The Psychological Effects of State Socialization: IGO Membership Loss and Respect for Human Rights

Gina Lei Miller  
*George Fox University*  
gmiller@georgefox.edu

Ryan M. Welch  
*University of Tampa*  
r.merrill.welch@gmail.com

Andrew J. Vonasch  
*University of Canterbury*  
andy.vonasch@gmail.com

July 18, 2018

Word Count: 11,940

---

1We wish to thank Babak RezaeeDaryakenari, Thorin Wright, and the ASU SPGS Workshop for helpful comments while writing this draft. The authors also received support from School of Politics and Global Studies at Arizona State University, Department of Political Science at Florida State University, Department of Political Science at University of Alabama, and Department of Psychology and Neuroscience at University of North Carolina, Chapel Hill while writing this paper.
We present an interdisciplinary theory that considers how loss of membership in international organizations affects states’ human rights practices. Drawing mostly from social psychology and international relations research, we argue that states are socialized into the international community through a process of social influence, whereby they are incentivized to comply with group norms by the promise (threat) of social rewards (punishments). Social influence occurs when states form social bonds through interactions with other states. When social bonds are severed, fewer opportunities for social influence occur due to lower information to both the remaining states and the state that lost those social bonds. Thus, we hypothesize that the loss of membership from IGOs reduces incentives to comply with group norms and adversely affects human rights practices at home. A combination of propensity score matching/regression and autoregressive distributed lag models on a global cross-section across the years 1978 to 2012 supports the theory. Specifically, losing at least one IGO membership leads to a long-run drop in human rights respect of about one quarter to one half standard deviation.
How does losing international governmental organization (IGO) membership affect state behavior? Past scholarship highlights the material benefits states derive from membership in various IGOs such as access to information, decreased transaction costs, and improved cooperation and coordination (Keohane 1984; Snidal 1985). We take a different approach, focusing on how social incentives affect state behavior, namely in the protection of human rights. By social incentives, we refer to the array of rewards and/or punishments that groups of states can use to incentivize compliance with pro-social behavior (Johnston 2008, 24-25), or behavior that is aligned with the group's normative expectations regarding members’ human rights practices. Examples of these social incentives include greater decision-making power within the group as a reward for normatively appropriate behavior or reduced status relative to other member states as punishment for counter-social behavior.

Like many other IR scholars, while building our argument, we find it useful to treat states as unitary actors (Keohane 1984; Waltz 1979). While many scholars that do so focus on states’ material interests, we take seriously states’ psychological preferences as well (Wendt 1999). Doing so allows us to view IGOs as forums where member states take part in a process of social bonding. As part of this process, states influence other member states within the IGO to act according to certain standards of behavior deemed appropriate for state identity. States receive not only material rewards for their compliance, but also social rewards such as increased status and prestige within the IGO. By gaining status and prestige within the IGO, states gain legitimacy within the broader international community and are subsequently better positioned to shape policy and pursue their interests in the international arena (Meyer, Boli, Thomas and Ramirez 1997; Wotipka and Tsutsui 2008).
When states lose the social bonds cultivated through IGOs, however, not only do they lose out on the material incentives that come with membership, but they also lose a valuable means of social influence that can incentivize normatively appropriate behavior. But, whereas states acting completely rationally may overdo their appropriate behavior to signal their social desirability, psychology literature suggests agents that lose these social bonds retract from the group and act less appropriately. If states act as psychological agents (Wendt 2004), we would expect a loss of social bonds to negatively impact states’ normative behavior, such as human rights practices. This is exactly what we find.

Importantly, states do not have to internalize IGO norms to follow normatively appropriate behavior like protecting human rights. While accounting for psychological needs, we still assuming states behave rationally – those that wish to receive the rewards that come with being a member of the organization will adapt behavior to reflect the group’s normative expectations and maintain access to the benefits of membership. If they fail to do so, states risk material punishments (for example, sanctions), as well as social punishments that jeopardize the rewards that come from shaping the international order by being in good standing within the IGO. Social punishments such as shunning may result in a loss of status or prestige within a group (Johnston 2008), and this damaged social standing may affect a state’s ability to enjoy the exclusivity and diplomatic platform that accompanies membership in an IGO (von Borzyskowski and Vabulas 2018).

We argue that these social incentives have a direct impact on human rights practices. The norm that human rights should be protected has proliferated along with IGOs since the end of World War II. As individuals and nongovernmental organizations promoted the idea of a universal set of human rights that all states would be expected to protect, states responded to these
domestic and international constituencies by adopting the protection of human rights as a goal in the common interest of all states. The legitimacy of this norm within the international community can be evidenced by the inclusion of human rights clauses in international agreements ranging from trade and labor to environmental conservation.

In this article, we examine the implications of losing these incentives for states’ human rights records. Applying insights from social psychology we expect states with severed social bonds to act counter to prevailing norms – in this case the international norm to protect human rights. When a state loses social bonds, information decreases to 1) the remaining states about the state that lost those social bonds, and 2) to the state that lost its social bonds with the other states. The remaining states can no longer monitor the state who lost its membership, thus it cannot give social rewards or punishments for human rights behavior. Additionally, the state who lost membership no longer has access to scripts that shape proper behavior and the social consequences of that behavior. A loss of IGO membership nicely captures this mechanism at work. Thus, we hypothesize that states that lose IGO membership will abuse human rights more often than those that do not lose IGO membership.

We use two statistical strategies to test our hypothesis. First, given that states do not lose IGO membership at random, we preprocess the data using propensity score matching to compare the effects of IGO membership loss on similar types of states. Next, we leverage an autoregressive distributed lag (ADL) model to glean longer-term effects of social bond loss. The two-fold strategy allows for stronger causal claims and estimated effects over time, as well as allaying concerns of model dependency. The matching/regression strategy allows us to identify that IGO membership loss leads to worse human rights behavior. The ADL models allow us to
estimate the long run effects of IGO membership loss of a 0.38 standard deviation decrease in human rights respect.

