Human Rights Change, Politics of Law and Order, and Targeting of Torture

Oskar Timo Thoms
The Simons Papers in Security and Development are edited and published at the School for International Studies, Simon Fraser University. The papers serve to disseminate research work in progress by the School’s faculty and associated and visiting scholars. Our aim is to encourage the exchange of ideas and academic debate. Inclusion of a paper in the series should not limit subsequent publication in any other venue. All papers can be downloaded free of charge from our website, www.sfu.ca/internationalstudies.

The series is supported by the Simons Foundation.

Series editor: Jeffrey T. Checkel
Managing editor: Martha Snodgrass


ISSN 1922-5725

Copyright remains with the author. Reproduction for other purposes than personal research, whether in hard copy or electronically, requires the consent of the author(s). If cited or quoted, reference should be made to the full name of the author(s), the title, the working paper number and year, and the publisher.

Copyright for this issue: Oskar Timo Thoms, oskar_thoms(at)sfu.ca.
Human Rights Change, Politics of Law and Order, and Targeting of Torture

Simons Papers in Security and Development
No. 66/2018 | November 2018

Abstract:
Human rights have improved but not everywhere and for everyone. Scholarship has focused on domestic conditions under which they improve but we know little about how they affect different groups. Whose rights are being protected? Under what conditions? I compare dissidents and criminals as targets of human rights violations – specifically torture. I also examine the effectiveness of human rights protections under conditions of public insecurity due to crime – as opposed to political or civil conflict or terrorism. I argue that mobilization and judicial enforcement are less effective in the face of public insecurity, and criminals benefit less than dissidents because courts provide less accountability for violations of those accused of crimes. Human rights treaties that depend on these mechanisms thus primarily benefit dissidents. My statistical analysis supports this argument and directly addresses concerns about measurement bias. The key finding is that commitment to the Convention against Torture enhances judicial protection only for dissidents.

About the author:
Oskar Timo Thoms is the Simons Research Fellow in International Law and Human Security at Simon Fraser University. He holds a Ph.D. in Politics from Princeton University (2017) and a B.A. in International Development Studies (2003) from McGill University. Prior to arriving at SFU, he was Research Associate at the University of Ottawa’s Centre for International Policy Studies. More information can be found at oskarthoms.net.

About the publisher:
The School for International Studies (SIS) fosters innovative interdisciplinary research and teaching programs concerned with a range of global issues, but with a particular emphasis on international development, and on global governance and security. The School aims to link theory, practice and engagement with other societies and cultures, while offering students a challenging and multi-faceted learning experience. SIS is located within the Faculty of Arts and Social Sciences at Simon Fraser University. Our website is www.sfu.ca/internationalstudies.
Human Rights Change, Politics of Law and Order, and Targeting of Torture

1. Introduction

While human rights have improved in many places, they have not in all places, and importantly, not for all groups. Many countries violate physical integrity despite their constitutional and international treaty commitments. One obstacle to better understanding where and why human rights improve is that human rights change is difficult to measure. Another is that research has focused much more on human rights everywhere than for everyone. Scholarship has focused on domestic conditions under which human rights improve but we know little about how they affect different societal groups. In this paper, I argue that to better understand mechanisms of human rights change, scholars need to ask: Whose rights are being protected? And under what conditions? To address these questions, I distinguish torture of dissidents from torture of criminals and examine the role of public insecurity due to crime – as opposed to political opposition, civil conflict or terrorism.

Scholarship attributes improvements in state respect for physical integrity rights to democratic institutions, domestic political mobilization and judicial enforcement of domestic and international legal obligations. In her agenda-setting study, Simmons argues that human rights treaties have modest positive impacts on a range of civil rights, including freedom from torture, in states that have democratized because they encourage domestic political and legal mobilization by civil society actors to demand human rights protection. Domestic judicial institutions in particular are considered an important mechanism. Studies find that independent

---

1 I thank Andrew Moravcsik, Grigore Pop-Eleches, Courtenay Conrad, Will H. Moore, Jeffrey T. Checkel, Chris Fariss, Emilie Hafner-Burton, Jillian Haglund, Aaron M. Hoffman, Yonatan Lupu, Brenda Lyshaug, Tamir Moustafa, Roland Paris, Irene Pang, Claire Peacock, David Petrasek, Joanna Quinn, Laurel Weldon, and participants at the CPSA Annual Meeting (May 30, 2017) and in the SFU School for International Studies Research Colloquium (October 1, 2018) for insightful comments on this project.


3 Simmons 2009.
and effective judiciaries are strongly associated with better physical integrity rights protection,\(^4\) and are key to the enforcement of human rights treaty commitments.\(^5\) Other recent research further elaborates on these mobilization and litigation mechanisms.\(^6\)

The literature on conditional human rights impacts largely assumes that governments use repression to control and weaken political opposition threatening leader tenures.\(^7\) Scholarship on the determinants of repression has also focused on the role of threatening dissent.\(^8\) This is theoretically appealing in authoritarian regimes and when dissent is violent in democracies,\(^9\) but human rights victims are often not political dissidents. While much repression is a response to dissent, it is far from clear that most repression is driven by political opposition or that threatening dissent is the most important driver of violations.\(^10\) Many physical integrity violations result from routine policing, criminal procedure, and social control, and not the targeted repression of political opposition. Referring to torture allegations against state officials in 150 countries between 1997 and 2000, a joint Amnesty International (AI) and Oxfam report notes: “The evidence strongly suggests that most of the victims were people suspected or convicted of criminal offenses. Most of the torturers were police officers who used armed threats and violence to subdue their victims.”\(^11\) Some policing scholars have written about such abuses as a democratization of state violence.\(^12\)

---

\(^4\) Keith, Tate, and Poe 2009; Keith 2012; Mitchell, Ring, and Spellman 2013; Hill and Jones 2014; Crabtree and Fariss 2015.


\(^6\) See, for instance, Conrad and Ritter 2013; Lupu 2015.

\(^7\) Neumayer 2005, 931; Simmons 2009, 122; Hill 2010, 1164, 1167; Conrad and Ritter 2013, 398–399; Conrad 2014; Lupu 2015, 582.

\(^8\) Moore 2000; Regan and Henderson 2002; Davenport 2007; Carey 2010.

\(^9\) Conrad and Moore 2010; Conrad 2014.

\(^10\) Rejali 2007; Hafner-Burton 2013, chapter 2; Conrad, Hill, and Moore 2018.

\(^11\) Hillier and Wood 2003, 27.

\(^12\) A review of policing scholarship notes: “In new, often unstable democratic systems the police also mirror social inequality and popular ambivalence toward the rule of law. Woven throughout these studies is a complex and seemingly paradoxical relationship among democratization, social values, the rule of law, and police brutality. In many cases not only has the transition to electoral democracy failed to strengthen the rule of law, but democratic politics have created social pressures that encourage police abuses. After the citizenry has won its electoral rights, it is frequently willing to tolerate abuses of civil rights and the rule of law in the name of law and order, targeted particularly against marginalized groups and classes.” Tanner 2000, 119; see also Uildriks and Van Reenen 2001.
Moreover, existing scholarship views human rights change primarily as driven by opposition politics and civil society pressure on governments to reform state institutions, and regards independent judiciaries as faithful enforcers of domestic and international legal provisions for human rights. The assumption that political opposition groups will pressure the government to equally protect the human rights of all citizens is difficult to maintain. Continued violations are often due to the political failure to reform the police. While demanding rights for themselves, opposition members in the legislature are not likely to block repressive policies against (suspected) criminals and marginalized groups when there is societal support for such policies. Marginalized groups, of course, do not have political influence. A recent study finds that electoral institutions are associated with higher levels of torture that scars the victim’s body, possibly because of public demand “for executives to ‘dirty their hands’ against criminals and dissidents to protect the body politic.” Yet, key electoral sectors can benefit from opposition contestation over human rights and at the same time be content to sacrifice the rights of marginalized societal sectors. Human rights organizations (HROs) have found it particularly difficult to advocate for the human rights of such societal groups, and experts note long-standing antagonism between the police and human rights advocates.

