestion might address these concerns to some extent.

Fourth, expungement aims to remedy particular social problems: the widespread inability of former prisoners to re-enter society in the wake of mass incarceration and the resulting poverty and recidivism caused by former prisoners’ homelessness and unemployment. These are serious problems, whose harms are disproportionately suffered by poor and minority communities. Excessive social punishments exacerbate the injustice of serving a sentence that itself may have been unreasonably severe.

Nothing about NDAs for sexual harassment resembles this pattern. NDAs do not target for protection groups that have suffered from excessive official punishment. Indeed, official punishment for sexual harassment is not notably excessive. Although data on outcomes for sexual harassers are not widely available, direct punishment seems relatively lenient. The vast majority of sexual harassment is never reported (and so perpetrators are not punished at all).\(^{40}\) When investigations uncover harassment, punishments are often mild. For example, one recent Australian study found that after sexual harassment complaints, 11% of perpetrators resigned and 5% were dismissed by the employer. Most perpetrators received warnings or informal conversations.\(^{41}\) Anecdotal evidence suggests that

---


\(^{40}\) See, e.g., International Bar Association, US Too: Bullying and Sexual Harassment in the Legal Profession 62 (2018) (Worldwide survey of lawyers finding that in 75% of cases, lawyers who were harassed did not report the incident).

even when perpetrators lose their jobs, they seem to be able to find new work because employers do not reveal the reasons that ended prior employment. Perhaps these punishments are well-calibrated to the wrongs, with the most serious wrongdoers receiving the most severe sanctions. But given the frequency of non-reporting, there is very little evidence that formal punishments are systematically excessive.

Apart from official punishment, the social consequences of sexual harassment are themselves not obviously unjust. This last assertion might be questioned by those who worry that the #MeToo movement has gone too far or that sexual harassment law is unduly restrictive. They worry that social punishment is imposed excessively on those who commit minor offenses, rather than reserving severe penalties for the worst offenders. Critics of #MeToo often cite as examples of excessive social punishment Aziz Ansari and Al Franken.

While one can argue that specific social punishments were excessive, or that our workplaces are over-regulated, it seems hard to conclude that social punishment for sexual harassment is systematically too harsh or inattentive to the seriousness of the offense. Among recent well-known perpetrators, those who committed the most serious offenses, such as Harvey Weinstein, are suffering more serious financial, career, and social losses than those accused of less serious things. Al Franken is considering running again for office; Aziz Ansari continues to work as an actor. And even if those who commit mild forms of harassment might have reason to fear excessive punishment, this would not justify enforcing NDAs. Those who commit less serious offenses are less likely to be sued and therefore less likely to

42 See Jessica Clark, If you Fire Someone, supra.

43 The leniency of official punishment might be facilitated by NDAs. Employers who do not want to discipline their harassing employees likely find it easier to avoid imposing punishments because NDAs allow the employer to avoid public pressure. Perhaps without NDAs, employers might err in the opposite direction, imposing unduly harsh punishments such as firing every employee accused of sexual harassment. But current punishments seem so far from excessive, this seem unlikely.


seek an NDA. As a result, NDAs are most likely to protect people who perpetrate more serious wrongs from social sanctions.

What about identity maintenance? Even if harassers are not systematically subject to unjust social punishment, perhaps they suffer unjustly if their reputations are dominated by a single bad act. NDAs might be a reasonable means of avoiding this harm to their identities.

I see two reasons to doubt this justification. First, it is not clear that wrongdoing displaces other aspects of a person’s life in how they are perceived. People accused or convicted of serious wrongs are often remembered not only for those wrongs but also for their accomplishments. Among them are Woody Allen, Bill Cosby, and Roman Polanski. Second, NDAs aim at more than preventing the distortion of their beneficiary’s image; rather they also aim to prevent accurate information from affecting that image, even in ways that are not distorting. In this respect, they overshoot any legitimate image-maintenance goal.

C. Corporate Defendants

The interests of NDA beneficiaries discussed above do not capture all that is at stake in wrongdoer NDAs, which often protect employers in addition to protecting individual harassers. In practice, employers often direct legal strategies and pay for settlements. In some cases, the corporation is defending a valued employee. It almost all cases, the company is a defendant in the lawsuit, charged with failure to detect or prevent the harassment or with retaliation against a complaining victim. Corporate interests in NDAs differ somewhat from the interests of individual perpetrators.

Corporations might be more often innocent of wrongdoing than individuals accused of harassment. This is so for several reasons. First, companies are included as defendants by plaintiffs’ lawyers, who see the corporate defendant as a likely source of money damages, even if the company was not especially culpable. Second, although victims of sexual harassment are in a fairly good position to know if they have been harassed, they are not always in a good position to know whether companies that employed the harasser took adequate precautions to protect them. For both these reasons,

48 Employer liability for harassment by employees sometimes depends on whether a company failed to fulfil obligations. When supervisors harass subordinates and the subordinate suffers an employment consequence, companies are strictly liable for supervisor actions. But when harassment does not produce an employment consequence, companies can defend against charges by showing that they exercised reasonable care and that the employee failed to take advantage of employer-provided corrective opportunities. See Burlington Industries, Inc. v. Ellerth, 524 U.S. 742 (1998); Faragher v. City of Boca Raton, 524 U.S. 775 (1998). When the harassment is
innocent companies may be more prevalent among those who are sued than are innocent alleged harassers. Perhaps this suggests that NDAs to protect companies from the effects of false charges are more warranted.