This paper contributes to the international relations literature in three important ways. First, our approach is consistent with the view that rational and social theoretical frameworks are compatible — rational states concern themselves not only with their material fate but with their socio-political situation as well (Fearon and Wendt 2002; Long and Hadden 1985; Thies 2003, 2010; Zürn and Checkel 2005). Unlike others who have tested empirical implications of increased socialization within this context (for example, Greenhill 2010; Schimmelfennig 2005), the focus here is on the effects of losing social bonds within the international community. Although World Polity theory (Meyer et al. 1997) suggests a gradual homogenization of state behavior over time, this paper seeks to explain why states act contrary to the liberal social order.

Second, this paper adds to the discussion of how institutional constraints affect states’ human rights practices. At the risk of oversimplification, one could view the field progressing in a somewhat linear fashion. Scholars held different views regarding the effectiveness of the international human rights regime (for example, Henkin 1979; Mearsheimer 1995). These disparate views gave way to empirical testing\(^2\) that produced discouraging results — specifically, there is little evidence that international institutions, on their own, exert much influence over states’ human rights practices (Hafner-Burton and Tsutsui 2005; Hathaway 2002). Further theorizing and testing led to the current consensus which takes one of two forms. The first highlights the importance of the domestic legal and political landscape leading to conditional successes for international institutions (Dai 2014; Simmons 2009). The second focuses on more direct material effects of international institutions, such as foreign aid (Lebovic and Voeten 2009) or human

\(^2\)Arguments still occur (for recent examples, see Dancy and Fariss 2017; Posner 2014).
rights clauses in trade agreements (Hafner-Burton 2005). Our work adds to the second in offering another direct effect of international institutions, albeit social rather than material.

Third, our theory and findings speak to why states behave according to social prescriptions (Finnemore and Sikkink 1998; Goodman and Jinks 2004). The international socialization literature provides evidence that social connections can influence state behavior, but less is known about why and how this process works. For example, do states alter their behavior as a result of internalizing international norms, or are they acting as expected to avoid censure? Our theory and results suggest that the latter process is at work: states, like individuals (Baumeister and Leary 1995; Sampson and Laub 1990), need contact with others (that is, socialization) to continue socially acceptable behavior. By viewing international relations as a social process that occurs within a community of states (Bull 2002; Meyer et al. 1997; Wendt 1999), we recognize the potential insights of social psychology research that investigates the effects of losing social bonds on agent behavior (for example, Baumeister, DeWall, Ciarocco and Twenge 2005). This approach adds to our understanding of why states abuse human rights and if the social processes that take place within international institutions influence human rights practices at the domestic level.

**International Institutions and State Behavior**

After World War II, the Allies established an international order that rested on liberal political and economic principles. Within this liberal regime, rational, purposive states create, design, and join international institutions to reap material benefits (Koremenos, Lipson and Snidal 2001). International institutions reduce the transaction costs of interstate relations by
establishing a priori rules (Keohane 1982). Instead of renegotiating interstate relationships every time cause to interact occurs, states can establish reasonable expectations of how to conduct their business with each other, thereby allowing for more efficient relationships (Krasner 1982).

Institutions also increase transparency, which in turn reduces uncertainty and ultimately increases the efficiency of interstate relations (Kratochwil and Ruggie 1986; Mitchell 1998).° If states violate the rules, the institutions can enact material costs themselves. For example, the World Trade Organization’s dispute settlement body can issue monetary sanctions to states that violate the free trade regime established by the organization. Skeptics argue that states can ignore these institutions at will given the lack of central enforcement (Mearsheimer 1995), but member states can self-enforce institutional rules by threatening reciprocity or retaliation against violators (Goldstein and Freeman 1990; Morrow 2007).

A central aspect of the prevailing liberal international order includes the protection of human rights. Enforcing the international human rights regime presents special challenges. The self-enforcement mechanisms — reciprocation and retaliation — do not lend themselves to governing state action that occurs solely within sovereign borders. Current scholarship supports this — when international institutions affect human rights practices, they rely on conditional factors. For example, trade agreements lead to better human rights, but only if they integrate strong human rights clauses (Hafner-Burton 2005). More often, scholars find conditions in domestic factors (Dai 2005; Simmons 2009). For example, domestic institutions such as elections (Davenport and Armstrong 2004; Henderson 1991; Poe and Tate 1994) and electoral rules (Cingranelli and Filippov 2010) can increase the political costs of abuse. Independent judicial

---

°Although see Barkin (2015) on transparency’s negative effects in international relations.
systems (Keith 2002; Mitchell, Ring and Spellman 2013), constitutional provisions (Ordeshook 1992; Vanberg 2005; Weingast 1997), and national human rights institutions (Cole and Ramirez 2013; Welch 2017) can increase the legal repercussions of human rights violations. The search for conditional mechanisms stems from privileging material costs and benefits. However, we join other scholars who focus on social mechanisms (for example, Bearce and Bondanella 2007; Greenhill 2010; Pevehouse 2002a; Thies 2010; Wendt 1999) given the expectation that international institutions can socialize their members (Johnston 2008).

State Socialization

International institutions result from and disseminate the culture of the international community (Boli and Thomas 1997; Meyer et al. 1997), and member states commit themselves to cooperate toward goals or behave in ways that are deemed appropriate as specified by this community (Johnston 2008, 14). By committing themselves to common goals — and the institutional rules and expectations set in place to advance those goals (Bull 2002) — states are exposed to social interactions that can change preferences and behavior.

Johnston (2001) identifies two processes through which states may be influenced to change behavior: persuasion and social influence. Persuasion is the process through which states internalize group norms. The factor motivating pro-social behavior through this process is a desire to do what is deemed "good" or "appropriate." States that are persuaded to hold certain values deem the rules that uphold these values as legitimate and worthy of being followed (Hurd 1999, 381). Consequently, they adjust attitudes and action willingly, even when explicit material or psychological sanctions are absent (Johnston 2001, 496).
Social influence, on the other hand, refers to the process of states complying with pro-social behavior in response to group-based social incentives. The factors motivating pro-social behavior include rewards such as status maximization or avoidance of punishments such as shaming or shunning. The process of social influence works independently of altruistic- or material-based incentives; instead, states’ social interactions induce them to conform to pro-social behavior in pursuit of social "status markers" or in avoidance of social costs. Notably, social influence depends on a shared understanding of what acceptable behavior looks like (that is, specific rules, goals, or expectations); otherwise, rewards such as public praise or punishments such as public criticism will not have the same behavioral effect (Johnston 2001).