Scholarship also suggests that there are limits to the human rights protecting role of judiciaries. Civil law systems are associated with worse physical integrity protection than common law systems, and are considerably less likely to stop using torture, regardless of their level of judicial independence. Some attribute this to inquisitorial criminal justice procedures in civil law systems and the high value they place on confessions in criminal cases. A recent study finds that greater judicial independence is associated with fewer allegations of “scarring” but

---

14 Powell and Staton 2009; Simmons 2009; Conrad and Ritter 2013; Conrad 2014.
17 Hollar 2013, 117.
19 Conrad and Moore 2010.
20 Conrad and Moore 2010, 465; see also Rejali 2007, 50.
more allegations of “stealth” torture,\textsuperscript{21} providing support for Rejali’s argument that increased human rights monitoring has led to the development and use of “clean” torture techniques in democracies.\textsuperscript{22}

I propose that the physical integrity-improving impacts of judiciaries and international treaties may have largely been focused on particular targets of violations: dissidents. Other societal groups may benefit less, particularly when there is domestic backlash against human rights due to public insecurity. Simmons recognizes that mobilization for human rights can be met by opposition and counter-mobilization and does not necessarily lead to human rights improvements, and she speculate that this accounts for the “less convincing impact on fair trials” in her study because providing fair trials “can be framed as being soft on crime.”\textsuperscript{23} However, due process protections affect both political dissidents and those accused of crimes, so it is not clear why fair trials protections would not benefit from mobilization when other rights do. Opposition to reform and counter-mobilization also undermine the protection of physical integrity, which according to Simmons and others, improve due to mobilization and judicial enforcement. This paper It is not only a question of whether protection of certain rights improves on average or not, but also whether these causal mechanisms function unevenly across countries and societal groups within them, and why. The present study begins to answer the call of a recent review article: “If repression eases, does it ease equally for all groups? Does it shift from one area to another? Who gains in the wake of rights treaties and trials and who loses, and through what mechanisms?”\textsuperscript{24}

This research contributes to comparative politics and international relations literatures. First, it builds on important case study research on policing and domestic legal accountability,\textsuperscript{25} and provides broader cross-national empirical support for their arguments. Second, it brings this

\begin{itemize}
  \item \textsuperscript{21} Conrad, Hill, and Moore 2018.
  \item \textsuperscript{22} Rejali 2007.
  \item \textsuperscript{23} Simmons 2009, 200, 148.
  \item \textsuperscript{24} Hollar 2013, 109.
  \item \textsuperscript{25} Rejali 2007; Brinks 2008.
\end{itemize}
literature into international relations scholarship on impacts of human rights treaties in order to better understand the domestic scope conditions for key causal mechanisms.

Section 2 reviews the human rights literature to which this study contributes, and outlines key alternative explanations. Section 3 draws on key contributions in comparative politics scholarship to argue that mobilization for human rights is particularly difficult under conditions of public insecurity due to crime, and this in turn limits legal accountability for police abuses. Moreover, dissidents are the primary beneficiaries of human rights change because mobilization is more likely in their defense. Section 4 discusses the empirical approach and explains how this study addresses key inferential concerns about measurement bias of human rights data and treaty selection effects. In Section 5, the last before my Conclusion, I find empirical support for three claims in analyses of disaggregated data on allegations of torture. First, public insecurity due to crime undermines judicial protection of physical integrity. Second, victim type matters to the effectiveness of human rights change. Judicial independence is more effective in reducing torture of dissidents than of criminals. Third, ratification of the Convention against Torture and other Cruel, Inhuman or Degrading Treatment or Punishment (CAT) enhances the judicial protection of only dissidents.

2. Literature

Human rights treaties prohibiting physical integrity and other civil rights violations are one of the most salient forms of international cooperation, nearing universal membership in the core treaties. Commitments to these treaties are at the center of the increasing global legalization of human rights. These human rights have also been almost universally constitutionalized at the domestic level, and many countries have established national human rights institutions such as ombuds offices and permanent commissions to monitor and publicize domestic developments.

---

26 Physical integrity rights are a subset of civil rights, including the rights not to be tortured, extrajudicially killed, disappeared or politically imprisoned.

27 Keith, Tate, and Poe 2009; Sandholtz 2012, 17; Law and Versteeg 2013; Chilton and Versteeg 2015.

28 Koo and Ramirez 2009; Cole and Ramirez 2013.
Two decades of research have found, at best, modest impacts of international human rights commitments on state practices, and the conditions under which treaties have positive effects continue to be debated. Early scholarship was unable to show positive impacts on physical integrity, and often surprisingly found negative impacts. Several quantitative studies find associations between commitments to the International Covenant on Civil and Political Rights (ICCPR) or the CAT and increased violations of these rights. Another study finds that the strong international anti-torture norm, measured as the global CAT membership rate, has not led to an overall reduction in reported torture. Constitutional prohibitions of torture are also associated with higher levels of commonly used torture measures, particularly in democratizing countries.

Early findings of associations between treaty membership and abuses are undermined by the two key inferential problems of measurement bias of human rights indicators and treaty membership selection effects. Recent studies have begun to address these analytical issues in various ways and find conditional positive treaty impacts: treaty commitments are associated with reduced violations under certain domestic conditions. These studies have moved the focus from the question of whether international law matters to how it matters under what conditions.

Subsequent scholarship focused on explaining or disputing the seemingly paradoxical association of human rights commitments and increased violations found in some studies. On one side of the debate, scholars argue that this association is due to authoritarian leaders making commitments while still having incentives to continue their repression. Some believe that the paradox reflects highly sophisticated strategic behavior by autocrats joining human rights treaties in order to raise the cost of repression, and by showing their willingness to bear this cost actually

---

31 Gilligan and Neshitt 2009.
32 Keith, Tate, and Poe 2009; Law and Versteeg 2013; Chilton and Versteeg 2015.
33 Landman 2005; Powell and Staton 2009; Simmons 2009; Conrad and Ritter 2013; Conrad 2014; Fariss 2014; Lupu 2015; Fariss 2018.
signal to their citizenry their commitment to repressing dissent. Others maintain that autocrats commit to human rights treaties as a concession to increasingly powerful and organized opposition groups, but that same increase in power reflects a threat to the regime and is thus met with further repression. Another variant of these arguments views treaty commitment as a way of bolstering domestic and international legitimacy and gaining support from activists and foreign governments; this is an attractive step for governments especially if they expect to continue or expand abuses. These theories are insightful in explaining why already abusive governments make treaty commitments.

On the other side of the debate, scholars argue that the findings of an association between treaty commitments and abuses are wrong due to biased measurement and inference. In key contributions, several scholars argue that this association is an artifact of poor but improving measurement of state behavior, because as governments become less repressive and more integrated into the global legal system, those within their borders become more likely to report abuses and HROs are increasingly likely to publicize them. These scholars maintain that the increasing quality and quantity of human rights information over time has biased studies to find negative human rights impacts. Others argue that the negative association between treaty commitments and abuses results from abusive states joining treaties at high rates but that the treaties do not have negative impacts. Thus, associations between treaty commitments and abuses are due to selection effects.

Several scholars argue that human rights treaties improve physical integrity under certain conditions, through domestic or vertical enforcement mechanisms such as mobilization by civil society and political opposition members and judicial enforcement of treaty obligations.

34 Hollyer and Rosendorff 2011.
35 Vreeland 2008.
38 Lupu 2013.
40 Neumayer 2005; Powell and Staton 2009; Simmons 2009; Conrad and Ritter 2013; Lupu 2015; on the distinction between horizontal and vertical enforcement of international law, see Moravcsik 2012; for a review of scholarship on human rights mobilization, see Tsutsui, Whitlinger, and Lim 2012.
Simmons theorizes that human rights treaties empower domestic groups and institutions. She argues that treaties provide resources for political mobilization by introducing ideas about the relationship between governments and citizens; translating existing grievances into human rights terms and proposing specific rights claims; signaling the legitimacy of rights demands by pre-committing the government to be receptive; providing focal points for the formation and expansion of pro-change coalitions; and providing benchmarks against which to hold government behavior to account. Recruitment of new influential actors, including legislators, to the cause is particularly important.\footnote{Simmons 2009, 139–145; see also Tsutsui, Whitlinger, and Lim 2012.} Regarding legal mobilization, Simmons argues that treaties provide specialized resources encouraging litigation if states have functioning and reasonably independent legal systems, which potentially entails reputational costs and a loss of government resources. Treaties enhance engagement by legal professionals, such as cause lawyers and judges, because this group is more likely to engage with human rights once a government has made treaty commitments.\footnote{Simmons 2009, 146–148; see also Powell and Staton 2009, 154; Keith 2012, 169.} According to Simmons, this enforcement mechanism is relevant only in democratic countries because in autocracies, political control over judiciaries is too pervasive.\footnote{Simmons 2009, 152.} While democracies are more likely to have independent judiciaries, other research finds that effective judiciaries are also associated with better protection in dictatorships, moderating dictators’ incentives both to commit to the CAT and to torture.\footnote{Conrad 2014.} 

