1 INTRODUCTION

Private governance initiatives such as product certification programs have targeted farm and factory working conditions, greenhouse gas emissions, and fishery, mine, and forest management (Auld, 2014; Bartley, 2003; Bozzi et al., 2012; Hudson and Hudson, 2003; van der Ven, 2015; Vince and Haward, 2017). Many of these initiatives were founded by activists who were dissatisfied with public regulations. With tactics like boycotts as "sticks" and brand-boosting praise as "carrots," activists create incentives for companies to comply with their preferred certification programs, thereby giving these programs rulemaking authority rooted in market power rather than governmental powers (Cashore, 2002). In some sectors, such as forestry, activist-backed programs meet resistance from industry groups that launch competing certification programs to offer more "business-friendly" alternatives. This results in hotly contested debates between supporters of programs founded by activists and supporters of industry-backed alternatives. These debates center on the relative stringency of each program's requirements.

In response to these emerging forms of regulation and ensuing public debates, scholars across political science, economics, and sociology have given sustained theoretical and empirical attention to the evolution of private regulation (Grabs et al., 2018). A key motivation behind this research is understanding whether the social and market forces behind private regulations lead to similar patterns observed in public regulations, such as a "race to the bottom" as governments attempt to attract capital, a "race to the middle" as shared expectations emerge, or a "race to the top" as companies operating in areas with more stringent regulations lobby to equalize requirements across jurisdictions (Berger and Dore, 1996; Rodrik et al., 2004; Vogel, 1995).

While this scholarship has made great strides, it lacks consistent concepts of regulatory stringency and methods to compare policies over time. Some scholars rely on broad characterizations of regulatory stringency without offering a precise definition. Others base

their characterizations on only a small subset of the issues that programs address. Such ambiguity, we argue, has led to seemingly contradictory findings, hindering efforts to compare private regulations and develop theories about why they change. To be sure, these questions interest students of both public policy and private governance, but whereas a rich public policy scholarship has emerged to measure and explain policy change as a dependent variable (Green-Pedersen, 2007; Hall, 1993; Howlett and Cashore, 2014), students of private governance have given much less attention to the concept of policy change.

To address this gap, we build on public policy concepts to offer a two-part framework to describe and compare regulations over time. Part one of this framework distinguishes three types of regulatory stringency: 1) How comprehensive is the scope of issues addressed? 2) How prescriptive are the requirements? 3) What are the specific levels of performance required? Part two offers a method to classify changes across programs, yielding nine possible patterns to describe both relative and absolute directions of policy change. This approach provides a common language to describe how various regulations in the same policy space may change over time. Such research is especially important where multiple programs, backed by different coalitions, compete to exercise regulatory authority. By distinguishing among types of regulatory stringency and patterns of change, many of the hypotheses advanced by the private governance literature can be restated in ways that are more conceptually precise and thus more tractable for empirical testing.

We proceed in the following steps. Section two maps the different concepts and measures of regulatory stringency in existing private governance scholarship. Section three presents our framework to distinguish three types of stringency and compare them across programs and over time. In section four, we use this framework to compare competing forest certification programs, one of the most institutionalized forms of private regulation. Overall, the activist-backed program is more prescriptive on more issues than the industry-backed program and has become more so over time. Comparing the regulations from each other yields a pattern of "ratcheting up" but also of "diverging" prescriptiveness as the activist-backed group imposes

2

increasingly stringent rules and the industry-backed group imposes rules that are increasingly less stringent. In particular, the activist-backed program increased requirements on ecological protection, while the industry-backed program increased requirements on issues relating to efficient business operations and public perceptions of forestry. However, in 2015 the industry-backed program introduced new requirements on three social and ecological issues that were historically emphasized by the activist-backed program. These results raise questions for future research, which we outline in section five.

2   REGULATORY STRINGENCY

Public policy scholarship has long been concerned with how different kinds of regulatory stringency might have different causes or effects (see Brunel and Levinson, 2016, for a review). Private governance scholars' limited attention to these questions is problematic because measuring regulatory stringency is necessary to assess theoretical debates about the role of policy content as both a dependent and explanatory variable. A measure of regulatory stringency as a dependent variable is necessary to test propositions about the way in which activist campaigns, market forces, and competition among alternative programs shape policy content. Conversely, a measure of regulatory stringency is also necessary to test propositions about how policy content shapes activist support, market adoption, impact, and how other programs respond.

Stringency as an explanatory variable: Scholars who study how private regulations gain legitimacy, trust, or support from various audiences posit that regulatory stringency is a key explanatory variable. For example, McDermott (2012) argues that stringency may reduce trust by mandating formulaic, top-down approaches. Atkinson and Rosenthal (2014) find that perceived stringency increases market demand for certified products, but Prado (2013) finds that it also reduces adoption by firms. Meidinger (2003) suggests that changes in stringency that disadvantage some firms or groups may catalyze these actors to create alternative private regulatory programs. Alternatively, those disadvantaged by changes to private regulation may

then opt to pursue their aims through public policy (Weimer, 2006). Such consequences are consistent with broader findings from literatures on "corporate social responsibility" (CSR) initiatives, such as environmental management systems (EMS), industry codes of conduct, and third-party certification programs, which find that more costly requirements are less likely to be adopted (Delmas and Montiel, 2008; Kollman and Prakash, 2001; Lyon and Maxwell, 2008).