These processes of socialization are not easy to observe, but both can lead to a convergence of behavior within international institutions. Within any given IGO, some states may take steps to protect human rights because this norm is considered right and appropriate (persuasion), while other states may be motivated to protect human rights in an effort to look legitimate and maintain social status by doing what the rest of the community expects them to do (social influence) (Wotipka and Tsutsui 2008, 736). The process of social influence — whereby states care about social rewards or the avoidance of social punishments — most closely fits both rational and social theoretical frameworks, as it recognizes the importance of social influence within international institutions but also acknowledges that states are motivated by material interests. For example, as states align behavior with the rest of the community in response to social incentives, they may gain approval and acceptance from other member states, which could translate into increasing influence within the group. States may then use their influence to facilitate cooperation toward important domestic or international objectives, to take the lead on
decision-making over economic or security issues, or to shape normative priorities of society (Meyer et al. 1997).

Once embedded in the community, states will continue to align behavior with normative expectations in order to be seen as a legitimate member and continue deriving gains from membership (Goodman and Jinks 2004; Wotipka and Tsutsui 2008). The social bonds created within the community are an important part of this process of social influence, as they promote a sense of shared identity among group members. As states come to self-identify with an organization, the group becomes an increasingly legitimate audience whose criticism and censure states seek to avoid (Johnston 2001, 501). When the social bonds that promote this sense of shared identity are severed, states are cut off from the social incentives that motivate behavior and essentially lose their reason for playing by the rules in the first place (Baumeister et al. 2005).

To summarize, we argue that one way in which international norms change state behavior is through social influence. Our argument is predicated on the idea that social influence often exerts similar pressures on states as it does on individual people. Whereas scholars have long debated whether treating the state as a unitary actor with agency appropriate,\(^4\) we find it a useful simplification in this case for modeling our argument about state behavior — specifically human rights respect. Many who view the state as agent usually focus on its propensity to seek material goals such as security and economic gains. But, states also seek psychological goals such as legitimacy and recognition (Wendt 2005). Therefore, we explore the theoretical insights the psychological literature offers on social influence in people to inform our model of social influence on state behavior.

\(^4\)For those scholars who treat the state as a unitary agent, see Keohane (1984); Waltz (1979); Wendt (1999). For those critical of this approach, see Gourevitch (1996); Milner (1997); Putnam (1988).
Social Influence and Human Rights Practices

Individuals are highly motivated to retain connections with other individuals given the many benefits connection offers (increased security, wealth, and status, for example). Therefore, they attempt to conform to social expectations so as to maintain their relationships with others (Hirschi 1969). Individuals are willing to make tremendous sacrifices — including death — to protect their reputations and thereby avoid losing social bonds (Vonasch, Reynolds, Winegard and Baumeister 2017). Behavioral experiments support the idea that the motivation to conform and obey is particularly strong, as many people are willing to blatantly lie or even administer lethal electric shocks if group norms or authority figures demand it (Asch 1956; Milgram 1963).

Furthermore, the more tightly a person is embedded within a group, the more he is incentivized to adhere to its values (Cialdini and Goldstein 2004; Gelfand, Raver, Nishii, Leslie, Lun, Lim and Aycan 2011). For example, an individual with more relationships to other people in his community, a good job, and long-standing membership in his church is likely to incur greater social costs for an extra-marital affair than someone without these social connections. The individual with more social bonds would likely consider all that is at stake before engaging in a relationship others in his social group would deem inappropriate, while the socially disconnected individual has less at stake in terms of these costs.

Like individuals, states enjoy substantial benefits from their connections via international institutions. States can improve their economies through trade, increase security through mutual protection and alliances with other states, and gain status and legitimation (Hall 1997; Keohane 1984; Krasner 1982; Kratochwil and Ruggie 1986; Meyer et al. 1997; Mitchell 1998). We argue that states, like individuals, also consider social bonds before deviating from acceptable
behavior that could sever those bonds and end the benefits they provide (Goodman and Jinks 2004).\(^5\)

One of the international community’s foremost values is the protection of human rights (Risse-Kappen, Ropp and Sikkink 1999; Wotipka and Tsutsui 2008). States that engage with the international community are expected to uphold the shared, legitimate international norm to respect the rights of their citizens. States that repress the rights of citizens risk losing international political legitimacy. Despite the protestations of other international actors, however, many states do repress citizens in order to protect and support political power (Poe 2004). By cracking down on dissenting citizens, states can maintain power — however, they also open themselves up to international criticism and risk losing face within the global community. The more embedded states are within this community, the more costly it becomes to violate its human rights standards. As a result, we expect that states with more social bonds to the international community through IGOs will be less likely to repress citizens in order to preserve these social bonds, remain in good standing within the IGO, and continue to reap the rewards of membership (Greenhill 2010).

**Losing Social Bonds**

The international socialization literature largely affirms the expectation that social bonds within the international community matter and can influence state behavior (Finnemore and Sikkink 1998; Goodman and Jinks 2004; Greenhill 2010; Thies 2010). But what happens when

---

\(^5\)As mentioned earlier in the paper, the assumption that states possess qualities akin to individuals underlies much IR work. It is a useful move to explore the behavioral implications of psychological and economic theories with respect to states. For explicit arguments on the subject, see Cederman (1997); Wendt (1999). For implications of the argument, see McGraw and Dolan (2007).
these bonds are severed? Although the consequences for losing social bonds is not well studied in political science, social psychology offers clues about what to expect. The international socialization literature shows states are social agents. Of course, people are also social agents (Baumeister and Leary 1995). When individuals lose social bonds, research shows they exhibit impaired self-regulation (Baumeister et al. 2005) and reduced pro-social behavior (Twenge, Baumeister, DeWall and Ciarocco 2007). Might states exhibit impaired self-regulation and reduced pro-social behavior when it comes time to make decisions about repression?