Others have elaborated on conditions for successful human rights mobilization. For instance, Conrad and Ritter point to the conflicting incentives regarding repression due to treaty-induced mobilization and judicial enforcement because political leaders may view increased mobilization as threatening dissent. They argue that leaders who are more securely in power are more constrained by the CAT because they would increase repression but are also more likely deterred from increasing repression by the prospect of costly litigation. By contrast, politically "vulnerable leaders will do what they must to remain in power” without regard for judicial consequences.\footnote{Conrad and Ritter 2013, 397.} In another study, Lupu proposes a legislative mobilization mechanism, finding
that the extent to which commitments to the ICCPR are associated with improved human rights practices increases with the number of legislative opposition groups. These veto players increase the cost of violations through their formal powers such as legally sanctioning abusive executives and controlling government spending. Lupu argues that the ability of the legislative opposition to use these powers is enhanced by treaty membership because they are more likely to obtain relevant human rights information due to the efforts of civil society actors.\textsuperscript{46} Veto player arguments are also prominent in comparative politics scholarship on the relationship between regime type and repression, arguing that democracies are internally more peaceful than authoritarian regimes because more numerous and more powerful veto players place constraints on the chief executive’s use of repression.\textsuperscript{47}

3. Theory

This study is concerned with decentralized state violence, in which specific violations of human rights are not directly ordered by political leaders, but result from law enforcement. As the primary coercive arm of the state, the police are tasked with upholding law and order, and they are also the most likely to violate physical integrity.\textsuperscript{48} These violations are often responses to public insecurity. Security is a fundamental public good to be provided by any state. In the Hobbesian social contract, citizens cede power to the sovereign in return for physical security. In order to remain relevant and maintain trust among the citizenry, police forces need to be seen to provide general public order and safety. High levels of public insecurity due to crime – especially violent crime – put pressure on the police to achieve results quickly.\textsuperscript{49}

Police may violate human rights in the course of law enforcement with arbitrary detentions, violence, and other extra-legal behavior for a variety of reasons. As responses to public insecurity due to crime, three motivations stand out. First, violations may be a response to the dangerous work of policing due to the unpredictability inherent in confrontations with

\textsuperscript{46} Lupu 2015, 583.
\textsuperscript{47} Davenport and Armstrong 2004; Bueno de Mesquita, Downs, Smith, and Cherif 2005.
\textsuperscript{48} Chevigny 1995; Tanner 2000, 119; Call 2002; Neild 2003, 277; Goldsmith 2003; Uildriks 2009.
\textsuperscript{49} Marenin 1985; Tanner 2000; Shaw 2002; Hinton and Newburn 2008.
citizens. Police officers may react violently to counter real or perceived threats to their own physical integrity.\(^{50}\)

Second, police may violate physical integrity in order to do their job more effectively. Here, human rights are regarded as obstacles to effective law enforcement. Torture is often employed to extract information used in criminal investigations. Such information can be used to gain investigative leads or full confessions. Police have strong incentives to torture in legal systems that place a high value on confessions as evidence in criminal trials.\(^{51}\)

Third, physical integrity violations may be used by police as punishment; the goal is to create order where the state has been unable to provide public security.\(^{52}\) Violations of physical integrity can be revenge for defiance against police authority or serve as deterrents against criminals. Punishment aims at substituting for real or perceived shortcomings of the criminal justice system, when the legal process is regarded as too slow and constrained by rights protections to deal with the problem of crime. Abuse also functions as an informal alternative to a criminal charge when there is not sufficient evidence to prosecute or serves as additional punishment where the criminal justice system is perceived as too lenient.\(^{53}\)

There are other possible motivations for violations in contexts of public insecurity, including personal gratification and profit motives.\(^{54}\) While these do not aim at establishing order and safety, they may also thrive when there are weak political and social constraints on abusive police actions. Moreover, police may violate the rights of individuals against whom they are biased or prejudiced on the basis of their social identity.

\(^{50}\) Chan 2000, 97; Uildriks and Van Reenen 2001.

\(^{51}\) Uildriks and Van Reenen 2001, 77; Rejali 2007, 50; Brinks 2009, 10.

\(^{52}\) Rejali 2007, 57.

\(^{53}\) Chan 2000; Uildriks and Van Reenen 2001, 76.

\(^{54}\) Uildriks and Van Reenen 2001.
3.1 Mobilization and opposition

Scholars often note the institutional stickiness of physical integrity abuses. Once states violate, they are likely to continue to do so.\textsuperscript{55} Human rights change requires the reform of abusive institutions, through vetting of state agents, refocusing the mission and capacity of law enforcement agencies, and strengthening their legal accountability. Law enforcement bureaucracies often oppose these reforms with much political support, particularly in states with recent histories of abuse.\textsuperscript{56} To succeed, mobilization for human rights must overcome such opposition, which involves forming a coalition by recruiting new actors to the cause. Particularly important is support from actors who do not directly benefit from human rights change but who have political influence.\textsuperscript{57}

Public insecurity undermines human rights mobilization. Societal backlash often includes demands for stronger law enforcement. Police abuses occur with at least tacit political and social support.\textsuperscript{58} For instance, in many Latin American democracies, public insecurity has been a hugely salient issue, leading to public support for heavy-handed (\textit{mano dura}) policing.\textsuperscript{59} Electoral accountability can translate into electoral support for repressive policing. Under conditions of insecurity, candidates seeking political office are likely to be voted in rather than out for supporting heavy-handed law enforcement.\textsuperscript{60} According to much scholarship, public and electoral support for increased police powers and decreased oversight has been a key obstacle to reform.\textsuperscript{61}

\textsuperscript{56} Tanner 2000; Frühling, Tulchin, and Golding 2003; Pereira and Ungar 2004; Bergman 2006; Fuentes 2006; de Mesquita Neto 2006; Frühling 2009; Ungar 2011, chapter 2; Hollar 2013, 117–118.
\textsuperscript{57} Simmons 2009, 145.
\textsuperscript{58} Chan 2000, 88.
\textsuperscript{60} Fuentes 2006; Malone 2010; Armstrong 2013.
Under these conditions, it is more difficult to enlarge the coalition of actors demanding universal rights protections for all societal sectors. Human rights organizations, which have traditionally focused their advocacy on politically motivated repression and centralized impunity under authoritarian rule, have found public debate more hostile to advocacy against post-transition police abuses than to accountability for past authoritarian abuses because they face strong countermobilization by pro-order activists. Conservative civil society groups and politicians often see the protection of human rights as an obstacle to the enforcement of public order, and argue that public insecurity necessitates a heavy-handed response.

3.2 Limiting legal accountability

Many studies find that judicial independence and effectiveness are associated with better human rights protection. Judicial independence requires judges to be free to make legal decisions based on their interpretation of the law and without political interference from the government in power. An independent judiciary with formally delineated authority and clear legal basis for human rights jurisprudence is expected to take seriously its mandate to review the actions of state agents to protect fundamental rights. Citizens are also more likely to bring litigation against the state if they have reason to believe that the courts actually constrain state behavior. Successful human rights litigation entails reputational costs and a loss of government resources, providing incentives to rein in abusive behavior.

Controlling violations in the course of law enforcement requires functioning institutional oversight and legal accountability. Judicial independence and effectiveness are the necessary prerequisites for legal accountability, but courts need to be motivated to hold law enforcement

---

62 For the case of fair trials, see Simmons 2009, 200, 357.
63 Neild 2003; Davis 2006; Fuentes 2006; Eaton 2008; Bergman and Whitehead 2009; Ungar 2011; Bonner 2014.
64 Cross 1999; Keith 2002; Apodaca 2004; Keith, Tate, and Poe 2009; Keith 2012; Crabtree and Fariss 2015.
65 Powell and Staton 2009; Conrad and Ritter 2013; Conrad 2014.
67 Keith, Tate, and Poe 2009; Keith 2012, 115, 169.
68 Powell and Staton 2009, 151; Conrad and Ritter 2013.
69 Powell and Staton 2009, 154; Simmons 2009; Keith 2012, 169.
actors accountable. An independent judiciary is usually assumed to be interested in protecting individual rights even when other government branches are not.\textsuperscript{71} This assumption is questionable. Judges can have the necessary autonomy and authority to rule against state agents, but they are not necessarily committed to holding them accountable for abuses.\textsuperscript{72} Judges use legal reasoning to make rulings on the basis of evidence brought before them, but courts are not simply instruments of the law; they are also political actors. All legal actors involved in prosecutions “have the potential and opportunity for strategic behavior.”\textsuperscript{73}