The effects of stringency on trust, legitimacy, compliance cost, and adoption matter because understanding the likely future impact of private regulations "on the ground" requires understanding their evolutionary trajectories (van der Ven et al., 2018). Even activist-backed programs that establish stringent requirements on one issue at one point in time may not do so on other issues and at other times (LeBaron and Burgoon, 2018). Nuanced gaps in otherwise stringent private regulations—"regulatory loopholes"—may also explain their lack of success in addressing problems like deforestation (van der Ven et al., 2018). Taken together, these studies suggest that changes in regulatory stringency may have a wide range of effects, but assessing them is often hampered by inadequate attention to defining and measuring stringency as an explanatory variable.

Stringency as a dependent variable: Regulatory stringency is also the primary variable of interest in studies of the reverse causal relationships: how ideological, economic, political, and social forces shape and constrain the policy content of private regulations (Bartley, 2003; Cashore et al., 2004; Fischer and Lyon, 2014). Here, regulatory stringency is the dependent variable. Unlike governments, which enjoy sovereign authority, private organizations must achieve and maintain legitimacy in the eyes of both those they aim to empower and those they seek to regulate (Bartley, 2007; Bodansky, 1999; Cashore, 2002), and one way they do this through claims about the stringency of their requirements.

Scholars theorize that various forces either promote or hinder stringent regulation. For example, ideas about the political responsibilities of businesses shape both activist demands for private governance and firms' responses to private governance efforts (Bartley, 2003; Djelic

and Etchanchu, 2017). These different ideas are then embodied in more or less stringent policies depending on which coalitions gain rulemaking authority (Botzem and Dobusch, 2012; Hsueh and Prakash, 2012). Bartley (2003) finds private regulations emerging when social movements target companies with tactics that aim to redirect, rather than challenge, neo-liberal ideas. Others find private regulations arising from collective action by industry to preempt or replace more stringent government regulations (Bartley, 2007; Cashore, 2002; Grabosky, 2013; Green, 2013; Loconto and Fouilleux, 2014; Lyon and Maxwell, 2008; Maxwell et al., 2000; Prakash, 2000). Abbott and Snidal (2009) suggest that the content of public and private regulations are a joint result of bargaining between activists and firms. The common thread is that each of these studies aims to explain relative differences or changes in policy.

Others seek to explain variation in regulatory stringency as a result of endogenous interactions among private authorities (DeLeon and Rivera, 2009; Eberlein et al., 2014; Green, 2017; Gulbrandsen, 2014; Howard-Grenville et al., 2008; Li and van 't Veld, 2015; Mills, 2016). For example, Smith and Fischlein (2010, p. 520) suggest that competing private regulations change frequently and often imitate each other. Similarly, Eberlein et al. (2014) identify "frequent rule revision" or "differentiation among rule systems" as potential effects of such interaction.

A related body of scholarship seeks to explain regulatory stringency as a result of strategic interactions among the coalitions backing different programs. Some focus on how competition may lead to more "weak or lax standards" as firms "shop" for lower-cost programs, potentially causing a "race to the bottom" (Abbott and Snidal, 2010; Fransen, 2011; Gulbrandsen, 2004). In contrast, others find competition causing "weak" regulations to be "revised upwards" as activists invite public comparisons with the requirements of "higher" regulations (Overdevest, 2005, 2010). And still others find both patterns occurring, depending on market and industry structures (Cashore et al., 2004; Hassel, 2008; van der Ven, 2015). Cashore et al. (2004) highlight how market and institutional logics initially work to pressure coalitions to "lower" stringency but then, later, work to maintain differences.

Concepts of regulatory stringency are also at the core of formal models of private governance. Models by Abderrazak and Youssef (2009) and Fischer and Lyon (2014) suggest that standards may increase or decrease stringency under different conditions, such as increases or decreases in compliance costs or market demand. Game-theoretic models (Fischer and Lyon, 2014; Li and van 't Veld, 2015; Poret, 2016) and empirical research (Cashore et al., 2004) both suggest that asymmetric incentives lead competing programs to adopt different levels of stringency in equilibrium. Where an activist-backed regulation competes with an industry-backed regulation, these theories predict a stable equilibrium where the activist-backed regulation is more stringent.