Despite this difference, there are several reasons not to sympathize with corporate claims that they need NDAs to protect themselves against reputational costs of false allegations. First, sexual harassment charges against an individual can, if they are believed and widely known, lead to job loss and significant impediments to future employment. For corporations, which settle discrimination suits regularly, the stakes are far lower. Except in extraordinary circumstances—such as the Weinstein Company—the public disclosure of settling one or two harassment cases is unlikely to have any significant effect on a company’s reputation, stock price, or profitability. Second, companies sometimes have malevolent motives for wanting NDAs, including a desire to continue employing a valuable employee who is a known harasser. Among the most notorious users of nondisclosure agreements were the Catholic Church in its settlement of sexual abuse cases, and the Weinstein Company to cover up Harvey Weinstein’s predation.

Third, the corporate financial interests that NDAs protect sometimes undermine important public goals. For example, some companies use NDAs to prevent their other employees from learning that someone sued the company over a particular harasser’s behavior. The company believes that copycat lawsuits will follow. Although the company will describe these as strike suits, in some cases, they will be valid claims. Avoiding these suits, while clearly in the company’s financial interest, is not in society’s interest. Publicity about a settlement may embolden other employees with valid claims to come forward, providing them with deserved compensation, and encouraging the company to prevent harassment and to remove serial offenders. As well, some corporate managers want to prevent stockholders from learning about sexual harassment within the company. Managers may want to prop up stock prices or to protect themselves from scrutiny. Insofar as some allegations of corporate malfeasance are correct, NDAs undermine the corporate-governance benefits of shareholder scrutiny. Finally, corporations often use NDAs repeatedly, including them in employment contracts and in settlement agreements. When companies use NDAs in systematic ways, they will sometimes have the effect of circumventing regulation of corporate behavior.

against a coworker or independent contractor, employers are liable if they knew or should have known about the harassment and failed to take corrective action. CFR §1606.11(d).


II. Victim Interests

Crime and tort victims have well-known reasons, often connected to privacy, for keeping secret the details of their victimization—for not reporting the wrong or for being circumspect after making an initial report. Coming forward to complain can be traumatic because it requires that one re-live painful past events. When crimes or torts are spectacular, victims may face lost anonymity. And of course, perpetrators may seek vengeance if victims complain or later publicize the wrongdoing. For sexual assault and harassment victims, the privacy reasons for silence can be particularly acute. Unlike most crime and tort victims, they are often blamed in part for the harms they suffer and often feel shame for having been victims. Sexual harassment victims might suffer career setbacks if employers or potential employers learn that they complained about past harassment.

Victim privacy interests are sometimes respected. The press usually does not name sexual assault or harassment victims without their consent. Failing to report a crime or a tort is not typically illegal for a victim. And although prosecutors can compel victims to testify, often they do not—both because unwilling victims do not make good witnesses, and because prosecutors respect victim reasons for not testifying.

Because victims care about their own privacy, NDAs are usually bilateral, precluding both victim and perpetrator from discussing the wrongdoing or the settlement. NDA bans rightly accommodate victim interests by enforcing NDA terms that prohibit perpetrators from identifying victims.\(^{51}\) Victims’ privacy interests thus need not be compromised when states ban NDA terms that protect perpetrators.

NDA bans might undermine victim privacy if perpetrators were unwilling to accept terms protecting victim privacy without the ability to get comparable confidentiality for themselves. Although there is no available data on this question, perpetrator refusal to promise victim anonymity seems unlikely. Perpetrators who value their own privacy should welcome any indication that the victim wants not to be identified, as it suggests that the victim does not plan to discuss the matter in public. Of course, perpetrators will want to defend themselves if victims discuss their sexual harassment. But a ban on identifying victims would not prevent perpetrators from responding to allegations and offering evidence of innocence.

Although NDA bans likely do not undermine victim anonymity, they might harm victims in several other ways. Without enforceable NDAs, perpetrators might not pay as much to settle disputes, leaving victims financially worse off; and cases might less often settle, which would require victims to litigate – suffering risk, expense, and loss of privacy. The following sections argue that these concerns do not justify enforcing wrongdoer NDAs.

A. Moral Problems with Exchanging Silence for Money

For reasons discussed below, NDA bans might not substantially lower payments to victims. But if NDA bans do reduce victim compensation, this outcome might be less objectionable than NDA supporters suggest. Compensation for silence (as opposed to compensation for the harms of sexual harassment) arguably constitutes an illicit gain, endangering future victims without good reason. Of course, negotiators do not conceive of payments in such discrete terms—some money for compensation and some money for silence. But if victim payments are lower in states that enact NDA bans than in states that do not, the difference is a premium for silence.

By arguing against NDAs, I do not mean to condemn victims who sign NDAs in states that permit them. Such victims likely have little choice; they are told that the only way to settle a case is to sign an NDA. My argument is not that signing an NDA is immoral; it is that the premium offered for the NDA (compared to a settlement offer received when NDAs are not allowed) is morally problematic. Unlike other reasons for victim silence—fear of retaliation or concern for personal trauma—the money paid for silence in an NDA is not an adequate justification for increasing risks to future victims. For this reason, we should not view lower settlement amounts as a reason against NDA bans.