Brinks argues persuasively that the preferences of criminal justice actors over legal accountability for police violations of human rights differ substantially from their preferences over ordinary criminal cases. In an ordinary violent crime, such as murder, the police should be, after the claimant, the most interested in a conviction, followed by the prosecutor who is interested in only prosecuting cases that she can win. The role of the judge is to be the neutral arbiter. By contrast, the preferences of these actors over convictions for police violations are biased in favor of the defendant. If violations routinely occur in the course of police carrying out their duties, then the police will side with the defendant because convictions limit their ability to employ repressive tools and expose them to future prosecutions. The preferences of prosecutors and judges in these cases derive from their relationships to the police. Courts rely on the police to collect, screen and forward relevant information. As the authoritative coercive institution, this means that the police can control information to their own advantage, for instance, by withholding evidence or intimidating witnesses.\textsuperscript{74} Moreover, courts do not have the resources to enforce their rulings, but rely on the police to do so.\textsuperscript{75}

Due to the operational dependence of courts on the police, the preferences of prosecutors and judges in misconduct cases are pulled closer to those of the police. Moreover, police and prosecutors can use their control over information to advance their preferred outcome. Abusive police have incentives to limit or undermine the collection of evidence supporting a prosecution;

\textsuperscript{71} See Larkins 1996; Keith 2012, 170.
\textsuperscript{72} O’Donnell 1999, 39; see also Cross 1999.
\textsuperscript{73} Kennedy 2004, 22; see also Brinks 2008, chapter 1.
\textsuperscript{74} Brinks 2008, 28–30.
\textsuperscript{75} Keith 2012, 149; see also Brinks 2008, 29; Simmons 2009, 134.
Prosecutors can assign insufficient resources to these cases; and judges can accept the resulting informational failures as grounds for dismissing cases or ruling in favor of the defendant. For these reasons, Brinks argues that “[e]ven a strictly neutral judge is likely to reproduce the pro-defendant bias caused by the failure of other supporting actors. In short, without some affirmative corrective, the typical criminal justice system is structurally predisposed toward failure in these cases.” Note that this endogenous judicial bias is not due to a lack of independence from the government in power; it is compatible with judicial independence.

These endogenous incentives interact with exogenous incentives provided by broader social and political pressures. In addition to direct political accountability, such as when judges and prosecutors hold elected positions or are put into their positions through political appointments, the politics of law and order produce indirect pressures. Since they lack the power of “sword and purse,” courts rely on diffuse public support for other actors to comply with their legal decisions. The judiciary builds and maintains legitimacy when its rulings are accepted and implemented by other actors. Scholars have argued that courts try to avoid losing legitimacy by issuing too many unpopular decisions. Therefore, courts are strategic in choosing which legal cases to admit and in deciding outcomes. In cases of police violence, courts are managing both the legitimacy of the criminal justice system and their relationship with the police.

Following Brinks, political and legal mobilization for human rights is crucial to overcoming the bias of the criminal justice system in cases of police violations. As discussed, mobilization is more tenuous in the face of public insecurity. When there are no threats to public security and public opinion supports universal human rights, court rulings in favor of these rights reinforce judicial legitimacy. Judiciaries enforce human rights obligations because broad, sustained social and political demands for rights protection lead prosecutors and judges to take police abuse cases more seriously, assign greater resources for prosecuting such cases, and use

---

76 Brinks 2008, 32.
77 Brinks 2008, 34.
78 Vanberg 2005; Gibler and Randazzo 2011, 698; Keith 2012, 189.
their authority to overcome informational failures resulting from police resistance to prosecutions. \(^{79}\)

By contrast, public support for heavy-handed policing reinforces the bias of the criminal justice system in favor of defendants in police abuse cases. When the public is not supportive of human rights protections for criminals, judicial enforcement will be limited, either because courts refuse to proceed with cases against police or fail to convict. Judicial independence is not a guarantee for legal accountability. More independent judiciaries generally protect human rights more than weaker judiciaries, but the rights-improving impact of judicial independence will be decreased in the face of public insecurity. Another key implication is that public insecurity particularly limits the judicial protection of (suspected) criminals. The human rights of dissidents should be more protected by judiciaries under these conditions.

In sum, public insecurity due to crime limits mobilization for human rights by strengthening public support for heavy-handed policing, putting political constraints on reforms. Moreover, while judicial institutions are key to protecting human rights, public insecurity makes courts less likely to hold law enforcement agents legally accountable for violations. Thus, these mechanisms of human rights change are expected to be less effective in the context of high levels of public insecurity, and criminals and suspects benefit less from human rights change than political dissidents, because courts will provide less legal accountability for police violations of those accused of crimes. Importantly, to the extent that human rights improving impacts of treaties depend on mobilization and judicial enforcement (see Section 2), treaty commitments should be associated with less improvement at higher levels of public insecurity and enhance protection against torture primarily for dissidents. The statistical analyses below test the following hypotheses derived from this theory:

H1 Judicial independence is associated with more protection for dissidents than for criminals against torture.

H2 Judicial independence is associated with decreased torture at low levels of public insecurity but not at high levels.

\(^{79}\) Brinks 2008.
H3 In contexts of public insecurity, CAT commitment enhances judicial protection against torture of dissidents but not of criminals.

To provide strong support for the theory, the analyses need to account for competing explanations, most importantly prominent arguments in the literature about information effects and changing standards of accountability. The former problem is that observations of violations are driven by the increasing availability of information on violations. The analyses therefore control factors that are expected to influence the likelihood that violations are detected and reported by HROs. The importance of such factors may vary across different targets of violations, because, for instance, HROs have found it more difficult to campaign for the rights of criminals than of political dissidents. The latter problem is that HROs have applied increasingly stringent “standards of accountability” in assessing the human rights practices of states. The disaggregated torture data used in this chapter is well suited to address these issues. The analyses also need to address potential treaty selection effects. These issues are discussed in the next section.

4. Empirics

4.1 Human rights data

The analysis requires a dependent variable that distinguishes between different targets of human rights violations. The Ill-treatment and Torture Allegations (ITT) dataset is currently the only cross-national dataset that disaggregates allegations of violations by victim type but these data are available for a shorter time period (1995–2005) than other cross-national data.\(^80\) This dataset quantifies specific AI allegations of violations by state agents prohibited by the CAT.\(^81\) These event data are coded based on content analysis of all AI publications, including annual reports, press releases, and action alerts, and distinguish both the state agency of control committing the violations and the type of victims. ITT’s typology of victims builds on Rejali’s


\(^{81}\) An allegation of state torture or ill-treatment occurs when AI alleges that the perpetrator is an agent of the state, the victim(s) is a person detained under the state’s control, and the alleged abuse meets the definition of torture and/or ill-treatment in the CAT. Conrad, Haglund, and Moore 2014, 430.
typology of motives for torture, which includes criminal investigation, national security interrogation, and social control. The categories of victims in the dataset are: criminal; dissident; marginalized individual; state agent; prisoner of war; and unstated/unknown. These categories are not mutually exclusive because victims described in AI documents at times fit more than one identity.

The fundamental problem of human rights measurement is that scholars infer the level of violations from biased subsets of reported violations. Violations are often deliberately hidden from the public and even from superiors within the state for moral and legal reasons. With the emergence of the modern transnational human rights movement, the quality and quantity of critical human rights information has increased dramatically since the 1970s, giving rise to potential measurement biases. On the one hand, the quality and quantity of information on physical integrity violations has increased over time as monitoring organizations have multiplied and gathered more information on abuses in more places. At the same time, HROs have limited resources for researching and reporting violations, further affecting the production of abuse allegations, because HROs have to choose strategically where to focus their efforts. On the other hand, in continually pressing governments to reform, activists have tended to classify more and more acts as violations. As a result, human rights reports became increasingly stringent assessments of state behavior over time. Since the standards-based measures commonly used in scholarly research, such as the Political Terror Scale and the Cingranelli & Richards (CIRI) indicators, are coded from such reports, they reflect a changing “standard of accountability,” and may not reveal actual changes in the level or types of abuses.

These measurement biases pose a considerable problem for inferences of impacts. It is particularly difficult to infer impacts of worsening human rights practices because the indicators are potentially biased toward finding exactly that. Similarly, event-based human rights

82 Rejali 2007.
84 Fariss 2014.
85 Clark and Sikkink 2013.
86 Clark and Sikkink 2013, 568.
indicators come with their own measurement biases, making their suitability for trend analyses also potentially questionable. As states lessen repression, those within their borders may be more likely to come forward with allegations of abuse. This means trends in event count measures may often fail to track actual trends. Moreover, most event-based measures are coded from news sources, which are particularly poorly suited to human rights measurement. Finally, some scholars note that measurement bias may be particularly a problem for measuring violations that target individuals other than political dissidents, such as police killings and torture of criminal suspects, because HROs have traditionally focused on state violence directed at political opponents. Given these concerns, to make convincing inferences, the analysis needs to model potential measurements biases.