Taken together these studies posit a wide range of theories to explain change in regulatory stringency, but assessing them has been hampered by poor attention to the dependent variable they seek to explain. The contradictions are most striking in the patterns of change these studies claim to find. Some posit—and find evidence for—a pattern where competing regulations "ratchet up" and less stringent regulations converge toward more stringent ones (Overdevest, 2005, 2010; Overdevest and Zeitlin, 2014). Other scholars posit—and find evidence for—the exact opposite pattern, in which competitive pressures lead a "race to the bottom" with more stringent programs decreasing stringency and converging toward less stringent ones in (Abbott and Snidal, 2010; Fransen, 2011; Gulbrandsen, 2004). Still others posit—and find evidence for—yet another pattern where programs maintain different levels of stringency, i.e., they remain distinct, neither converging to the "top" nor the "bottom" (Fischer and Lyon, 2014; Li and van 't Veld, 2015; Poret, 2016; Cashore et al., 2004). While these three sets of findings seem incompatible, we argue that they are the result of different measurement strategies. Reconciling them thus requires a set of shared concepts and measures of regulatory stringency.

2.1 Concepts & Measurement of Variation in Private Regulations

The diversity of private governance scholars' conceptual and empirical approaches to measuring regulatory stringency makes this literature vibrant but confusing: Some scholars evoke vertical notions of variation, describing standards as high or low or more or less stringent (Fischer and Lyon, 2014; Li and van 't Veld, 2015). Others evoke horizontal notions of variation, describing the width or breadth of issues covered (Auld, 2014; Heyes and Martin, 2017). Cashore and Howlett (2007) calls attention to variation in prescriptiveness versus flexibility, i.e., the extent to which regulations use mandatory and substantive performance thresholds. Other scholars combine concepts of breadth and prescriptiveness into one broader notion of stringency (Fransen, 2011). Some measure height in a relative sense, defining the "benchmark" as the higher standard (Overdevest, 2005, 2010). These distinct dimensions of stringency are often conflated. For example, formal models often assign each program a single overall "quality" or "stringency" parameter that could be measured multiple ways yielding different results. And these are only a few of the many measures of stringency used in this literature, ranging from so broad that they conflate multiple concepts to so narrow that they measure only a few select components (see Table 1).

Overall, concepts of stringency in existing work tend to be either insufficiently precise to be consistently applied across programs, insufficiently comprehensive to yield consistent results, or completely absent.

In the absence of consistent measures of regulatory stringency, scholars have turned to proxy measures that can be more easily compared across programs. For example, Darnall et al. (2010) suggest that a program's sponsor is a signal of its stringency. In the broadest study to date, van der Ven (2015) uses another common proxy for stringency–compliance with perceived "best practices," which are often also considered "benchmarks" for measuring stringency but are based on a variety of different notions of "rigor" and "credibility." However, these approaches do not allow scholars to examine relationships between stringency and program sponsorship or between stringency and perceived stringency.

Table 1: Concepts and measures of regulatory stringency

Selected Scholarship	Concept	Measurement orientation	
García-Montiel et al. (2017)	"more rigorous," "higher level," "higher quality," and thus "Greater complexity/ Effectiveness/ Cost" vs. "More Simplicity/Lower Cost"	Number of indicators. Descriptions of consistency, coherency, and completeness.	**More Specific**
Moore et al. (2012)	Management practices changed	Survey of self-reported number and type of management practices implemented	
McDermott et al. (2010; 2008)	"comprehensiveness and prescriptiveness"	Number of key issues with most prescriptive language	
Overdevest and Zeitlin (2012)	"far apart" or "closer" on select "Program Characteristics"	Binary table of select issues and descriptive examples	
Overdevest (2005, 2010)	"comparative quality"— "weaker" standards are "revised upwards" to be "equivalent" to "higher and more prescriptive standards"	Descriptive theory, examples, and review of previous comparisons	
Fransen (2011), Fransen & Conzellman (2015)	"stringency" as "comprehensive in scope, specific in content, and prescriptive in terms of requirements"	Descriptions based on "leading policy analysts per issue area"	
Hansen et al. (2006)	Select "general features" and "six aspects" of management	Descriptive table of select issues	
Auld (2014)	"Policy scope and regulatory domain" "policy changes," "character of the rules developed"	Description of the set of problems addressed and how	
Cashore et al. (2004)	"stringency"	Descriptive theory and examples	
Smith & Fischlein (2010)	"stringency" of "weightings across multiple, and often conflicting, attributes," also "excellence in content"	Descriptive theory and examples	
Porter (2014)	"hard law" or "soft law"	Descriptive examples	
Gulbrandsen (2004)	"variations in the strength of standards"—"more stringent and less discretionary," "more rigorous and wide-ranging" vs. "weak or lax" "and allow far wider flexibility" Some "regulations have become more flexible" while others are "changing upward"	Descriptive examples	
Eberlein et al. (2014)	"differentiation among rule systems" along "dimensions of regulatory governance," e.g. "more or less stringent requirements" or "regulatory capacity"	Descriptive typology and examples	
Hassel (2008)	"high and low quality regulation", "higher standards" vs. "lower standards"	Descriptive theory and examples	
Bartley (2003)	"more credible claims" vs. "lax standards"	Descriptive theory and examples	
Abbott & Snidal (2008)	"substance and form of regulatory outcomes"— "stringent" "higher standards" vs. "less stringent" "business-friendly" "weaker standards"	Descriptive theory	
Bernstein & Cashore (2007)	Pressure to "raise" or "lower" requirements "explains convergence/ divergence"	Descriptive theory	
Kollman & Prakash (2011), Potoski & Prakash (2004), Prakash & Potoski 2006)	"lax" or "processes-based" vs. "more stringent" "outcome-based" or "product-based" "types of regulations"	ISO14001 classified as process-based, stringency assessed only for public regulations	
Prakash & Potoski (2007)	"stringent" vs. "lenient" standards	Proportional costs, social externalities, and branding benefits	
Formal models of "stringency" or "quality"	"sustainability quality level" (Poret 2016), "more ambitious" (Fischer et al., 2017), "stricter rules" (Schmitz et al., 2017), "stringency" (Fischer and Lyon 2014; Hayes and Martin, 2017)	Proportional costs & benefits to programs & firms	
Formal models of issue scope	"issue-width" in an "issue space" (Hayes and Martin 2015)	Proportional costs and benefits to programs and funders	**Broader**