One set of arguments against compensation for silence derives from the literature on blackmail. This literature addresses a puzzle called the paradox of blackmail: Why is blackmail a crime if most blackmailers are legally entitled to divulge or to keep their secrets?52 If the blackmailer is making a

52 This puzzle is often subdivided into two, one moral and the other legal. The moral version asks why it should be immoral to seek payment for silence if one has no moral duty to disclose or to keep silent. Some writers deny that supposed cases of this moral puzzle are real, claiming that in most cases the blackmailer has a moral duty either to disclose or not to disclose. The moral wrong of blackmail is in asking to be paid for something they already have a moral duty to do or to refrain from doing. Such theories have faced several challenges. Among them is the problem of explaining why blackmail is coercive and why it is a wrong to the particular victim. See generally, James Shaw, The Morality of Blackmail, 40 Phil. & Pub. Aff. 165 (2012); Stephen Galoob, Coercion, Fraud, and What is Wrong with Blackmail, 22 Legal Theory 22 (2016).
choice that he or she is entitled to make, why is it wrong to take payment as the way to decide? The bargain seems to benefit the target, who prefers to pay for secrecy rather than have the secret known.

Although NDAs differ from blackmail, NDA bans raise a parallel question. Sexual harassment victims are generally free to discuss the harassment or to keep silent. If victims have no legal duty to disclose the facts of their harassment, exactly what is wrong with deciding whether to keep silent based on a payment? Both parties to an NDA regard themselves as better off with the agreement than without it.

One might think there is no puzzle here because victims should always publicize wrongdoing. Despite the sacrifices involved in coming forward—personal trauma or risk of job discrimination—victims ought to report wrongdoing and discuss it openly. Failure to do this is morally wrong because it endangers future victims. The law cannot feasibly force victims to complain about abuse and then publicize that abuse to anyone who will listen. So it bans NDAs as a smaller, practical, step to encourage public disclosure.

I do not think this interpretation is correct. We could punish victims who do not report crimes and torts if we regarded non-reporting as a serious wrong. We could revive the nearly-defunct crime of misprision of felony. Or we could make victims financially liable if they fail to report wrongdoing and the perpetrator later reoffends. That such proposals seem absurd suggests that we immunize victims who remain silent from criminal and tort liability for more than practical reasons. We take seriously the privacy harms that victims risk when coming forward and regard their decisions to remain silent as often excusable.

That the law treats paid victim silence very differently than unpaid silence does not fit easily with the traditional view of NDA bans, which is that they override the important interests of private parties in order to protect future victims. If the law were committed to protecting future victims, despite recognizing that legitimate private interests are being compromised, why would it stop with NDA bans rather than imposing broader duties for victims to speak? One plausible answer is that most

---

53 NDAs of course differ from blackmail. Notably, the transaction is likely initiated by the payer – with the sexual harasser seeking to pay for silence, rather than being asked by the victim for this payment. On some accounts, this shows the bargain not to be coercive. Shaw, supra at 176. In the blackmail literature, the differential treatment between blackmail threats (which are illegal) and bribery for silence (which often is not) has been regarded as itself puzzling. See Sidney DeLong, Blackmailers, Bribe Takers, and the Second Paradox, 141 U. Penn. L.Rev. 1663 (1993).

harassment victims who remain silent are regarded as justified, or at least excused, but that victims who take payment for silence are not.55

Actions taken for pay are sometimes morally different from comparable actions that do not involve payment. Indeed, we sometimes regard them as different acts. The most obvious example is bribery. Admitting a student to a university-based on a bribe is wrong even if it would have been reasonable to admit the student absent the bribe. The justification for distinguishing these actions is easy to see; allowing money to influence admissions decisions is widely seen as undermining a fair process. We need a parallel explanation for why the law might treat victim silence in exchange for money as a fundamentally different act than silence for other reasons.

Why might we regard victims who accept compensation for silence as less justified than victims who remain silent for other reasons? In part drawing on discussions in the blackmail literature, I offer several answers.

By agreeing to an NDA, victims might reveal that the usual justifications for silence do not apply to them.56 Generally, harassment victims need not come forward because we respect the important privacy rights mentioned above. If someone signs an NDA, they are apparently only willing to keep a secret if paid, which provides evidence that those privacy considerations did not weigh heavily on them. Lacking a privacy-based justification for secrecy,57 the victim has a moral duty to disclose. For them, the

55 Another possible answer is that the legal system seeks to avoid complicity with wrongdoing. The law does not want to be the tool that facilitates harm to future victims—so it denies contract enforcement. But it does not mind purely private action to that outcome, such as victim silence.

56 This account draws on a theory of blackmail first proposed by Mitchel Berman. Mitchell Berman, The Evidentiary Theory of Blackmail: Taking Motives Seriously, 65 Chicago L.Rev. 795 (1998); Mitchell Berman, Blackmail, 34 The Oxford Handbook of Philosophy of Criminal Law (Deigh & Dolinko, Eds 2011). He says that the wrongness of disclosing information sometimes depends on the discloser’s reasons for deciding. Typically, we cannot know enough about underlying reasons to warrant legally requiring disclosure or nondisclosure. Any simple rule about the matter would forbid too many of the morally acceptable decisions. But someone who is willing to forego disclosure if paid will regularly be the sort whose disclosure or nondisclosure would have been immoral. A parallel argument is made in Shaw, The Morality of Blackmail, supra. He argues that blackmailers frequently show impermissible disregard for the interests of their victims, while often those who choose to disclose have adequate reasons for disclosure. See also, Ram Rivlin, Blackmail, Subjectivity and Culpability, 28 Canadian Journal of Law & Jurisprudence 399 (2015); Peter Westen, Why the Paradox of Blackmail is so Hard to Resolve, 9 Ohio St. J. Crim. L. 585, 626 (2012).