There are more than 16,000 torture events in the ITT dataset. (In this paper, torture refers to all ill-treatment and torture allegations in the dataset.) These events do not reflect the intensity of torture or the number of victims involved but refer to reported torture incidents, which can include several or more victims. Figure 1 shows the yearly totals for four victim types, which make up most of the allegations in the ITT dataset; there are very few allegations involving state agents and prisoners of war. These data show that by the mid-1990s, AI did not focus its reporting more on dissidents than on other victim types. The similar magnitude and trends over time in allegations of torture of dissidents and criminals provide confidence in the validity of statistically comparing the determinants of these events counts. For the present study, the specific allegations data are aggregated to event counts at the country-year level to create three dependent variables. The first consists of all torture allegations, regardless of victim type, while the others consist of allegations of violations targeting criminals and dissidents, respectively.

87 Davenport and Ball 2002.
88 See Marchesi and Sikkink 2015, quoted in Fariss 2018, 253-254.
89 The coding manual defines a criminal as “One believed to have contravened statute, excluding crimes that are considered threats to national security. We do not code Victim Type as Criminal in instances where a victim has broken a law that is in opposition to the articles in the Universal Declaration of Human Rights. Absent other information, we consider Prison Populations (pre- and post-sentence) to be Criminals. Asylum seekers being deported were coded both as Marginalized Individuals and Criminals.” A dissident is defined as “One believed to be a threat to the state or be willing to engage in illegal activity to challenge policy. Note that we consider Prisoners of Conscience, Human Rights Defenders, and
The relatively short time period covered by the ITT data comes with an advantage for controlling the standard of human rights accountability that may bias measures. Schnakenberg and Fariss derive a latent measure of physical integrity practices from a measurement model designed to take into account the changing standards of accountability by combining a range of standards-based and events measures of repression and state mass violence and incorporating over-time effects. Fariss finds that the ITT level of torture measure, which is coded from the ITT specific allegations data used here, is the only measure in the measurement model that is not subject to temporal change. According to this evidence, the standard of accountability reflected in the ITT data is generally static. Thus, these data already mitigate concerns about

Protesters to be Dissidents unless otherwise noted in the report. We also consider individuals who AI notes have been disappeared as Dissidents. Terrorists were coded as both Criminals and Dissidents. Guerrillas were coded as Dissidents only. In countries where AI notes that the government persecutes the news media, we code members of the press as Dissidents.”

90 Schnakenberg and Fariss 2014; Fariss 2014.
91 Fariss 2014.
measurement bias. Moreover, to address concerns about other information effects, the statistical model below directly accounts for factors believed to influence the production of torture allegations.

**4.2. Methods**

The ITT dataset explicitly distinguishes observed violations from their true level, recognizing that allegations do not provide an unbiased record of violations. How can we make inferences about all violations based on biased allegations data? While available information used to produce allegations varies across time and space, it is safe to assume that allegations of human rights violations are always a subset of all violations by state agents. This means that, with a theory of which factors influence human rights reporting, the measurement bias can be modeled.\(^92\) Some previous research has tried to account for the production of allegations by simply adding relevant covariates as controls to their statistical analyses. This is less than satisfactory. For one, a covariate that affects both the likelihood of an allegation and the government’s human rights practices can only be entered into the model once, and thus its estimate combines two different effects, which makes it difficult to interpret statistical results.\(^93\) More importantly, this approach still assumes that allegations of violations are randomly drawn from the underlying distribution of all violations, a highly questionable assumption.

To improve on this approach, I follow Conrad, Hill, and Moore in using a detection-controlled “undercount” model to examine the ITT event counts.\(^94\) This mixture model combines a count model for the total number of events with a binary response model which addresses the measurement error by letting the probability of detecting an event vary across all observations as a function of covariates. It assumes that the number of observed events is a proportion of the actual number of events and estimates a parameter for the probability of decreased detection of an event for each observation. In this way, inferences can be drawn about the effects of

---

\(^92\) Hill, Moore, and Mukherjee 2013; Conrad, Hill, and Moore 2018.

\(^93\) Conrad, Haglund, and Moore 2014, 434.

\(^94\) Conrad, Hill, and Moore 2018.
covariates on violations rather than on allegations.\textsuperscript{95} The coefficients from the detection equation are interpreted as the effects of the covariates on the probability of detecting an event, while the coefficients of the count equation are interpreted as the effects of covariates on the expected total number of events, observed or not.

The “undercount” model has limitations. First, the model’s performance will only be as good as the theory regarding the recording probability, i.e. the choice of covariates for the detection equation. Second, the model assumes independence of the recording and occurrence mechanisms. Here this means that it assumes that the probability of detecting a human rights violation has no influence on its occurrence; this may be a strong assumption, but relaxing this assumption would lead to a much more complicated (and difficult to estimate) mixture model.\textsuperscript{96} Finally, this model cannot account for the unobserved country heterogeneity that may be present in panel data. Fixed or random effects detection-controlled negative binomial models have not been published.\textsuperscript{97} In this analysis, I chose to model the production of torture allegations rather than any unobserved country-level heterogeneity.\textsuperscript{98}

4.3. Variables expected to affect torture event counts

De facto judicial independence is commonly conceptualized in terms of autonomy or power, although scholars question whether it is possible to measure the former without the

\textsuperscript{95} Cameron and Trivedi derive such a mixture model by assuming a negative binomial distribution with an exponential conditional mean for the occurrence of events and a logit distribution for the probability of their recording. Cameron and Trivedi 2013, 492; see also Winkelmann 2008. The parameters of this model are identified as long as the sets of covariates in the detection and count equations are not identical. Since no software package exists for implementing maximum-likelihood estimation of the “undercount” negative binomial model, I wrote software for all the analyses in the \texttt{R} statistical programming language. \texttt{R} Core Team 2016. I thank Daniel Hill and the late Will H. Moore for sharing their code of the log-likelihood function for comparison.

\textsuperscript{96} See Cameron and Trivedi 2013, section 13.5.

\textsuperscript{97} For a discussion of count models of panel data, see Cameron and Trivedi 2013, section 9.4.4. In any case, a fixed effect negative binomial model is subject to the incidental parameter problem. Regarding this and other problems with existing fixed effects negative binomial models, see Allison and Waterman 2002.

\textsuperscript{98} Future research would benefit from the derivation of a statistical model that can address both issues simultaneously.
latter. Autonomy means that the legal decisions of judges reflect only their reasoned evaluation of the legal record, and not any external influences. When judicial independence is conceived of as power, judges need to be autonomous and their legal decisions are implemented properly by the government. From the power perspective, legal decisions constrain other actors. The reason that it is difficult to measure autonomy without taking into account power is that “[e]xpectations about the compliance process can influence decision making in the sense that an inability to fully control the implementation of orders can undermine decisional autonomy.” Judicial behavior is strategic, in that courts can select and decide legal cases in light of their expected implementation, in order to avoid conflict with government actors. Thus, inferences from observed court behavior may be misleading.

De facto judicial independence is difficult to measure; I use Linzer and Staton’s latent measure, which is derived from a measurement model that treats it as an unobservable concept that can be measured with uncertainty. It operationalizes the judicial power concept by synthesizing eight commonly used measures of judicial independence and judicial effectiveness and extracting the information that is common to all the indicators to infer an underlying latent measure. The original measure ranges from 0 to 1, but I put it on a normal scale; higher values indicate more independence.

Measuring perceptions of public insecurity cross-nationally over time is difficult because regular and consistent public opinion data on questions of crime and criminal justice is not available for many countries. Instead, I use the law and order measure from the International Country Risk Guide (ICRG) available from Political Risk Services (PRS). This has previously been used as a measure of rule of law or judicial capacity. However, it is problematic as a

99 Larkins 1996; Cameron 2002; Linzer and Staton 2015. De jure independence refers to “the existence of a set of formal institutions – such as fixed budgets or cumbersome removal procedures – that are thought to provide incentives for independent judging.” Linzer and Staton 2015, 225.
100 Powell and Staton 2009; Linzer and Staton 2015.
101 Linzer and Staton 2015, 225.
102 Powell and Staton 2009.
103 Linzer and Staton 2015; for a review of the measures in the model, see Ríos-Figueroa and Staton 2014.
104 Goodliffe and Hawkins 2006; Powell and Staton 2009.
measure of legal and judicial institutions, because it conflates two distinct concepts in one index: criminal behavior by societal actors and state responses to crime. Referring to the underlying reports, regional experts note that

A closer inspection of the information in these reports, however, suggests that something other than judicial performance is being assessed. A prime example of this is the highly regarded International Country Risk Guide, used by some as a measure of a country’s “rule of law,” a concept that is then closely linked to judicial performance (...). The publisher of the ICRG, though, insists that its data were intended not to measure judicial quality but instead to tap evidence of what ICRG calls ‘law and order,’ a term suggesting law enforcement and prevention of crime more than judicial system performance.