We expect losing social bonds will negatively affect state behavior by decreasing information, both to the other states in society and to the state who has lost the bonds. First, international institutions increase transparency and information about state actions. Although, those thinking of international institutions with respect to international economy show this can lead to more efficient interaction between states (Keohane 1984), it can also allow states to monitor each others’ actions (Mitchell and Hensel 2007). This monitoring allows members of society to dole out social rewards, such as legitimation (Johnston 2001); or social punishments, like naming and shaming (for example Murdie and Davis 2012). Losing social bonds decreases the monitoring that can produce the social rewards and punishments that incentivize compliance with group norms. As a result, compliance with norms becomes less beneficial. States, like individuals (Bateson, Nettle and Roberts 2006; Ernest-Jones, Nettle and Bateson 2011), act more appropriately under watching eyes (Rejali 2007).

While we do not consider the source of lost bonds in this study (that is, voluntary vs. involuntary), evidence from psychological experiments on the effects of social exclusion leads us to expect that lost bonds that result from states being kicked out of an IGO may cause states to interact more violently with citizens (e.g. Underwood 2002). Future studies should explore the source of lost social bonds.
Secondly, losing social bonds decreases information available to the state losing the bonds regarding norms and the expectations associated with them. Through international institutions, states regularly (both explicitly and implicitly) discuss policy issues and challenges with other affiliated states and receive feedback as to how others in the international community will view their actions. Like individuals (Bateson, Nettle and Roberts 2006), states collect information about other states’ behavior when deciding how to handle future interactions and when deciding how other states will handle future interactions with them. For example, states may observe the noncompliance — and subsequent punishment — of fellow member states or hear negative discourse about states that violate norms, thereby deterring similar behavior. Losing social bonds decreases access to this type of information and subsequently increases the state’s probability of behaving counter to prevailing norms.

To summarize, we expect that losing social bonds decreases information to both the community and the state who loses those bonds. Decreased information about the state’s behavior decreases the community’s ability to give social rewards and punishments, thereby making pro-social behavior less likely. Losing social bonds also decreases information to the state who lost those bonds. Without access to discussions, excluded states lose valuable opportunities to update their beliefs about current normative scripts that lay out appropriate behavior and the social rewards/punishments associated with that behavior. IGOs represent particularly relevant forums through which states form social bonds, and so we turn next to how IGOs fit into this process.
**IGOs as Social Forums**

Similar to the ways in which individuals become socialized in groups, states embed within international society by interacting with other states in groups. In so doing, they exchange both policy and normative ideas (Schmidt 2008, 305). States may embed themselves in the global community in more or less formal ways with varying degrees of contact (Bull 2002, 12). Less formally, states may organize in informal groups (for example, G groups) or hold impromptu conferences to exchange ideas about a specific policy or problem (Vabulas and Snidal 2013). More formally, states join international governmental organizations, giving them membership in a group with shared interests and goals. We focus here on the more easily observed and measured membership in formal international governmental organizations (IGOs).\(^7\)

In addition to being easier to observe and measure, highly formalized organizations more effectively define legitimate behavior (Zucker 1977, 1983). These organizations represent shared interests, and they stipulate rules which identify behaviors that sustain or advance these interests (Bull 2002). Since World War II, states have relied more heavily on these formal institutions to organize international life, resulting in the management of banal, everyday interactions as well as more dramatic crises (Abbott and Snidal 1998; Meyer et al. 1997). States create and join IGOs to make interaction more efficient, including reducing transaction costs and increasing information sharing (Keohane 1984). In this way, IGOs act as arenas where states can achieve more desirable outcomes predicated on efficient interaction (Moravcsik 1991).

\(^7\)We follow Abbott and Snidal (1998) in defining IGOs as organizations of more than two states with a formal structure such as a secretariat, rules for admission, and regular meeting times.
But, IGOs also act as social environments where states become socialized into world culture as they interact and share information with each other (Johnston 2001; Meyer et al. 1997). Participating in IGOs acculturates states, allowing them to shape and learn the dominant norms associated with the global community (Goodman and Jinks 2004; Meyer et al. 1997). While values such as the protection of human rights are often pioneered by actors outside of IGOs (for example, transnational advocacy networks or nongovernmental organizations), IGOs play an important role in endorsing and legitimizing these norms (Bull 2002; Franck 1990; Hurd 1999). Citizens and groups concerned with promoting norms may pressure state actors to take action consistent with normative goals; one way states can respond to these constituencies is to codify normative expectations into the rules that govern the behavior of states within IGO communities. By doing so, states within the IGO tie their own hands to play by these rules (thereby deeming them a legitimate interest held by member states) and also commit future members to uphold the same standards of behavior. States that wish to gain admittance to or remain in good standing with the IGO are expected to play by these rules.

Salient international norms like the protection of human rights are not only relegated to issue-specific organizations. For example, financial institutions like the World Bank increasingly consider the human rights record of states applying for aid (Abouharb and Cingranelli 2006), and environmental organizations like the International Union for the Conservation of Nature focus on the rights of marginalized populations with respect to climate change (IUCN 2016). So even if states create, join, and operate IGOs for their own rational interests such as increased trade or financial assistance, they become exposed to cultural scripts that shape their belief about appropriate behavior (Johnston 2001) and reap the benefits of membership discussed above. Even more illiberal IGOs will consist of states that share information with each
other about what actions they believe other, more liberal, states deem appropriate. Losing membership, whether voluntarily or not, severs states’ social bonds to the international community, thereby reducing the social incentives to comply with community norms and rules.

The recent economic and political crisis in Venezuela provides an illustrative case. In early 2017, Venezuela signaled its intent to leave the Organization of American States (OAS) (BBC 2017). At the same time, as food insecurity and deaths from malnutrition intensified throughout 2017, Venezuela’s ruling United Socialist Party has refused humanitarian aid from neighboring states, the international community, and nongovernmental organizations like the Catholic Church (Kohut and Herrera 2017).