More importantly, different approaches to measuring regulatory stringency prevent us from adjudicating between claims that competing regulatory schemes will "race to the bottom," "ratchet up," "converge," or "diverge." Indeed, different measurement strategies explain the seemly contradictory evidence in favor or each theory. While van der Ven (2015) does find support for the prediction that activist-backed private regulations are more likely to align with "best practices" but does not find support for the prediction that industry-backed regulations are less likely to do so. The latter finding seems to contradict findings by Cashore et al. (2004) that industry-backed programs set less stringent requirements. This is due to differences in measurement; Cashore et al. focus on substantive prescriptiveness rather than procedural "best practices."

Even the handful of scholars who have developed direct and precise measures of stringency (the top of Table 1) tend to fall short because they only focus on a few salient components, rather than attempting to specify a full range of relevant comparisons. This approach can lead to conflicting results if scholars select different policy components as indicators of stringency. For example, Cashore et al. (2004) examine prescriptiveness on seven issues related to ecological protection (plantations, chemicals, clearcuts, exotics, reserves, streamside riparian zones, and genetically modified organisms) and find large enduring differences between activist-backed and industry-backed regulations. In contrast, Overdevest and Zeitlin (2014) find that these same private regulations "all moved closer" by assessing whether or not each program addressed six other features—two substantive requirements on firm behavior (public reporting and stakeholder consultation), two on compliance mechanisms (auditing and supply chain tracking), and another two on decision-making and marketing strategy—finding policy convergence on all six. Hence, choices about which issue to measure may lead to different conclusions and appear to support conflicting theories of change.

Efforts to identify patterns of change face two common challenges. First, results vary depending on the policy components we study. McDermott et al. (2008) identify 48 key types of substantive requirements in the forestry sector alone—selecting any limited combination of

them may lead to different conclusions. If Overdevest and Zeitlin had chosen Cashore et al.'s set of issues or vice versa, each might have found the opposite pattern.

Second, binary indicators such as whether or not a program addresses a given topic—i.e. "is this issue in the program's scope?"—fail to capture variation in degree—e.g., "how high is the threshold set" (what is the required frequency of public reporting or prohibited amount of pollution?) and "how prescriptive are they?" (How much is voluntary versus mandatory?). The scope of requirements, degree of prescriptiveness, and levels of thresholds are important but orthogonal dimensions of variation that may exhibit different patterns of change for different reasons. Overdevest and Zeitlin (2014) assert that the industry-backed program moved in the direction of the activist-backed program within the scope of issues related to public reporting and consultation, while Cashore et al. (2004) found that these competing programs did not converge in prescriptiveness on issues related to ecological protection. We can thus resolve the apparent conflict between Overdevest and Zeitlin's study and Cashore et al.'s study by distinguishing between the scope of issues covered and the prescriptiveness of requirements as two distinct dimensions of regulatory stringency.

To address these challenges, we distinguish three types of regulatory stringency and illustrate how doing so enables scholars to compare programs across issues and over time more systematically.

3   A Framework to classify change in private regulations

Students of public policy have long recognized the need to break policies into their component parts, finding different explanations for change for different policy goals and means to achieve them (Cashore, 1997; Hall, 1993; Weimer and Vining, 2005). While private governance scholars have shown that private regulations resemble public laws (Meidinger, 2003, Meidinger (2006)), they have paid less attention to distinctions between policy ends and means. Similarly, policy change, a core concept in public policy scholarship, remains underdeveloped in research on private regulation. We thus draw on public policy scholarship to address these gaps.

3.1 Step 1: Measuring scope, prescriptiveness, and policy settings

We focus on three dimensions of variation: (1) the comprehensiveness of a regulation's scope (i.e. which policy problems it addresses), (2) the extent to which requirements are prescriptive versus flexible (i.e. whether they have mandatory and substantive thresholds), and (3) the levels of those thresholds or similarly specific policy settings. Our framework thus combines qualitative issue-by-issue comparison of policy settings with two measurement concepts, policy scope and prescriptiveness, that can be applied across issue areas and thus aggregated to measure overall stringency. That is, by comparing the number of issues covered by a regulation and the number of prescriptive requirements on those issues, scholars can assess aggregate trends.