57 Some scholars criticize the evidentiary theory of blackmail on the ground that actions ought to be evaluated based on justifications rather than on motive. See Rivlin, supra and Westen, supra. This concern might not be relevant to the analysis in text. As Westen notes in discussing threats to reveal the commission of a crime unless paid, one need not inquire into the motivation of someone willing to keep this secret. We can reasonably conclude that they lack key justifications for secrecy—such as fear of retaliation—from the fact that they would be willing to disclose if not paid. Westen supra at 607.
NDA is an illicit offer to violate a moral duty in exchange for payment. In this respect, paying someone to keep silent about wrongdoing amounts to an improper bribe—improper because it offers money to breach a moral obligation. Both payers and recipients of such bribes act corruptly. By contrast, those who remain silent without an NDA likely do so for privacy reasons that could excuse imposing risks on others. Agreeing to an NDA thus provides evidence that the person lacks good reason for silence.

This evidentiary theory faces at least two problems. First, it does not explain why the desire for money is a worse justification for keeping a sexual harasser’s secret than other, allegedly good, reasons, such as victim privacy. Harassment victims might remain silent because they do not want to lose a job, miss out on a promotion, or have trouble finding a new job. These harms, though not only financial, are significantly tied to money. Why is money an acceptable reason for not disclosing sexual harassment when tied to privacy, but not otherwise?

Perhaps money is a problematic reason for imposing risks when it is disconnected from preventing a wrong. Sexual harassment victims are entitled to be put in a position where they are as nearly as well off as had they not been harassed (including not having to fear losing job prospects because they complained about harassment). Harassment victims who remain silent to protect a current or future job act to protect themselves against wrongful discrimination. Imposing risks on other people for private gain might be worse than imposing similar risks to protect oneself from injustice. Harassment victims who sign NDAs might be understood to seek benefits beyond compensation.

That we may more appropriately impose risks on others to avoid wrongs than to make profits may seem clear by considering an example unrelated to NDAs. If I pull to the side of a busy highway to exchange insurance information with the driver of a car that rear-ended me, I might increase the risk that another driver will have an accident. If I am miles from the nearest exit, I will be justified in imposing this risk (and may be required by law to do so) to ensure that I can be compensated by the driver who hit me. But if I pull to the side of a busy highway so that I can negotiate a business contract before the fiscal year ends, which will ensure a good bonus, I will not be justified.

---

58 There is debate in the blackmail literature on what sort of wrong is revealed by a willingness to accept money in exchange for keeping a secret. Berman suggests that only moral reasons (not self-interest) can justify harming someone by revealing a secret. **Blackmail** supra at 68. Shaw disagrees, describing the issue as whether the blackmailer shows impermissible disregard others’ interests. The question is not whether the blackmailer aims at self-interest, but whether the blackmailer is pursuing sufficiently valuable ends. **The Morality of Blackmail**, supra. Ram Rivlin suggests that the wrong derives from intending to harm the victim, rather than intending an act that harms the victim as a side effect. **Blackmail, Subjectivity and Culpability** supra.
Advocates for NDAs might accept this position, but deny that NDA payments represent private gain. They can point out that settlements are frequently under-compensatory. The only way for harassment victims to prevent the injustice of under-compensation is to exchange silence for money. In this respect, endangering future victims by exchanging silence for money resembles pulling to the side of the road to collect insurance information. It is the only way to ensure adequate compensation for a wrong. As a justification for silence, it is thus comparable to the privacy concerns that justify non-disclosure.

Does the need for full compensation justify accepting money for silence? Perhaps full compensation is simply not important enough to justify imposing risks of sexual harassment on future victims. The privacy interests that justify nondisclosure (avoiding trauma, workplace retaliation, and discrimination) might matter more than increasing compensation from partial to full. Certainly, our laws suggest as much. Workplace retaliation against sexual harassment victims is illegal. This right cannot be waived by contract. But our legal system makes partial compensation the norm. Plaintiffs settle for what their case is worth, given uncertainty at trial and the expense of litigation. Some plaintiffs must pay their own attorneys out of settlement funds. As a result, compensation for plaintiffs in all areas of law is typically less than full. If full compensation were important, we could take steps assure greater compensation without putting future victims at risk, for example by increasing damage awards for harassment claims, making attorney fee provisions more generous, or making damage awards for sexual harassment immune to federal taxation (in a way comparable to physical injury awards).

If under-compensation for victims is an unfortunate (but accepted) part of our legal system, then perhaps victims who sell silence for money are unjustly enriched. Victims put others at risk for inadequate reasons when they seeking greater compensation than our system of justice has generally assured. Alternatively, if seeking full compensation is regarded as an adequate reason for victim silence when full compensation cannot otherwise be achieved, we might ask whether the perpetrator wrongs a

60 In the blackmail literature, this problem is sometimes seen in hypotheticals in which a crime victim threatens to report the crime to the police unless the criminal compensates the victim for the harm done. See Berman, Blackmail at 74-76.