Notably, in response to Venezuela’s announcement that it would withdraw from the OAS, some member states expressed concern about Venezuela’s decision to leave. Uruguay’s president commented: "We do not think Venezuela should be isolated… It needs to be offered a hand" (BBC 2017). When Venezuela announced its intent to withdraw from the OAS, it had already been suspended a few months before from the South American trade bloc Mercosur over violations of the group’s bylaws, specifically its failure to incorporate trade and human rights rules into its national law. Venezuela’s foreign minister was barred from participating in a Mercosur meeting shortly after the suspension. The effect of being excluded was apparent when the foreign minister attempted to "gate-crash" the meeting (BBC 2016). As the Economist (2018) remarked of the incident, "That it [ostracism] hurts the regime’s pride was clear”.

Our theory suggests that this social punishment (suspension from Mercosur and exclusion from the meeting) led to Venezuela’s non-cooperative behavior. Instead of adopting pro-social behavior (that is, incorporating trade and human rights rules into its national law) to regain good standing in the organization, Venezuela has repudiated community norms. This can been
seen through its rejection of humanitarian assistance to help those within its borders in dire need of food and medicine. Its suspension from Mercosur and exclusion from the organization's meeting was followed by the announcement that it would withdraw from the OAS. As Venezuela weighs its response options to the opposition, its no longer bound by the organization's social expectations.

To summarize, we argue that the focus on material constraints for influencing states' behavior toward human rights does not fully consider social processes that may be at work (Thies 2010). We suggest that international institutions, while also functioning as forums for influencing behavior through material rewards and punishments, exert an additional social influence that induces compliance with international norms and standards of behavior. As such, we allow social psychology to guide our thinking about state normative behavior. Since the end of World War II, the international norm proscribing abuse of human rights has become a primary prerequisite to state legitimacy. When membership in international society is lost, information decreases to the states in international society and to the state that lost membership. In the first instance, other states cannot monitor the excluded state's behavior, and thus cannot socially reward or sanction that state. The opportunities for social influence to occur are reduced, states no longer reap the rewards of inclusion and so lose the incentive to continue to act in normatively desirable ways (that is, protecting human rights). Secondly, lost social ties decrease information to the excluded state which decreases the ability for the state to keep current with normative discussions and scripts that inform proper (improper) behavior and social rewards (punishments) for that behavior. A way we can observe this loss of social ties is when states lose IGO membership. For these reasons, we produce the following testable hypothesis:
**Hypothesis 1.** All else equal, when states lose membership in international governmental organizations (IGOs), they engage in more repressive behavior.

**Research Methods**

Testing the hypothesis requires measures of state respect for human rights, as well as membership in IGOs. We create a dataset with these variables as well as a battery of controls listed below. We first preprocess the data using matching methods. Doing so allows us to more closely simulate a randomized control trial, decreasing selection concerns. We then estimate linear regression models with the preprocessed data to draw inferences about the effect of IGO membership loss on human rights behavior. Additionally, we estimate a general autoregressive distributed lag (ADL) model to understand the effects of lost membership over time.

**Data**

Our dependent variable is respect for human rights. To measure respect for human rights, we used the latent variable created by Fariss (2014). It incorporates many of the most commonly used measures of human rights into a dynamic model that yields a continuous measure of respect for human rights ranging from -3.117 to 4.370 in our dataset (higher numbers represent higher levels of respect).

Our main explanatory variable of interest is IGO membership loss. To measure loss of IGO membership, we create a dichotomous indicator variable taking the value ‘1’ if the state lost membership in at least one IGO in a given year, and ‘0’ otherwise. To create this variable, we use the International Organization v2.3 data from the Correlates of War Project (Pevehouse,
Nordstrom and Warnke 2004). The data include country-year membership for 529 IGOs. To be included in the data as an IGO, the organization must consist of three or more states that hold regular plenary sessions at least once every ten years and possess a permanent secretariat/headquarters (Pevehouse and Nordstrom 2003, 2).

States may fall into one of four different associations with IGOs: no membership, full membership, associate membership, or observer. We collapse full, associate, and observer status to create a dichotomous variable indicating any level of membership with the IGO to capture the socialization associated with membership. We sum the number of IGOs a state belonged to in a given year. The dichotomous indicator used as the explanatory variable takes the value ‘1’ if the total number of IGO memberships (at time, t) is less than the total IGO memberships from the prior year (at time, t-1), and ’0’ otherwise. Though a rough measurement, it captures our theoretical concept of interest of whether or not a state loses social connection(s) in the global community, thus allowing us to test the empirical implication of our argument. Future work should collect more fine-tuned data that allows testing interesting research extensions.8

As one would expect, IGO membership has steadily increased over time. However, Figure 1 shows some states have lost membership at least once, and some up to nine times. Although our measure does not distinguish between voluntary and involuntary loss of membership, we expect that any loss of social bonds will produce the exclusion effects that lead to human rights abuse. The IGOs included in the construction of the variable represent organizations concerned with a plethora of issues (for example, education, social development, industry, environment). We include IGOs from various issue areas (instead of only human rights IGOs, for example) be-

8For example, how does the manner in which states lose IGO membership differently affect human rights behavior?
cause our theory posits that states are influenced by the world polity and derive benefits from the social bonds formed across a whole range of international organizations that share liberal scripts of appropriate behavior (Greenhill 2010; Ingram, Robinson and Busch 2005). Theoretically, the discourse taking place within these organizations revolve around those ideas that hold sway in the global polity (Meyer et al. 1997). Thus participants in IGOs not officially driven by human rights — even more illiberal IGOs — consider their organizational environment and engage the dominant ideas (for example, human rights) when justifying their actions (Barnett and Finnemore 2004; Lipson 2007; Meyer and Jepperson 2000). Anecdotal evidence strengthens our choice to include even those IGOs that some may consider less concerned with human rights. For example, the Arab League suspended Syria in 2011 (AlJazeera 2011). Syria’s behavior did not improve. Rather, it ramped up repression, including the use of chemical weapons on civilians.