I contend that the law and order measure is a reasonable proxy of perceptions of public insecurity, which depend on the state’s perceived capacity to constrain crime and enforce order. The ICRG reports are a commercial product marketed to multinational corporations. Like its other measures, the law and order measure is coded by PRS country experts to be used in quantitative risk assessments to guide business investment decisions. Whereas the ICRG reports also include measures of legal certainty, contract enforcement, and political violence, it is clear that the law and order measure is intended to assess business risks due to crime and social turmoil. The measure is coded “on the basis of subjective analysis of the available information.”  

These subjective assessments by PRS employees likely heavily depend on official statistics and local news media. These sources of information reflect local perceptions of the extent of crime and the state’s capacity to control crime. Indeed, Englehart uses the law and order index to measure law enforcement capacity. The measure is available monthly since 1984 and ranges from 0 to 6; I average it annually and invert it such that higher scores reflect perceptions of more public insecurity. These data are available for less countries than the ITT data, thus lowering the number of observations in the analyses.

In order to examine the effect of international human rights law, the analyses include a binary indicator of whether a country was a member of the CAT in a given year, as the allegations data measure violations specifically prohibited by this treaty. This variable is coded

---

105 Political Risk Services 2012; see also Englehart 2009, 168; Powell and Staton 2009.
106 Englehart 2009, 168.
based on state membership data from the United Nations Treaty Collection.\textsuperscript{107} In order to model potential treaty selection effects, I also include estimated CAT commitment preferences; this variable is explained in Section 4.5.

The analyses include several control variables to account for other factors influencing human rights practices. A binary indicator of nondemocracy is coded from the Polity IV dataset.\textsuperscript{108} Its polity2 variable aggregates component measures of competitiveness and openness of executive recruitment; competitiveness and regulation of political participation; and institutional constraints on executive power. Ranging from −10 to 10, it is commonly used to measure the level of autocratic and democratic institutions, with higher scores indicating more democracy. Political regimes are coded as democratic when the polity2 variable is greater than six, and as nondemocratic otherwise; this threshold is common in comparative politics and international relations research.\textsuperscript{109} An indicator of the publicly stated economic policy orientation of the political party holding the national executive office controls for ideological differences; this variable comes from the Database of Political Institutions.\textsuperscript{110} I also include several variables commonly used as controls in statistical analyses of human rights practices. The natural log of real gross domestic product (GDP) per capita (in constant US dollars) controls for a country’s level of national wealth. Population size accounts for the greater opportunity to torture more people in larger populations. These data are from Gleditsch’s expanded GDP, Trade and Population dataset.\textsuperscript{111} To account for national security threats as drivers of human rights abuses, I

\textsuperscript{107} United Nation Treaty Collection 2016.

\textsuperscript{108} Data from Polity IV Project 2016.

\textsuperscript{109} I also conducted the analyses using an alternative measure of regime type, which focuses only on the vertical, electoral dimension. See Boix, Miller, and Rosato 2012. The results are very similar to the ones presented in this paper.

\textsuperscript{110} Orientation with respect to economic policy was coded based on the following criteria: right for conservative, Christian democratic, or right-wing parties; center for centrist parties; left for communist, socialist, social democratic, or left-wing parties; no information for parties which do not have a clearly articulated orientation on economic policy. Since the no information category of this variable has the substantive meaning that there is no clear orientation articulated, this category combined with the centrist orientation is used as the omitted reference category in the analyses. Beck, Clarke, Groff, Keefer, and Walsh 2001; Cruz, Keefer, and Scartascini 2016.

\textsuperscript{111} Gleditsch 2002, 2013.
include an indicator of internal conflict, coded from the UCDP/PRIO Armed Conflict Dataset.\textsuperscript{112} This dataset distinguishes between low-level conflict (at least 25 battle-related deaths) and war (at least 1000 battle-related deaths). Finally, to control for the institutional stickiness of abuse, I include Schnakenberg and Fariss’s latent physical integrity measure (lagged one year), discussed in Section 4.1.

4.4. Variables expected to affect the detection of torture events

The detection equation includes several covariates that are expected to affect whether AI observes and reports a violation. To address concerns about information effects, these variables address arguments regarding the quantity of human rights information.\textsuperscript{113} First, I include the count of ITT events in the current year for which the type of victim was unknown or not stated, in order to account for the possibility that these mask other types of victims. For instance, human rights reports may disproportionately fail to identify certain types of torture victims.

Second, international HROs are likely to particularly turn their attention and limited resources to countries that violated human rights in the recent past. Beliefs about a state’s respect for human rights influence the likelihood that AI investigates and publishes an allegation. In order to proxy AI’s prior beliefs about conditions in a given country, I use Schnakenberg and Fariss’ latent physical integrity measure (lagged one year).\textsuperscript{114} As noted in Section 4.1, there is evidence that the ITT data is less subject to a dynamic standard of human rights accountability than other measures. Including the latent physical integrity measure as a covariate further controls for any remaining variation in the standard of accountability that may be biasing the analyses. Note that the version of the latent measure used here includes the standards-based ITT level of torture measure in its measurement model, which is distinct from, but closely related to, the specific allegations event data.\textsuperscript{115} This is useful, since AI’s beliefs will be influenced by both its own prior research in the country and other available information. Like Conrad, Hill, and

\textsuperscript{112} Gleditsch, Wallensteen, Eriksson, Sollenberg, and Strand 2002; Themnér and Wallensteen 2014; Data from Uppsala Conflict Data Program 2015.
\textsuperscript{113} Clark and Sikkink 2013; Fariss 2014.
\textsuperscript{114} Schnakenberg and Fariss 2014; Fariss 2014.
Moore, I expect that the more a state was found in the recent past to violate human rights, the more likely AI is to search for and report new violations.

Third, AI’s ability to observe violations in a country depends on the existing research and advocacy capacity in a given country. To account for in-country capacity, I include Murdie’s measure counting international HROs that have permanent offices or secretariats within a country in a given year, coded from a list of 432 international HROs.\textsuperscript{116} This measure is distinct from HRO membership data, which captures whether an organization has members or volunteers within a state. I further include a measure of international nongovernmental organization (INGO) membership ties to proxy a country’s general linkages to transnational advocacy networks; this is a count of INGOs in which citizens of a country hold memberships in a given year.\textsuperscript{117} I take the natural log of these measures to account for diminishing information returns with increasing numbers of organizations.

Fourth, HROs face varying general levels of information freedom. Government censorship undermines the ability of HROs to collect information and produce allegations, because it generally lowers the availability of public information and may make citizens less likely to provide information due to fear of government retaliation. Conrad, Hill, and Moore include the CIRI measure of freedom of speech and press,\textsuperscript{118} but this variable is not statistically significant in their models.\textsuperscript{119} Nonetheless, freedom of information and expression remains an important control variable for the detection equation. I use Whitten-Woodring and van Belle’s media freedom indicator, which measures to what extent the press “is able to hold those in power accountable.”\textsuperscript{120} This is a de facto measure; it is not focused on legal protections.

Fifth, I also include the natural log of GDP per capita in the detection equation, to account for a country’s level of wealth. While it is an imperfect proxy, per capita income is correlated with existing transportation and communications infrastructure, both of which affects

\textsuperscript{116} Murdie 2014.
\textsuperscript{117} Hafner-Burton and Tsutsui 2005.
\textsuperscript{118} Cingranelli and Richards 2010.
\textsuperscript{119} Conrad, Hill, and Moore 2018.
\textsuperscript{120} Whitten-Woodring and van Belle 2017, 2; Whitten-Woodring and Van Belle 2015.
AI’s ability to collect human rights information within a country. Finally, I include the public insecurity measure discussed in Section 4.3 to allow for the possibility that the politics of law and order influence not only human rights conditions but also human rights reporting, because nongovernmental organizations may have a harder time advocating rights for (suspected) criminals.

4.5. Treaty selection effects

The standard selection bias story in scholarship on international institutions is that the decision by states to commit to a treaty is endogenous to any interest in complying with it. In the prominent realist version of this argument, compliance is indicative of shallow cooperation rather than effective institutions, because states are more likely to commit to a treaty when they expect to comply with it. Thus, treaty commitments screen rather than constrain state behavior. Accordingly, Simmons argues that one of the most serious inferential problems in research on the effects of human rights treaties is that “with few exceptions, the same factors that lead to ratification are likely to explain compliance.” Yet, the standard selection argument is not consistent with scholarship finding high levels of abuses in member states, as it does not explain why abusive states commit to treaties that require extensive changes in policies and behavior. Therefore, some scholars have proposed theories of negative selection, in which states that are likely to continue their abuses are more inclined to make treaty commitments than other states for domestic reasons (see Section 2). The important point is that selection effects may cut both ways.