20
victim by refusing to provide full compensation unless the victim violates a duty to others by keeping silent. Perhaps the victim is blameless for accepting this deal, while the perpetrator (and our legal system) are blameworthy for making it necessary.

A second problem for the evidentiary theory is that offering to keep a secret if paid does not actually demonstrate that the victim regarded privacy as unimportant. The evidentiary theory seems to suppose that anyone who signs an NDA would have disclosed information had they not been paid, and that they, therefore, must not have valued their privacy highly. But this is surely not the case. Three sorts of victims might sign an NDA without regarding privacy as unimportant. One group—perhaps including most harassment victims—regards privacy as very important in the immediate aftermath of harassment. Trauma may be at its worst during this period. The fact that victims might later want to discuss the harassment more publicly does not mean that privacy was unimportant when they signed the NDA. We can call this group delayed disclosers. Unless such victims anticipated this change, it is not clear why we would say that their privacy-based reasons for silence were inadequate to justify silence. Perhaps binding themselves to keep silent even after privacy-based concerns no longer weighed so heavily can be understood as breaching a duty. But the argument for this conclusion seems less clear than it would if privacy had been relatively unimportant when signing the NDA.

Another sort of victim who might sign an NDA despite caring deeply about their privacy is a bluffer. Bluffing victims, who would not disclose the harassment, with or without an NDA, might regard their own privacy as extremely important. A final group of victims is ambivalent good citizens. They fear the loss of privacy that would occur were they to make claims public, but would overcome these concerns—out of a sense of duty toward other possible victims—were they unable to negotiate an NDA. Ambivalent good citizens see money as a tie-breaker between significant personal concerns for privacy and a sense of duty. Money is a decisive reason for silence, but it is not the only or most important reason.

If delayed disclosers, bluffers, and ambivalent good citizens care deeply about their privacy, then their signing an NDA does not demonstrate that they lack adequate justifications for silence. Because privacy is viewed as an acceptable reason for silence, they had no duty to disclose. So they are not offering to violate a duty in exchange for money. Many victims likely do not know which of these categories best describes them. Nor can we say with any confidence how many victims fall into each

---

63 See Rivlin, supra at 410.
category. For this reason, the evidentiary theory seems to fail; signing an NDA does not show that the person is violating a duty to disclose.

The only way to salvage the evidentiary theory would be to explain why all victims—those who care little about privacy, delayed disclosers, bluffers, and ambivalent good citizens—should not sign NDAs. The account might begin with the claim that those who care little for privacy breach a duty by keeping silent for money. Perhaps, as suggested above, this argument could be expanded to include delayed disclosers, who breach a duty by binding themselves not to disclose even after privacy reasons no longer justify their silence. But a different explanation would be needed for bluffers and ambivalent good citizens.

Perhaps bluffers’ entitlement to payment for silence can be questioned based on a different theory. Bluffers gain a financial benefit from a bargain that gives the other side little of value. Although they purport to trade silence for money, they are not really trading anything; since the silence was going to be provided even without the payment, the bluffer profits from a kind of fraud. Their fraud might seem unimportant. After all, their wrong is at most a wrong to the harasser. But bluffers might be criticized as gaining unjust enrichment at the expense of future victims.

This critique of bluffers might be questioned on several grounds. First, we do not usually think that bluffing in negotiation is fraudulent. Were I to accept payment for a non-compete agreement when I sell my business, no one would condemn me for taking this money if they learned that I never intended to compete with my old firm. Second, it is unclear that bluffers give perpetrators nothing by promising silence. Very few people know for certain that they will never want to disclose. Even victims who currently plan not to disclose might change their minds (which is to say that the victim might be a delayed discloser). If so, the harasser gains certainty from the NDA even when making an agreement with someone who is bluffing. Third, this account is subject to the challenge that bluffers’ enrichment is not unjust—it is a means of getting full compensation. So the argument against bluffers depends, as did the argument against those who disclose without the NDA, on whether seeking full compensation justifies imposing risks on future victims.

64 Some theories locate the wrong in blackmail in a kind of fraud. Absent a chance to charge for secrecy, the blackmailer would have kept the secret (or not even discovered it). In such cases, the bargain only appears to benefit both parties. In fact, the opportunity to bargain imposes a cost on the payer with no corresponding benefit. See, e.g., Scott Altman, A Patchwork Theory of Blackmail, 141 U. Penn. L.Rev. 1639 (1993).
Evaluating the ambivalent good citizens is not easy. Condemning them along with those who do not value privacy seems wrong. The two groups have in common that absent an NDA they would disclose the harassment. But the ambivalent good citizen may have the kind of strong privacy concerns that would generally justify silence. What makes them ambivalent good citizens is that they have an unusually strong (albeit limited) commitment to the public good. It seems odd to condemn them for being somewhat committed to others’ welfare.