We include a number of possible confounders as controls. We control for the number of non-governmental human rights organizations (HROs) with permanent offices in a given state (Murdie and Bhasin 2011) and the total number of Amnesty International press releases and background reports (Ron, Ramos and Rodgers 2005). The presence and shaming activities of HROs may influence human rights behavior by mobilizing actors against the government (Keck and Sikkink 1998; Murdie and Davis 2012; Risse and Ropp 1999). They may also influence a state’s decision to join international organizations. Boli and Thomas (1997) argue that states exist in a society in which appropriate behavior is learned from other actors including non-governmental organizations, suggesting that those states most influenced by NGOs may imbed themselves more in IGOs. Descriptive data show the number of NGOs and IGOs concomitantly grew together in number since the end of World War II (Slomanson 2011, 125).
Democratic states abuse their citizens less often and less harshly than autocratic states (Davenport and Armstrong 2004; Poe and Tate 1994), so much so that Davenport (2007) coined the term “domestic democratic peace” to describe the pacific effect of democratic institutions on citizen-state relations. Democracy also increases interstate cooperation, especially participation in IGOs (Gaubatz 1996; Mansfield and Pevehouse 2006; Mansfield, Milner and Rosendorff 2002; Russett and Oneal 2001). However, Hill and Jones (2014) argue that the most common measures of democracy and repression are strongly correlated due to their tautological relationship. Instead, they recommend including a measure of executive constraints, as the measure represents institutional barriers to repression while not including repressive behavior (for example, competitiveness of participation which is affected by political imprisonments). Accordingly, we include XCONST from the Polity IV project as a measure of a state’s level of democracy (Marshall, Gurr and Jaggers 2012).9

We include a measure for the presence of a National Human Rights Institution (NHRI). Past work shows that NHRIs lead to better human rights respect (Cole and Ramirez 2013; Welch 2017), and those states most embedded in the global polity appear more likely to adopt NHRIs (Koo and Ramirez 2009).

We include the logged GDP of each state as a measure of economic capacity, using data available from the World Bank (2013). The monetary resources a state possesses may influence human rights outcomes (Poe and Tate 1994), as well as the state’s ability to join IGOs (Beckfield 2003; Jacobson, Reisinger and Mathers 1986; Shanks, Jacobson and Kaplan 1996). States with larger budgets can afford to hire more repressive state agents and equipment, for example; 9Using the complete Polity index yields similar results (presented in Appendix).
however, larger budgets also enable states to contribute to IGO activity and more effectively implement domestic policies or policy changes required for membership.

We include a logged measure of state population collected from the World Bank (2013). States with larger populations are more likely to repress citizens, which may be due to the increased probability of rebellion (real or perceived). Large populations strain resources, which is likely to create grievances that motivate rebellion (Henderson 1993). Additionally, large populations provide more potential rebel recruits, which in turn provide more opportunities for the onset of rebellion. The onset of rebellion may then lead to increased repression (Moore 1998). Lastly, population size may affect a state's ties with other states and thus its embeddedness (Beckfield 2003, FN 10), as larger states have the incentive and ability to engage in global governance structures more often than smaller states.

We control for number of IGOs in which a state is a member in a given year. Different states may be more or less inclined to embed in the global polity through IGOs, which could affect their behavior with respect to lost IGO membership or their human rights behavior (Greenhill 2010).  

Finally, we include a one-year lagged dependent variable. State repression behavior from the prior year predicts much of the variation of the present year's repression. For instance, Conrad and Moore (2010, 459) found “[s]tates whose agents engage in torture in a given year have a 93% chance a continuing to torture in the following year”. Including the lagged dependent variable accounts for possible autocorrelation in the data (Beck and Katz 1995).  

10 We include an interaction between lost IGO membership and number of total IGO memberships to test whether the affect of losing IGO membership on human rights is conditioned by the number of memberships. The coefficient does not reach standard significance levels (p = 0.65).
We also estimated a robustness check with other commonly used control variables for human rights models to further probe the plausibility of results. Given that our argument hinges on states’ existence in a global society — often referred to as the "world polity" — we consider the extent to which that community has spread throughout the world (Boli and Thomas 1999) by including a measure for the number of INGOs in a given year (Union of International Associations 2015).

We also include dummy variables for whether or not a state is engaged in a civil or international war. Violent conflict can result in instability that leaders combat with increased repression (Boswell and Dixon 1990; Gurr 1986; Poe, Tate and Keith 1999). The dichotomous variables come from Sarkees and Wayman (2010) and take the value ‘1’ if a war occurred in a country-year.

We did not include the INGO control in the main analyses given the theoretical and empirical reasons to believe this variable is highly correlated with IGO membership, thus possibly affecting the inferences for our main explanatory variable (correlation = 0.79) (Wonnacott and Wonnacott 1990, 502-504). We hesitated to add the conflict controls to the main analyses since international war does not consistently predict repression (Hill and Jones 2014; Poe and Tate 1994) and Hill and Jones (2014) caution against civil war as a covariate due to its operational definition's overlap with repression.

Adding these control variables actually makes a stronger case for the effect of lost membership — doubling the magnitude of the coefficient estimate and decreasing the p-value by an order of magnitude — but we present our theoretically-preferred and more conservative model as the main results. We present the robustness check results in the Appendix.
**Preprocessing**

To isolate the effect of losing membership in IGOs on human rights practices, we would need to administer the control (not losing membership) to \( state_i \) and measure the outcome (that state's respect for human rights) associated with receiving the control. At the same time, we would have to administer the treatment (losing membership) to \( state_i \) and measure the outcome associated with receiving the treatment. The difference in these outcomes tells us the difference in effect of the treatment compared to the control. Of course this contrived situation is impossible to achieve, leading to the “fundamental problem of causal inference” (Holland 1986).