The first step for scholars who wish to make claims about stringency or direction of change involves three tasks: describing policy content according to policy settings, scope, and prescriptiveness (Table 2). Comparing across programs requires a second step: measuring relative stringency and change on each dimension (see Table 3). First, we elaborate on step one.

Table 2: Measures of Regulatory Stringency

|	Program Level	Issue Level
Policy Ends	How **comprehensive** is the **scope** of issues addressed?	What are the **specific requirements (i.e. policy settings)** on each issues? (e.g. the specific size of stream buffer zones)
Policy Means	In aggregate, across all issues, how **prescriptive** is each regulation? To what extent (e.g. on what portion of issues) are there mandatory and substantive thresholds?	1. How **prescriptive** is each requirement? 2. How are they enforeced?*

*Beyond the scope of this paper

Scope: To assess a regulation's scope, we must inductively derive a set of distinct policy

"issues" that are addressed by one or more regulatory texts in a given policy domain. We can then measure the comprehensiveness of each regulation's scope by asking which of these issues it addresses. While this approach reduces the risk of omitting key issues on which regulations may vary, it is costly. Scholars may thus opt for a limited scope, as long as they clearly describe their scope relative to the potential set of comparisons. A comprehensive approach is necessary, however, to assess claims about the scope of regulations (such as the hypotheses from section 2.3). We can measure a regulation's scope in an absolute sense (how many issues it addresses), in a relative sense (how many more or fewer issues it addresses than its competitor), and in change over time (on how many issues were requirements added or subtracted).

Prescriptiveness: Second, we measure the extent to which each requirement is prescriptive, i.e., has substantive and mandatory features like performance thresholds (see Table 3 adapted from Cashore and Howlett (2007)). Because "prescriptive versus flexible" refers to how each issue is addressed, not the ends of the policy, we can compare prescriptiveness across different substantive requirements. Prescriptiveness is a continuum from discretionary guidelines, which allow maximum flexibility, to procedural requirements that define processes that must be followed but do not prescribe outcomes, to mandatory substantive requirements, which prescribe precise actions, such as quantitative performance thresholds. In contrast to mandatory thresholds, even mandatory requirements to follow local "best management practices" are less prescriptive because these practices may not include substantive requirements. Discretionary practices, processes, or plans are even less prescriptive. On each issue, we code requirements in an absolute sense—as "no prescriptive requirements" or "some prescriptive requirements"—and then, if the latter, in a relative sense—whether they are "more prescriptive" than another regulation or "most prescriptive" (requiring as much as or more than any other regulation). Coding prescriptiveness across issues creates an additional measure of policy scope: how many key issues have "some prescriptive standards." Coding prescriptiveness across programs creates a measure of the relative level of prescriptive requirements. Additionally, we classify

changes as becoming more prescriptive or less prescriptive on each issue, thus capturing the direction of change in prescriptiveness.

Table 3: Potential patterns of change with two regulations

		Relationship Among Standards		
		Converging	Parallel	Diverging
	Increasing	↗ ↗	↗ ↗	↗ ↗
Directions of Change (in comprehensiveness of scope, prescriptiveness, *or* levels of requirements)	Opposite or Equilibrium	↘ ↗	⟶ ⟶	↗ ↘
	Decreasing	↘ ↗	↘ ↘	↗ ↘

Policy settings: Finally, the third type of stringency—specific performance levels (what policy scholars call "policy settings")—allow us to interpret differences in scope or prescriptiveness substantively. For example, forestry certification programs have different requirements for how close loggers can harvest near streams. In this example, all standards prescribing minimum stream buffer widths are equally prescriptive since all are mandatory requirements, albeit with different thresholds. Yet buffer widths and other specific policy settings are a meaningful type of variation. Unfortunately, most specific policy settings, even prescriptive ones, cannot be quantified and are thus difficult to aggregate. Even numeric stream buffers are difficult to compare because they often vary in different landscape contexts, for example in mountainous or flat areas, and involve different levels of harvest restrictions based on different criteria, such as whether fish live in the stream (see Figure 5 in section 4). Measurement strategies that allow program-level aggregation cannot replace issue-specific qualitative comparison. It is crucial to both quantify absolute and relative differences and describe the most meaningful differences that capture the overall trends. We thus suggest that scholars combine aggregate measures with descriptive comparisons of important requirements, assessing each issue in an absolute sense, in a relative sense (if possible), and in how the required level of performance changed.