Has any of these arguments—all drawn from the blackmail literature—shown that victims who sign NDAs lack adequate reason for selling silence, or that their reasons for silence are less appropriate than victims who remain silent without an NDA? Two main arguments were considered: (1) some victims who sign NDAs lack sufficient reason for silence because privacy concerns do not weigh heavily on them; and (2) other victims, for whom privacy is very important, had good reason for silence, but lack much complaint if they are denied money, since they are not giving up anything valuable in exchange. Significant objections to these ideas have been noted. Victims are more diverse than this division suggests; they likely change over time and may regard NDA payments as needed to overcome the inadequate compensation they would otherwise be offered for the harms they suffered. These objections may be enough to reject the arguments drawn from the blackmail literature altogether. Nonetheless, one might think there is some truth to the following: all victims fall into one of two categories (though of course neither they nor we can know which victim belongs to which group). Either they would eventually disclose facts about their perpetrator absent an NDA or they would not. For those who would disclose, they have a duty to disclose once the core privacy reasons that lead them to secrecy have declined enough that they would disclose without an NDA. For them, the NDA inhibits them from fulfilling this duty. For the other group, who would never disclose, the NDA payment compensates them for giving up a right that they would not have used. For that reason, it is a windfall. In combination, these two observations suggest that victims who agree to silence in exchange for an NDA are not wrongfully deprived of money by NDA bans. Whether this broad argument is persuasive likely depends on one’s view of NDA compensation as a needed supplement to inadequate compensation for the underlying wrong. Even if seeking full compensation justifies signing an NDA, it does not justify our collective failure to change compensation rules so that NDAs are not necessary. Rather than allowing NDAs, we should ban NDAs and more fully compensate victims. This would encourage broader disclosure and thus reduce sexual harassment risk.
An alternative to theories drawn from the blackmail literature is a theory based on complicity. Perpetrators may be wrong to offer victims money (and victims wrong to accept it) because NDAs make victims complicit in the perpetrator’s wrongdoing. To be effective, this theory must explain why we ban NDAs but do not mandate victim speech, and thus must show why NDA participation makes victims more complicit than those who remain silent for other reasons, including reasons connected to money in other ways.  

As an analogy, consider someone deciding whether to rescue a struggling swimmer despite significant personal risk. If an enemy of the swimmer offers to pay the potential rescuer not to save the victim, accepting the payment would be wrong. This is so even if the decision to rescue is morally optional. The choice not to rescue might be reasonable even if the rescuer decided in part based on money—for example, because the rescuer feared job loss after the rescue made the rescuer late to work. Financial self-interest does not make the rescuer complicit in the same way as financial cooperation with the wrongdoer.

By accepting a wrongdoer’s money as a reason not to rescue, the rescuer becomes complicit in the enemy’s wrongful aims. The rescuer’s complicity would not disappear even if the enemy owed money to the potential rescuer. Imagine that some years earlier the swimmer’s enemy had stolen the potential rescuer’s car and was now offering to compensate the rescuer for the stolen car on condition that the rescuer would not save the victim. That the rescuer was merely trying to rectify a past injustice would not absolve the rescuer of complicity.

Are NDAs comparable? One difference is that because not all harassers are serial harassers, not all NDAs facilitate future wrongs. But anyone who signs an NDA with a harasser risks facilitating such future wrongs. Additionally, NDAs may help perpetrators—even those who will not re-offend—avoid deserved social punishment. Victims who sign NDAs facilitate this wrong. Finally, NDAs sometimes undermine rule-of-law virtues and allow companies that use them systematically to avoid regulation,

---

65 A related point is suggested in the blackmail literature by James Shaw, though Shaw seems to rely on a theory of corruption rather than on complicity. He says “[F]orming an intention (even a conditional one) to pursue [a moral end] can place one under additional obligations not to be deterred in certain ways. . . . Now that you are involved in the issue and 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.” Morality of Blackmail 184-85. One problematic aspect of this theory is that it is unclear that most blackmailers (or NDA recipients) have made any commitment—even conditional—before negotiating for payment. The beach example in text suggests that receiving funds may be wrong even if the person bribed has not yet made a decision about whether to rescue. So the wrongfulness cannot depend entirely on having formed an intention.
perhaps contributing to the prevalence of sexual harassment. Victims who sign NDAs may be complicit in these harms also.

Does complicity distinguish harassment victims who keep harassers’ secrets in exchange for money from victims who are silent for other reasons, such as avoiding trauma or employment discrimination? After all, they all risk facilitating future harassment and evasion of social punishment. And all act for self-interested reasons.

The answer might depend on one’s theory of complicity. One important difference between silence for privacy and silence for money stems from how one benefits from imposing risks on others. In the case of NDAs, the benefit of silence is provided directly by the wrongdoer. To see this, return to the beach example. Imagine that the swimmer’s enemy wants the swimmer to die because the enemy is the beneficiary of the victim’s large life insurance policy. The bribe is effectively drawn from anticipated profits of this crime. This benefit is more tightly tied to the wrongdoing than other benefits that might lead the rescuer to refrain from saving the victim, such as concern for personal safety. In both cases, victims are self-interested. And in both cases, victims might knowingly facilitate the perpetrator’s wrong. But only the payment is tightly tied to that wrong because the wrongdoer provided the funds to advance the wrongful goal.66