If we were able to randomly assign to states in any particular year if they were going to lose membership in an IGO, we could assume the states in the control and treatment group are similar on average and isolate an average treatment effect (Rubin 1974). Of course, different reasons could influence why states lose membership, so the treatment does not occur at random. Matching methods allow us to more closely simulate that world with our observational data. Using nearest-neighbor (1:1) propensity score matching,\(^{11}\) we create a database in which half of the observations receive treatment and half receive control with balancing of the covariates (propensity score of treated and control after matching: 0.15, 0.15). By comparing like units with only the treatment/control condition differing between them, we are less concerned with selection issues. For instance, one may appropriately worry that certain states may be excluded more often (for example, non-democracies or rights offenders). By balancing the data, we dis-

\(^{11}\)We choose this matching algorithm for a number of reasons. Since all of the matching algorithms used produce similarly balanced datasets, we turn to theory. Theoretically, nearest-neighbor matching is intuitive and works best when the original data contain many more control than treated units (ours is around 8.5:1).
pense with these concerns. Table 1 includes the balance of the covariates\textsuperscript{12} before and after matching.

[Table 1 about here.]

\textbf{Linear Regression Results}

With our data appropriately balanced, we then estimate a linear regression model (Gerber and Green 2012; Guo and Fraser 2010).\textsuperscript{13} Table 2 shows the estimated coefficients and associated \(p\) values. Losing membership in at least one IGO leads to worse human rights practices (\(\beta = -0.03, p = 0.03\)). The estimated coefficient for the treatment (a state losing IGO membership) can be interpreted in a straight-forward way: a one unit increase in \(X\) results in a \(\beta\) increase in \(Y\).\textsuperscript{14} If a state loses at least one IGO membership compared to the prior year, their human rights score will decrease by 0.03 points.

\textit{Effects Over Time}

Recall that the dependent variable is a continuous latent measure of human rights respect that ranges from -3.13 to 4.69. The decrease in human rights protection due to lost membership, while statistically significant at standard levels, is substantively small. Even though substantively small, we believe any decrease in human rights protection a substantial finding. But,\textsuperscript{12} We use the same control variables noted above. The sum of IGO memberships enters as a control to take into account the possibility that some states may be more apt to join IGOs.\textsuperscript{13} We include a regression model with the full dataset in the Appendix. The results remain similar.\textsuperscript{14} We cannot interpret the control variables from the regression coefficients as one would do without preprocessing. To draw causal inferences about any of the control variables, we would need to re-specify the matching algorithm with the variable of interest as the treatment (Gerber and Green 2012).
we also wish to take advantage of the dynamics in the data to estimate the full effects of losing IGO membership. The interpretation of the coefficient represents the short run change in human rights respect. We expect lost membership to also exhibit effects over time. To estimate the long-term effects of lost membership we turn to time-series methods. More specifically, we estimate an unrestricted autoregressive distributed lag (ADL) model (Hendry 1995) of the form\(^\text{15}\)

\[
Y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{i=0}^{q} \beta_i X_{t-i} + \gamma C_t + \epsilon_t. \tag{1}
\]

where \(Y\) is human rights respect, \(X\) is whether the state lost IGO membership, \(C\) is a vector of controls, \(t\) indexes year, \(p\) refers to the number of lags of \(Y_t\), and \(q\) the number of lags of \(X_t\). The coefficients to be estimated are \(\alpha_0\) and the vectors \(\alpha_i, \beta_i,\) and \(\gamma\).

Social science theories, including our own, do not usually contain enough information to know the lag structure a priori. To determine our main model, we used the “general to specific” strategy to settle upon the appropriate lag structure for both \(Y\) and \(X\) (De Boef and Keele 2008; Hendry 1995). We start with an ADL(4,4)\(^\text{16}\) and iteratively pare the model down by noting which lags achieve statistical significance. We test for serial autocorrelation each time. The final model is both sufficient (no serial autocorrelation in the residuals) and parsimonious (does not include any unnecessary lags) (Webb 2017). In our case, that model is the ADL(3,0).\(^\text{17}\) We present results from the model in Table 3.

\(^{15}\text{We drop the index for each country for ease of exposition.}\)
\(^{16}\text{Autoregressive distributed lag notation lists the number of lags of } Y \text{ followed by the number of lags of } X \text{ in parentheses.}\)
\(^{17}\text{We fail to reject the null hypothesis of no serial correlation using the Durbin-Watson test } (p = 0.12).\)
The coefficient remains negative with a slightly larger $p$ value ($p = 0.08$) than the $p$ value from the matching/regression model. The strength of the matching/regression analysis lies in the identification strategy to make causal claims about IGO membership loss on state human rights behavior. Estimating the ADL allows us to leverage the dynamic nature of the data to calculate that relationship over time.\textsuperscript{18} To do so, we need to calculate the long-run multiplier (LRM) (Wooldridge 2009). Using the notation for the ADL above, we calculate the long-run multiplier as \[ \frac{\beta_0}{1-(\alpha_1+\alpha_2+\alpha_3)}. \] Using this equation we estimate a long-run effect of losing IGO membership on human rights scores as -0.55. Although the matching/regression strategy estimate of human rights respect decreases only slightly the year of IGO membership loss (-0.02), as time goes on, human rights respect decreases by 0.38 standard deviations of the dependent variable.

**Conclusions**

States that lose membership in the institutions of international society repress more often than those that do not lose membership. Our findings confirm the empirical implication of our argument that when states lose membership in IGOs, the social incentives to behave appropriately decrease as social bonds are lost.