At its most stylized, step one, comparing two hypothetical programs (A and B) in a policy space with two issues (Hazardous Chemicals and Worker Training) might look like this: A researcher examines regulations in this policy domain and inductively identifies a total of two issues. Both programs have some prescriptive requirements on both issues, so they are equal in scope. Program A bans using chemicals above certain quantitative toxicity thresholds, whereas Program B bans "hazardous" levels which one may interpret several ways, so Program A is more prescriptive on the issue of Chemicals. For policy settings, the two programs ban slightly different lists of chemicals, Program A focusing on ecologically harmful chemicals and program B targeting those most harmful to humans, so one can only compare their specific requirements on chemicals qualitatively. On the second issue, both programs require mandatory worker training programs, and neither specifies how many hours, so they are equally prescriptive on Training. Each program suggests a slightly different list of topics for training to cover. Program A focuses more on skills needed to avoid ecological harm, and Program B focuses more on worker safety, so again, one can only compare their policy settings qualitatively. Yet a pattern emerges: Program A, the overall more prescriptive program, is also more focused on ecological protection, possibly due to a strong influence from environmental activists. In contrast, program B is more focused on worker safety, possibly to reduce the risk that worker injuries at one firm will impose reputational or regulatory costs for the whole industry.

As this example illustrates, the combination of precise and comprehensive measurement can avoid problems with using any one approach alone. Measuring scope alone risks overlooking variation in prescriptiveness and levels of performance required. Measuring prescriptiveness alone risks capturing a kind of stringency that is void of content. And comparing a few specific performance levels alone risks missing the broader picture, or worse, making overly broad generalizations where a different set of issues would yield different overall conclusions.

3.2 STEP 2: CLASSIFYING PATTERNS OF CHANGE

Drawing on Baumgartner (2002) and Howlett and Cashore (2007), we also note the importance of the direction of change. Assessing patterns of change like punctuation or equilibrium requires measuring change on each dimension because there may be equilibrium on one dimension but punctuation on another. In absolute terms, requirements may be increasing, decreasing, or neither, and, in relative terms, competing regulations may be converging, in equilibrium, or diverging on each dimension over any given period (Table 3). Thus, in aggregate, nine relationships fully capture the possible dynamics for each dimension of change. All of the concepts of regulatory stringency from Table 1 should be able to be expressed in terms of the dimension(s) to which a theory applies, the absolute directions of change they predict, and relative relationships they anticipate.

Conclusion: This framework helps researchers accomplish four key tasks that Brunel and Levinson (2016) set out for a measure of regulatory stringency (1) measure change over time, (2) assess both relative and absolute magnitudes, (3) aggregate across multiple issue areas, and (4) it can be theoretically related to compliance costs. To these criteria, we add that any measurement approach should also capture qualitative differences in policy settings.

3.3 THEORIZING IN TERMS OF SCOPE, PRESCRIPTIVENESS, AND POLICY SETTINGS

Our core methodological critique is that different dimensions of stringency may exhibit different patterns of change. Precise and testable hypotheses about the causes and effects of change must distinguish among types of policy change. If different dimensions of regulatory stringency vary independently, a vast array of theories that use stringency as an explanatory or dependent variable must be revised to specify the dimension(s) to which they apply. Revisiting theories in terms of scope, prescriptiveness, and policy settings may yield different predictions on each dimension. It is beyond the scope of this paper to revisit all hypotheses in this vast literature in light of our methodological critique, but, for illustrative purposes, we offer

examples of such a restatement for hypotheses rooted in compliance cost and differentiation.

Compliance costs and competition: By breaking down stringency into three distinct dimensions, we expand on two related propositions: (1) that compliance costs cause competing programs to set different levels of stringency and (2) that programs change in response to changes by their competitor. Cashore et al. (2004) and Fischer and Lyon (2014) theorize that industry-backed programs set less stringent regulatory requirements than activist-backed programs because industry-backed programs are less willing to impose costs on firms. Likewise, we can expand upon the proposition that, when private authorities compete for market share, if one changes its requirements, the other will change in a similar direction (Fischer and Lyon, 2014; Smith and Fischlein, 2010). Yet, these studies do not specify which dimensions of stringency ought to be affected by compliance costs. Do incentives rooted in compliance cost affect each dimension in the same way? Are competing programs more responsive to changes in the scope, prescriptiveness, or policy settings of competing standards? Disentangling policy settings, scope, and prescriptiveness suggests more precise hypotheses to assess such theories rooted in compliance cost and competition.

Revised compliance cost hypotheses: If broadening scope is low-cost for firms but increasing prescriptiveness and performance levels are high-cost,

- **H1.1:** An industry-backed regulation will be more similar to an activist-backed regulation in scope than in prescriptiveness or performance levels.
- **H1.2:** An industry-backed regulation will be more likely to respond to changes in an activist backed regulation by converging in scope than in prescriptiveness or performance levels.

Differentiation: Another core theoretical claim is that different coalitions will establish qualitatively different policies (Botzem and Dobusch, 2012; Hsueh and Prakash, 2012). By distinguishing types of stringency, we may identify qualitative differences in how stringency varies across programs.

We suggest that the relative stringency of an industry-backed program on a given issue depends on whether it results in net costs or net benefits to the industry. Activist-backed programs drive stringency on issues where requirements impose costs on firms in order to achieve social or ecological goals. We call these "costly issues." On these issues, an industry-backed program must balance achieving legitimacy through perceived stringency with minimizing compliance costs. The result is likely to be a lower level of stringency than that of an activist-backed program, especially where industry-backed programs can more easily create an impression of stringency or where compliance costs are high.