The NDA case is not entirely comparable. Funds used to pay the harassment victim do not derive from future harassment. Nor do payers seek victim cooperation only in order to facilitate future wrongdoing; they likely also want to protect their jobs and reputations. Nevertheless, victims who accept payment for silent are participating more extensively with perpetrators than are victims who remain silent to protect their own privacy. If the perpetrators do not plan future wrongdoing, and so are seeking silence only to protect their job and reputations, victims are participating in the perpetrators’ plans to evade social sanctions. If the perpetrator does contemplate more harassment, the victim cooperates with that plan by accepting money that is paid for this purpose.67

Victims who remain silent to protect their own jobs, although they may facilitate future harm (and might be to some degree complicit), participate less directly in that harm. The benefit they seek

---

66 For a comparable example, see Gregory Mellema, *Complicity and Moral Accountability* 107-08 (2016).
67 For discussions of shared purpose and complicity, see Chiara Lepora and Robert Goodin, *On Complicity and Compromise* 73-77 [2013]; *Complicity and Accountability* at 102-115.
from silence—such as avoiding the trauma of reliving the event—is not provided by someone seeking to do future harm or to avoid deserved punishment.68

None of this discussion should distract us from the wrongdoing of perpetrators. If some victims are wrong to accept payment for silence because they do not value privacy highly and therefore have inadequate reason to impose risks on future victims, then surely the perpetrators who seek to buy their silence are also to blame. Similarly, if victims who accept NDAs can be criticized for their complicity with wrongdoers, this means that wrongdoers should not be making offers that make others complicit in their wrongs. My point in focusing on victims is not to emphasize their problematic behavior, which is surely far less wrong than the behavior of perpetrators. Rather, it is to suggest that NDA bans do not necessarily ask victims to sacrifice income to which they are entitled.

B. Failed Settlements and Lower Payouts

Will banning NDAs in sexual harassment cases decrease settlement rates or payouts to victims?69 Predictions of these results rely on a simple interpretation of case-settlement incentives. Both plaintiffs and defendants have an economic interest in settling disputes. Doing so saves on the costs of trial and avoids the risk of unfavorable trial outcomes and of privacy losses imposed by trial publicity. The concern for lower payments makes sense because a defendant who cannot negotiate for secrecy will no longer see settling as a way to avoid the publicity harms of a trial. So the defendant will cap settlement offers at the point that equals the expected costs of trial (litigation costs plus expected damages). If permitted to use an NDA, the defendant would offer a premium for the value of avoiding publicity. The second harm—fewer settlements—might happen if the plaintiff’s expected benefit from trial is greater than the defendant’s best offer, which is lower because NDAs are unavailable. Sophisticated settlement models predict that at least in some instances, banning NDAs will indeed reduce victim payments and decrease settlement rates.70

69 Moss, supra at 878.
There are, however, reasons to doubt these predictions. First, some defendants may be more interested in avoiding litigation costs and the risks of large judgments than in avoiding publicity. If so, limits on NDAs will not lead to markedly lower payments or fewer settlements.

Second, insofar as avoiding publicity is important to defendants, settlements without NDAs might remain more attractive than litigation. This is so because litigation increases the risk of publicity, while settlement (even without an NDA) might be accomplished quietly.\textsuperscript{71} Because most victims value privacy, defendants can expect victims not to disclose even in the absence of an NDA, making settlement a sensible way of protecting privacy—certainly more effective than litigating. If guaranteed secrecy is not a legally available option, defendants who care deeply about their reputations may be willing to pay to avoid guaranteed disclosure. In other words, the driving force behind settlement payments (apart from avoiding trial costs and risk) may be securing the benefit of avoiding trial publicity.\textsuperscript{72} The marginal difference in expected levels of secrecy with and without an enforceable NDA may be small in comparison to the secrecy benefit of trial avoidance.\textsuperscript{73} These considerations suggest that payment reductions and settlement rate declines may be relatively small.\textsuperscript{74} Of course, if NDA bans do not significantly reduce payment to victims—because even without payments, perpetrators expect victim silence—then NDA bans also do not succeed in one of their main aims, which is to help disseminate information about harassers.

\textsuperscript{71} The sort of publicity that matters will vary from one defendant to another. Celebrities and public figures may be most concerned with press coverage. Other defendants and their employers may be concerned with whether sexual harassment allegations become known to friends and family, to other employees of the firm, or to stockholders. For all these groups, trial likely increases the chance of allegations becoming known. For example, trial preparation will involve discovery, which might include calling other employees as potential witnesses. And trials may require harassers to spend time in court, which could more easily be discovered by friends and family.

\textsuperscript{72} Whether the secrecy benefit of settlement without an NDA is large, compared to trial, will depend on many factors. This includes, of course, the victim’s inclination to publicize the harassment. It also depends on institutional details. For example, if civil complaints in the state are easily searched online (and not subject to being sealed), then settlement brings a risk that someone could find and read the complaint. But absent a trial, fewer people are apt to think about looking for such a document. Additionally, the complaint likely contains fewer details than would be revealed at trial. As well, many cases settle before a civil case is filed, perhaps after an administrative complaint or a demand letter. These may be less subject to easy access via online search than are civil complaints. Finally, many employment cases are resolved in arbitration. Depending on how they arrived in arbitration, they may not leave behind any searchable record.

\textsuperscript{73} Formal models of settlement incorporate this idea. The extent to which NDA bans reduce payment to plaintiffs and increase litigation depends in part on extent to which settlement reduces publicity risk compared to trial. See Hush Money, supra.