\textsuperscript{18}This is true only if the data do not contain a unit root. We perform a number of tests to assure no unit root exists. Using an augmented Dickey-Fuller test, we reject the null hypothesis of a unit root ($p < 0.01$). An Im-Pesaran-Shin test can only tell us whether some panels are stationary compared to the null of all panels containing unit roots. At least some of our panels are stationary ($p = 0.0004$). Fisher-type tests based on augmented Dickey-Fuller tests ($p = 0.01$) and Phillip-Perron tests ($p = 0.00$) suggest at least one panel is stationary as opposed to all panels containing a unit root. Since many of these tests lack specificity about every panel, we also estimate a first-differenced regression (Wooldridge 2001). The reader can find the results in the Appendix, but the significance level of the coefficient on $\Delta X$ suggests the relationship we see in the ADL(3,0) is not a spurious relationship caused by a unit root ($p = 0.09$).
Skeptics may wonder about the appropriateness of leveraging individual-level social psychological theory and findings when studying international relations. We believe it so for the following reasons, though ultimately the reader must decide if these reasons suffice.\textsuperscript{19} Most international relations scholars have personified the state since the eighteenth century (Wendt 2004). Some even argue, convincingly, that individual attributes such as integrity more logically apply to states than individuals (Nili 2018). We do not go that far, but rather see simplifying the state as an agent as a useful theoretical move. Like Waltz (1979) we do not claim that individuals and domestic politics do not matter, but rather that viewing the state as the primary actor can yield important insight, and just as important in our view, testable implications. Following Thies (2010), we do not deny states act as individually rational agents in pursuit of material benefits, but also recognize and take seriously their social nature and preferences. We also agree with Wendt (2004) that states can be viewed as psychological agents. All of this leads us to take seriously the work of those who spend their careers considering social psychological interactions and outcomes, and to apply those insights to states as social psychological actors. To be sure, exploring how domestic politics might condition the arguments made here may lead to interesting insights. We leave this for future work.

As for other future work, the idea that social influence affects states’ human rights practices suggests the possibility that social psychological mechanisms could also affect other state behaviors. If IGO membership loss disincentivizes appropriate behavior, then what other behaviors can be affected? According to our argument, states may break other international norms as a result of lost social bonds. Future research should explore the effects of lost social bonds.

\textsuperscript{19}A rich debate exists in international relations about how to view state agency and why. The citations are too many to list here. Interested readers may wish to consult the special 2004 forum on the topic in Review of International Studies as a starting point.
on norms of state sovereignty violation (that is, aggressive territorial conflict) and treaty compliance, for example.

The current project represents our attempt to understand an interesting feature of IGO membership we felt has been unexplored. While much scholarship has asked why states join IGOs (Mansfield and Pevehouse 2006; Pevehouse 2002a, b, 2005; Poast and Urpelainen 2013) and the consequences of doing so (Abbott and Snidal 1998; Bearce and Bondanella 2007; Greenhill 2010; Poast and Urpelainen 2015; Russett and Oneal 2001), and newer research is tackling why states exclude other states from IGOs (von Borzyskowski and Vabulas 2018), we explore the consequences of exclusion. Like much research on IGOs, we started by looking for average effects over many IGOs. Future research should explore how IGO variation — for example, levels of institutionalization (Boehmer, Gartzke and Nordstrom 2004), type of IGO (Mansfield and Pevehouse 2008), IGO design (Koremenos, Lipson and Snidal 2001), how democratic the IGO membership is (Pevehouse and Russett N.d.), IGO age — might influence the effects we find here. Future research should also dig deeper into the mechanisms of action. We propose a plausible link between lost IGO membership and human rights behavior – decreased information for both the state that lost ties and the states still members in the IGO, which affect the state that lost membership’s social psychological decision calculus. When the state loses membership it loses monitors and access to normative scripts. Without these social influences, states resort to repression more often. We propose the decreased information mechanism as it is shared across the psychology and international relations literature. Future research could probe this further.

Our findings also add to the debate on inclusiveness in international relations. As international society has become more interdependent, states continue to create IGOs to govern the

\[\text{\footnotesize 20}\text{We thank an anonymous reviewer for this suggestion.}\]
relationships between states. Although IGOs restrict membership to solve enforcement, uncertainty, and distribution problems (Koremenos, Lipson and Snidal 2001), many see inclusiveness as a way in which to socialize other states through acculturation and learning (Goodman and Jinks 2004). Our research highlights the importance of inclusiveness and cautions against using exclusion as a punishment without considering the repercussions for those living within the punished state. For example, although excluding states like North Korea from the United Nations could be construed as punishing them (Rosett 2010), the likely consequences would be to increase, rather than diminish, human rights violations in those states. Inclusion in IGOs not only encourages appropriate behavior through material incentives, but membership also opens states up to social influences that incentivize appropriate behavior. Our results suggest that the social influence mechanism may be more important in driving conformity to human rights norms than expecting that states will internalize these norms and uphold them for their own sake, since it is the continued membership in the community that influences state behavior. We do not doubt that some states come to internalize human rights norms, but fewer states should make it to the internalization phase from the norm cascade in which legitimacy and reputation drive behavior (Finnemore and Sikkink 1998).

Our paper highlights the costs of restricting state membership of IGOs. However, this represents just one piece of states' overall utility calculations. States, IGOs, and international society may also experience benefits from the exclusion of recalcitrant states, such as increased functioning capacity of IGOs, preservation of legitimacy, or countering the spread of alternative norms. For example, “rogue” states could sabotage effectiveness or create counter-cultures with like-minded states, thereby sowing dissent within organizations that could affect state practice
throughout the world polity. Further research should explore these costs and benefits of IGO inclusion/exclusion and how they affect state outcomes that shape global society.

Lastly, this paper demonstrates the potential gains from interdisciplinary scholarship across the social sciences (for other examples in political science, see Druckman, Kuklinski and Sigelman 2009; Hafner-Burton, Victor and Lupu 2012; Thyne and Schroeder 2012). The present research question and the resultant discoveries would not have been possible without communication across disciplines. Social psychology may be an especially fruitful area for collaboration, given its diverse array of perspectives on individual behavior that could be applied to explain the behavior of state-level actors.

We hope this paper pushes scholars closer to an understanding of why states (do not) abuse human rights. Human rights violations are an unfortunate consequence of states’ incentives to maintain power. The international community should use any and all ways it can to disincentivize such malevolent behavior. Although past work has emphasized the material punishments and rewards that organizations can use, the present work suggests that social influences on state behavior are a potential tool for furthering human rights protections. Although organizations are often tempted to kick out wayward members, doing so comes at a heretofore unseen cost. When states lose membership in international organizations, they lose affiliative reasons to avoid repression and protect their reputation, they lose other states’ surveillance of their actions, and they become more likely to violate human rights.