In order to address potential treaty selection effects, I modify the approach by Lupu, who uses the method of roll-call ideal point estimation common in studies of legislative choices to estimate a spatial model of latent preferences based on treaty commitment choices, and then creates statistical matches based on these preferences and analyzes the matched samples with regression (without including the estimated preferences).

---

122 Simmons 2009, 267; see also Hill 2010.
123 Vreeland 2008; see also Hollyer and Rosendorff 2011.
124 Lupu 2013; on roll-call ideal point estimation, see Poole 2005.
preferences is a useful approach to accounting for treaty membership selection to the extent that it accurately predicts ratification decisions, because it enables researchers to separate the effects of the decision to commit from the legal commitment itself, but the combination with matching is not necessary to address the issue of treaty membership endogeneity. Matching does not provide any advantage over regression alone, as it does not address the underlying problem: both methods rely on a “selection on observables” or “no omitted variable” assumption to avoid selection bias.\(^{125}\) Moreover, matching introduces other methodological choices with unclear trade-offs in the case of panel data.

I estimate annual CAT commitment preferences for each state and include these in the analyses as a control variable, but not as a basis for statistical matching. To recreate Lupu’s measure, I collected data on all state actions regarding all multilateral treaties and protocols in the United Nations Treaty Collection, and determined which of these were open to all states.\(^{126}\) On this basis, I created commitment matrices of all countries and open treaties in existence for each year, estimated the annual spatial models, and used the resulting estimates of state and treaty positions to calculate the probabilities that states commit to the individual treaties.\(^{127}\) The measure of treaty commitment preferences included in the analyses consists of the annual probabilities of commitment to the CAT.

### 4.6. Multiple imputation

I employ multiple imputation to address missing data, to use as much of the available information in the existing data as possible and to incorporate the estimated uncertainty of the latent measures derived from measurement models. With one exception, missing data are not extensive in this dataset. The ITT dataset, the INGO ties variable and the indicator of executive party economic policy orientation have missing data in less than 2% of observations, and the law and order index in less than 3%, but 16.5% of observations do not have data on HRO secretariats. Multiple imputation provides a better approach to missing data than the most

---

\(^{125}\) Miller 2015.

\(^{126}\) United Nation Treaty Collection 2016.

\(^{127}\) The spatial models were estimated in the \texttt{wnominate} package in R. Poole, Lewis, Lo, and Carroll 2011.
common alternative, list-wise deletion of observations with missing values, which violates statistical assumptions due to the likely non-randomness of missing data.\textsuperscript{128}

The final dataset has 1,419 country-year observations for 129 countries from 1995 to 2005.\textsuperscript{129}

5. Results

Table 1 presents the base model. Figure 2 shows select results for the three dependent variables: country-year counts of all AI torture allegations and counts of allegations of torture of dissidents and criminals. The first measure is included to show the aggregate effects on all victim types, and provides a baseline for comparing to the individual victim types. The figure presents results as marginal percentage changes with 95% confidence intervals. A negative effect means a decrease in torture; the confidence intervals of statistically significant estimates do not touch the zero-lines.

The results show that the aggregate findings for all victims mask differences between victim types. In this model, several key predictors of human rights protection in the literature have different statistical effects for dissidents and criminals. For instance, the results indicate that

\textsuperscript{128} Honaker and King 2010; for the \textit{Amelia} software, see Honaker, King, and Blackwell 2011.

\textsuperscript{129} The countries are: Albania, Algeria, Angola, Argentina, Armenia, Australia, Austria, Azerbaijan, Bangladesh, Belarus, Belgium, Bolivia, Botswana, Brazil, Bulgaria, Burkina Faso, Cameroon, Canada, Chile, China, Colombia, Republic of Congo (Brazzaville), Congo (DRC), Côte d’Ivoire, Costa Rica, Croatia, Cuba, Czech Republic, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Estonia, Ethiopia, Finland, France, Gabon, Gambia, Germany, Ghana, Greece, Guatemala, Guinea, Guyana, Haiti, Honduras, Hungary, India, Indonesia, Iran, Iraq, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Kuwait, Latvia, Lebanon, Liberia, Libya, Lithuania, Madagascar, Malawi, Malaysia, Mali, Mexico, Moldova, Mongolia, Morocco, Mozambique, Myanmar, Namibia, Netherlands, New Zealand, Nicaragua, Niger, Nigeria, North Korea, Norway, Oman, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Poland, Portugal, Qatar, Romania, Russian Federation, Saudi Arabia, Senegal, Serbia & Montenegro, Sierra Leone, Singapore, Slovakia, Slovenia, South Africa, South Korea, Spain, Sudan, Suriname, Sweden, Switzerland, Syria, Taiwan, Tanzania, Thailand, Togo, Trinidad & Tobago, Tunisia, Turkey, Uganda, Ukraine, United Arab Emirates, United Kingdom, United States of America, Uruguay, Venezuela, Vietnam, Yemen, Zambia, and Zimbabwe.
Table 1: Detection-controlled undercount model (without interactions).

<table>
<thead>
<tr>
<th></th>
<th>all victims</th>
<th>dissidents</th>
<th>criminals</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Count equation:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Judicial independence</td>
<td>-0.19 [0.06]</td>
<td>-0.33 [0.10]</td>
<td>-0.14 [0.09]</td>
</tr>
<tr>
<td>Public insecurity</td>
<td>-0.14 [0.05]</td>
<td>-0.31 [0.07]</td>
<td>-0.10 [0.07]</td>
</tr>
<tr>
<td>CAT ratified</td>
<td>0.02 [0.07]</td>
<td>-0.21 [0.11]</td>
<td>0.20 [0.10]</td>
</tr>
<tr>
<td>CAT commitment preferences</td>
<td>0.27 [0.09]</td>
<td>0.39 [0.14]</td>
<td>0.53 [0.12]</td>
</tr>
<tr>
<td>Nondemocracy</td>
<td>-0.10 [0.08]</td>
<td>0.64 [0.14]</td>
<td>-0.01 [0.12]</td>
</tr>
<tr>
<td>Right economic policy orientation</td>
<td>0.01 [0.08]</td>
<td>-0.01 [0.13]</td>
<td>-0.10 [0.10]</td>
</tr>
<tr>
<td>Left economic policy orientation</td>
<td>0.14 [0.06]</td>
<td>-0.09 [0.11]</td>
<td>0.14 [0.09]</td>
</tr>
<tr>
<td>Latent physical integrity</td>
<td>0.15 [0.08]</td>
<td>0.41 [0.17]</td>
<td>-0.14 [0.12]</td>
</tr>
<tr>
<td>Internal conflict</td>
<td>-0.06 [0.09]</td>
<td>-0.01 [0.15]</td>
<td><strong>-0.26 [0.13]</strong></td>
</tr>
<tr>
<td>Internal war</td>
<td>0.15 [0.13]</td>
<td>-0.13 [0.22]</td>
<td>0.05 [0.18]</td>
</tr>
<tr>
<td>Population size</td>
<td>0.13 [0.02]</td>
<td>0.19 [0.04]</td>
<td>0.16 [0.03]</td>
</tr>
<tr>
<td>GDP per capita</td>
<td>0.18 [0.07]</td>
<td>0.26 [0.13]</td>
<td>0.23 [0.10]</td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.19 [0.73]</td>
<td><strong>-3.47 [1.23]</strong></td>
<td><strong>-2.86 [0.97]</strong></td>
</tr>
<tr>
<td><strong>Detection equation:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Torture events: victim not stated</td>
<td>-0.53 [0.15]</td>
<td>-1.47 [0.36]</td>
<td>0.18 [0.24]</td>
</tr>
<tr>
<td>Latent physical integrity</td>
<td>0.59 [0.06]</td>
<td>0.45 [0.15]</td>
<td>0.49 [0.07]</td>
</tr>
<tr>
<td>HROs secretariats</td>
<td>-0.05 [0.08]</td>
<td>0.19 [0.15]</td>
<td>-0.14 [0.12]</td>
</tr>
<tr>
<td>INGOs membership ties</td>
<td>0.36 [0.06]</td>
<td>0.49 [0.15]</td>
<td>0.43 [0.08]</td>
</tr>
<tr>
<td>Free media</td>
<td>-0.32 [0.11]</td>
<td>-0.78 [0.24]</td>
<td>0.17 [0.18]</td>
</tr>
<tr>
<td>Imperfectly free media</td>
<td>0.15 [0.09]</td>
<td>0.34 [0.22]</td>
<td>-0.11 [0.14]</td>
</tr>
<tr>
<td>Public insecurity</td>
<td>-0.09 [0.11]</td>
<td>-0.41 [0.30]</td>
<td>-0.11 [0.17]</td>
</tr>
<tr>
<td>GDP per capita</td>
<td>0.42 [0.11]</td>
<td>-0.04 [0.30]</td>
<td><strong>0.90 [0.17]</strong></td>
</tr>
<tr>
<td>Intercept</td>
<td><strong>-1.99 [0.90]</strong></td>
<td>-0.10 [1.92]</td>
<td>0.33 [1.36]</td>
</tr>
<tr>
<td>observations</td>
<td>1419</td>
<td>1419</td>
<td>1419</td>
</tr>
<tr>
<td>countries</td>
<td>129</td>
<td>129</td>
<td>129</td>
</tr>
</tbody>
</table>

Note: Estimates [standard errors]; significant at 0.05 alpha level in bold.