\textsuperscript{74} See Helland & Lee, supra, at 465 (2010) (“Some argue that confidential settlement bans will in reality have little effect because the default position for most settlements is in effect non-disclosure, and a confidentiality ban would leave that position unaffected.”)

27
Third, states have enforced various limits on non-disclosure for several decades, albeit not in sexual harassment cases.\textsuperscript{75} Anecdotal evidence suggests that these laws have not chilled settlement.\textsuperscript{76} One empirical paper finds that mandated disclosure reduced payments to plaintiffs and decreased settlement rates.\textsuperscript{77} But this paper examined medical malpractice settlements, which involved mandated public disclosure on publicly searchable web pages designed for the purpose of disclosing malpractice information. Those malpractice disclosures included both the alleged facts and the financial amount of settlement. The authors acknowledge that this evidence does not necessarily show what would happen if two key features were absent: aggressive public dissemination and disclosure of settlement amounts.

Whether sexual harassment NDAs will reduce victim compensation or decrease settlement rates are empirical questions, which will perhaps be answered over time now that several states have enacted bans on NDAs in sexual harassment cases. Even if NDA bans do reduce settlement offers to plaintiffs, this might not be a bad outcome. It suggests that NDA bans are increasing publicity about harassment. And perhaps (as argued above) the lower payments are appropriate.

C. Employer NDAs with Wrongdoers

Employers sometimes enter into NDAs with the harassing employee, often when negotiating the employee’s departure. For example, the employee might agree to resign (or not to sue the employer over a termination) if the employer promises not to disclose the reasons for the employee’s departure.

In this circumstance, the employer occupies the role of harassment victim in our earlier discussion—promising the perpetrator not to disclose in exchange for something the employer wants—in this case avoiding litigation and bad publicity. Just as with NDAs between perpetrators and victims, these NDAs impose risks on future victims. The employer has a far weaker justification than the harassment victim for imposing this risk. This is so for several reasons. First, the employer may be trying to hide its own wrongdoing, conspiring with the harasser to prevent anyone from discovering that either of them engaged in bad behavior. Second, unlike harassment victims, who rightly fear personal trauma and unjust employment discrimination if they reveal the facts of their harassment, employers aim primarily to minimize costs, such as defending against claims of wrongful termination. Although

\textsuperscript{75} For a review, see Richard Zitrin, The Laudable South Carolina Court Rules Must be Broadened, \textit{55 S. Carolina L.Rev.} 883 (2004).

\textsuperscript{76} See James Rooks, Jr., Settlements and Secrets: Is the Sunshine Chilly?, \textit{55 S. Carolina L.Rev.} 859(2004); Zitrin, supra.

\textsuperscript{77} Eric Helland & Gia Lee, supra.
employers also care about their reputations, they often will face limited reputational harm from having employed a harassing employee, unless they tolerated bad behavior for an extended period. Indeed, we might think that (unlike victims) employers always have a duty to disclose information about their employee’s harassment.

D. Does #MeToo Change Everything?

It is, of course, difficult to know how the #MeToo movement will shape our society over the coming years. We might be at the start of a social transformation that dramatically reduces sexual harassment. But like many movements, this one may face backlash and retrenchment. Despite this uncertainty, we can say something about the importance of possible rapid social change for the analysis above.

First, if sexual harassment becomes far less tolerated than it was just a few years ago, victim interests in keeping secret the details of their harassment could become less important. Part of the trauma of public disclosure comes from victim shame and from fear of social condemnation. These could be less severe in a more supportive environment. As well, victims fear that becoming known as a victim, or as someone who complained about harassment, could undermine future job prospects. Any social transformation could reduce these effects. If so, bans on NDAs would be more justifiable because non-disclosure itself would be less reasonable.

Second, if social stigma surrounding victimhood were reduced, reporting of harassment and victim willingness to disclose the details of harassment could increase. If this happened, NDAs bans might dramatically reduce the amount offered in settlement compared to what might be offered without such bans. If most victims could be expected to discuss harassment willingly absent an NDA, perpetrators might come to see settlement without an NDA as little better than trial as a means of avoiding publicity. When NDA bans lead to widespread publicity, they also reduce the publicity-avoidance value of settling with victims.

Third, if changed social practices do undermine the significance of victim privacy, we likely should not worry about the lost compensation for victims who cannot sign an NDA. The amount of money foregone will be larger. But the claim that this money represents an unjustified betrayal of future victims will also be stronger.
Conclusion

Banning non-disclosure agreements for sexual harassers and other wrongdoers can help prevent future wrongs. Although these bans are sometimes thought to undermine privacy interests of accused perpetrators and to deprive victims of deserved compensation, these concerns are less serious than NDA advocates suggest. NDAs are an excessive means of protecting the wrongly accused, as they likely more often shelter the guilty than the innocent. And they are not appropriate as a means of protecting guilty harassers from excessive social punishment. Unlike the people protected by criminal expungement statutes, NDA beneficiaries are not systematically subject to excessive official and social punishments that contribute to a serious social problem. As to victims, arguably many of them should not seek compensation for silence. Some of them breach their duty to alert future victims to the harassers’ behavior, if not immediately, then as soon as their legitimate privacy interests fade. Others breach no duty since their privacy interests justify their silence. But for them, NDA payments represent a windfall. Further, all victims who sign NDAs become complicit in harassers’ wrongdoing by accepting payment for their silence.