The opposite result is likely on issues where requirements provide net benefits to the industry. Here, activist-backed programs have little incentive to develop stringent requirements because activist pressure is redundant. These "business-friendly" issues are frequently addressed by industry associations. They include coordinating resources and solving collective action problems related to industry reputation (e.g. through public image campaigns) and capacity (e.g. by developing collective goods like technical knowledge or a skilled workforce). By "collective action," we merely mean actions across individuals that have net benefits but that requires a coordinating institution. Another business-friendly strategy to create perceptions of stringency may be to add requirements to do things that firms would do anyway. If observers fail to distinguish among different types of stringency on different issues, such a strategy may be a low cost and effective way to shape perceptions of overall stringency.

Revised differentiation hypotheses: Where activist-backed and industry-backed private regulations compete,

- **H2.1:** Activist-backed regulations will have more comprehensive coverage, more prescriptive requirements, and higher performance thresholds on costly issues.
- **H2.2:** Industry-backed regulations have more comprehensive coverage, more prescriptive requirements, and higher performance thresholds on business-friendly issues, such as those that firms do anyway or those related to industry collective action problems.

These hypotheses illustrate how many of the theories reviewed in section 2 could be

revised in light of our core methodological critique. One may assess whether our framework is useful in two ways: (1) Does restating theories in terms of the predicted direction of change in scope, prescriptiveness, and policy settings improve our understanding of past research? (2) Does applying the framework reveal patterns of change that other methods failed to discover? Sections 4 and 5 show that our framework meets both tests: its application reveals that the scope, prescriptiveness, and policy settings of forestry certification programs do follow different patterns and that existing theories cannot fully account for these changes.

# 4 COMPETING US FOREST CERTIFICATION PROGRAMS

We illustrate our methods for measuring stringency through an analysis of forestry certification in the United States, one of the most advanced cases of private regulation. Like many substantive domains, forestry scholars have carefully dissected components of forestry regulations, both public and private. Yet the unit of analysis in political science scholarship still tends to be broad characterizations of entire policies or only a few of their constituent parts. By drawing on domain-specific scholarship, we conduct a more systematic analysis. The results of this analysis offer the most comprehensive and detailed description of changes in forestry certification standards to date.

Forest certification illustrates how market-based authority can involve formal decision-making modeled on government rulemaking processes, legalistic requirements, and powerful enforcement mechanisms. When product certification programs gain power with consumers and retailers, a timber company's contracts may depend on an audit of their compliance with hundreds of requirements. Noncompliance may be costly. For example, Resolute Forest Products claimed damages of $100 million CND related to auditor findings of non-conformance (Tigar, 2017). This scale of impact makes forest certification an important case.

For over 20 years the Forest Stewardship Council (FSC) and Sustainable Forestry Initiative (SFI), have been developing written Forest Management Standards (standards) that promote different conceptions of "sustainable" forest management. The SFI and FSC play a major

role in regulating the forest products industry in the United States, regulating a third of commercially harvested timberland including most corporate-owned timberland (see Figure 1). Many U.S. states support certification as a compliment or alternative to public regulation. For example, some state regulators forgo inspections of FSC-certified forests as legal compliance is part of their FSC audit (Judge-Lord, 2013).

Figure 1: U.S. Timberland by ownership and certification scheme
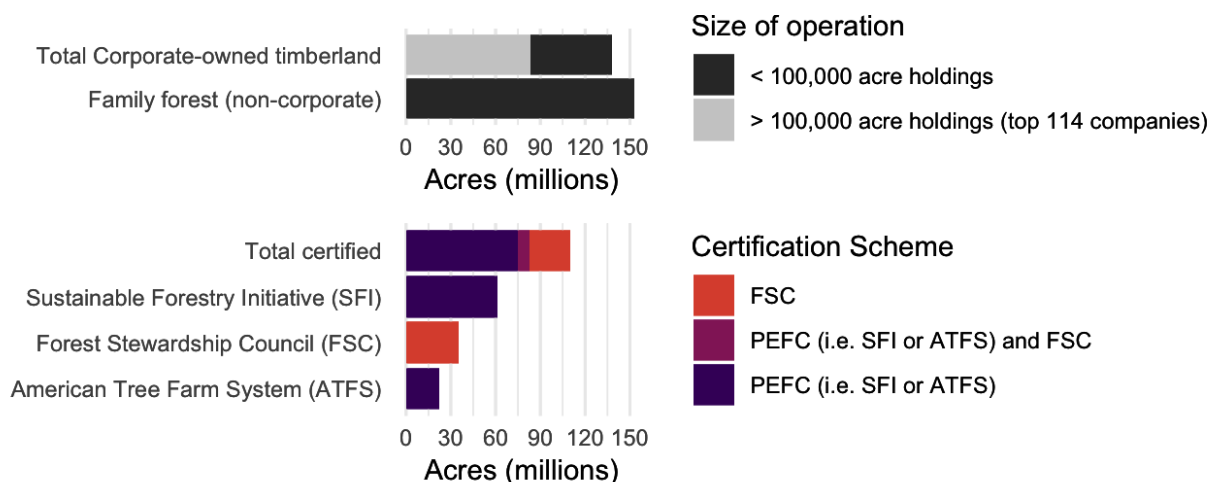
Sources: AF&PA, 2016; PEFC, 2017; U.S. Forest Service, 2009