Nondemocracies clearly torture dissidents more than democracies (the reference category), as one would expect, but this is not the case for criminals or the aggregate counts of all torture events. According to this evidence, there is no statistical difference in torture between democracies and
nondemocracies except for dissidents. By contrast, judicial independence is associated with a decrease in torture for dissidents and all victims, but not for criminals. This provides initial evidence that judicial protection against torture benefits dissidents but not criminals, as proposed by H1, which implies that the effect of judicial independence is negative and statistically significant for dissidents and the effect for criminals is smaller in magnitude or not statistically significant.

Figure 2: Base model.

CAT commitment is associated with decreased torture of dissidents as expected, although the estimate narrowly misses statistical significance. By contrast, treaty commitment is associated with increased torture of criminals. I do not argue that CAT membership leads to torture of criminals. Rather, this latter finding is possibly due to a selection effect, which needs to be investigated further. The results for the latent measure of treaty commitment preferences
show that states that are likely to ratify the CAT are also more likely to subsequently torture all victim types, but actual treaty commitment decreases torture of dissidents and increases torture of criminals. This lends some support to a negative selection explanation whereby countries make human rights commitments while being unable or unwilling to stop violations, but once they make the commitment, they treat dissidents better than criminals.

Public insecurity is associated with decreased torture of dissidents (and all victims) and has no association with torture of criminals. While this finding seemingly cuts against the proposed theory, it is important to note that the theoretical claim is not that public insecurity increases abuses directly but that it undermines human rights change and particularly judicial protection against abuses. In order to test the hypotheses of conditional effects, the figures below present results of additional models with two sets of interactions.\(^{130}\)

H2 implies that the interaction between judicial independence and public insecurity is associated with reduced torture at low levels of public insecurity but has no statistical effect at high levels. Figure 3 shows the effects of judicial independence across all levels of the public insecurity variable for both dissidents and criminals; it shows the marginal percentage changes due to a one-unit change in judicial independence. The thicker middle lines are the estimates and the thinner outer lines show the 95% confidence intervals. A positive slope indicates that improvements in torture associated with judicial independence decrease with increasing public insecurity; a flat line would mean that this effect is not conditional on public insecurity. The dashed vertical lines show the level of public insecurity above which judicial independence no longer has a statistical effect on torture. The result for dissidents suggests that judicial independence decreases torture at low and intermediate levels of public insecurity but not at high levels (more than 3 on the scale of 0–5.5). For criminals, this decreasing effect only holds at very low levels of public insecurity. Thus, both H1 and H2 are supported, because judicial independence has no effect in the face of high levels of public insecurity, but it is associated with decreased torture at higher levels of public insecurity for dissidents than for criminals.

\(^{130}\) The complete regression table for these analyses will be included in the online appendix accompanying publication.
To further examine the role of public insecurity in human rights change, a triple interaction of judicial independence, public insecurity and CAT membership status tests whether treaty commitment enhances the effect of judicial independence in the face of public insecurity. H3 implies that for CAT members, judicial independence is associated with decreased torture at
higher levels of public insecurity for dissidents than for criminals. The analyses provide clear supporting evidence. The estimate of the triple interaction for dissidents is statistically significant, while the one for criminals is not, which means there is no statistical difference in the interaction effect of judicial independence and public insecurity between treaty members and non-members for criminals.

Rather than present the estimates in a table, Figure 4 again shows the marginal percentage changes, but broken down by treaty membership status for both types of torture victim. In the upper panel, regardless of whether or not a state has committed to the CAT, judicial independence is associated with decreased torture of dissidents at low levels of public insecurity. The effect line for treaty members is flatter than for non-members, which indicates that judicial independence still has a torture reducing effect for dissidents at intermediate levels of public insecurity. This is consistent with the claim that treaty commitment enhances judicial protection, but at very high levels of public insecurity, judicial independence is not associated with improvements even for dissidents. By contrast, in the lower panel there is little difference in effects on torture of criminals between CAT members and nonmembers, except that judicial independence is associated with a decrease for treaty members only at a very low level of public insecurity.

Finally, I briefly note the results from the detection equation, shown in Table 1 in the appendix. Torture events for which the victim is not stated are associated with less reporting of torture of all victims and of dissidents, but has no association with the detection of torture of criminals. It could be that dissidents are often not identified as such in reports of human rights. The result for criminals suggests that the victim not stated category is not systematically used for torture of criminals. As expected, prior physical integrity violations raise the likelihood that AI will report new allegations across all victim types. INGO ties increase the likelihood of detection across victim types, but the presence of in-country HRO secretariats has no statistical effect on detection. Counter to expectations, more media freedom decreases the detection of violations, except for criminals. Per capita wealth increases the detection of violations of criminals but not

---

131 The Appendix is not included with this Working Paper.
Figure 4: Effect of judicial independence by public insecurity and CAT commitment, (a) for dissidents and (b) for criminals.

(a) Dissidents

(b) Criminals
of dissidents. As discussed in Section 4.2, the theory regarding the detection probability is crucial to the model’s ability to account for information effects. Thus, the fact that several of the included covariates have the expected associations provides some confidence in the validity of the detection model. However, the results for the media freedom measure are surprising, since they suggest that free media detect less violations, except for those of criminals. This may be the case because human rights organizations are particularly motivated to report violations of dissidents when freedom of expression is restricted.

6. Conclusion

The “undercount” model of allegations of CAT violations against dissidents and criminals provides support for the theory. Under what conditions does judicial independence reduce torture? Judicial protection against torture is stronger for dissidents and at low levels of public insecurity and decreases with increasing public insecurity for both types of victim. According to these findings, a key mechanism of human rights change primarily benefits dissidents, compared to criminals. Moreover, CAT commitment is associated with less torture of dissidents and more torture of criminals, but the latter finding should not be used to conclude that CAT commitment leads to more torture of criminals. Importantly, CAT commitment enhances judicial protection only for dissidents.

The analyses try to account for selection effects by including the estimated treaty commitment preferences. The association between treaty commitment and increased torture of criminals is likely due to selection into the treaty by states that are prone to continue torturing. This interpretation is partially supported by the statistical effects of the treaty commitment preference and treaty membership variables on torture of criminals. There may be other unobserved differences between countries leading to selection effects, and these should be investigated in further research. Nonetheless, the analyses here allow one to infer relative differences in effects across the targets of violations. Country-level selection is constant across different societal groups within countries, and does not pose an inferential threat to assessing whether certain societal groups experience better protection than others. The key contribution of this study is to show that CAT commitment is associated with improved protection from torture.
for dissidents but not for criminals. Moreover, the analysis shows that public insecurity undermines judicial protection even for dissidents. These findings are in line with the theory outlined in Section 3.

By using the ITT data and modeling the production of torture allegations, this study directly addresses important concerns about measurement bias. This approach to human rights events data is promising for future research but the model’s performance is only as good as the underlying theory of human rights reporting. Further theorizing should improve the detection model. For instance, new data on organizational factors and resources affecting AI’s and other HROs’ decision-making regarding research and reporting would help.

The analyses support the key theoretical claim that international law affects the human rights protection of different societal sectors in different ways and that public insecurity undermines judicial protection of human rights. The findings on the role of treaty commitment are noteworthy because they suggest nuanced effects on human rights protection: CAT membership does enhance judicial protection against torture, but this impact is limited and does not apply to all societal groups. This study only distinguishes between dissidents and criminals. Further investigations of treaty impacts should take seriously the notion that domestic human rights enforcement mechanisms may have heterogeneous impacts not only between countries but also within them across societal groups.

It is a core principle of international law that human rights are universal: they apply to all people. However, this study reveals that certain social groups may be excluded from the benefits generated by key mechanisms of human rights change. By revealing disparities in physical integrity protection among social groups within states, this research sheds new light on the gap that persists between the principles and the practice of human rights. Future research should further examine the workings and limits of human rights change. What about other societal groups, such as marginalized communities? When do judicial institutions protect the human rights of particular groups and when do they not? Such research will require new disaggregated and more fine-grained human rights data.
References