Origin of the FSC: The FSC was established as an international non-profit organization in 1993 by a group of environmental and social advocacy organizations, academics, indigenous groups, and companies. FSC's founders designed its rulemaking procedures as a "democratic" process where members vote on decision-making rules as well as substantive policy (Meidinger, 2003). FSC standards begin with a set of international "Principles and Criteria" (FSC–P&C) that are used by national-level organizations to develop more specific indicators. Our analysis of the FSC in the U.S. thus assesses both the international P&C and the FSC-US national standard.

Origin of the SFI: In 1995, in response to the emergence of the FSC, the U.S.-based industry association, the American Forest & Paper Association, established a forest management standard and required its members (most of the U.S. forest products industry) to support it. Optional third-party auditing was added in 1998, which became mandatory in 2002, the same

year that the American Forest & Paper Association made the SFI a legally distinct entity with a rulemaking process that is formally independent, though still largely governed by business stakeholders. The SFI has since been endorsed by the global Program for the Endorsement of Forest Certification (PEFC). The PEFC maintains a set of Sustainable Forest Management Benchmarks intended to guide participating programs, many of which are industry-backed alternatives to the FSC. Unlike the FSC–P&C, the PEFC does not require the SFI and other national-level programs to adopt its benchmarks verbatim. Instead, they are expected to demonstrate the "equivalence" of their standards with PEFC benchmarks.

"Sustainable" Forestry: Like many sectors, there are ongoing debates over acceptable business practices and the appropriate role of public and private regulation in forestry. "Sustainable" forestry has many meanings (McDermott, 2012). For example, some programs use "natural" conditions or functions as benchmarks for sustainability, involving complex choices about what is "natural" and what degree of naturalness is appropriate. In other conceptions, "sustainable" is less associated with naturalistic management and more about the long-term efficiency of production. Such differences manifest in distinct goals and different means to achieve them. A regulation targeting efficiency may require high levels of utilization of trees and tree-parts, whereas a regulation targeting naturalistic management may include requirements to leave economically valuable timber behind for animal habitat or soil health. Disagreements become concrete in the details of such requirements. Thus, a meaningful assessment of similarities and differences between regulations requires attention to detail.

4.1 Scope, prescriptiveness, and policy settings of forest certification programs

To measure comprehensiveness of scope, we reviewed all FSC, PEFC and SFI standards in effect between 2008 and 2016 to assess their coverage across 48 distinct "key issues" covering a broad scope of forestry requirements, from employee wages and resource utilization to protections for endangered species and indigenous peoples' rights. These issues were selected

in 2008 using an iterative process to disaggregate forestry policies to capture all of the major issues addressed by FSC, PEFC, and SFI requirements (McDermott et al., 2010).

To measure prescriptiveness, we assess the precise wording of the text on each issue. If firms have discretion among performance levels, only the least demanding levels are prescriptive. For example, if firms are required to "maintain or enhance" water quality, the option to merely "maintain" means that there is no mandatory requirement to "enhance" water quality.

To measure policy settings, we offer detailed issue-by-issue comparisons of performance requirements on most of our 48 key issues in the text below and all of them in the online appendix. This approach is similar to how previous scholars have descriptively compared the SFI and FSC standards on select sets of issues, except with a comprehensive scope of potential issues. Doing so allows us to classify each specific change, the types of issues that changed, and difference on issues that may be important but not (yet) salient in the public debates.

4.2 RESULTS

Here we compare each standard to its previous version and the contemporary version from its competitor. We assess revisions in the FSC- International's 2012 Revised Principles and Criteria 01-001 Version 5-0 (FSC–P&C), and we compare them to the PEFC's Sustainable Forest Management Standards (1003:2010). Similarly, we compare the 2010 FSC-U.S. Forest Management Standard Version 1.0 to the FSC-US National Indicators and regional standards it replaced, and we compare these to the 2005-20092010–2014, and 2015-2019 SFI standards. Unless otherwise specified, "FSC-US" and "SFI" refer to the version of each standard in effect in 2016. We do not fully capture subnational variation. The FSC-US standard recognizes nine different sub-national regions. Some have additional indicators, meaning that, in some states, FSC standards were more prescriptive than our findings reflect (see online appendix).

4.2.1 Comparing FSC's and PEFC's international requirements

Scope: The FSC-P&C and PEFC maintained a similar scope of issues covered (see the top panel of Figure 2). The PEFC once covered slightly fewer issues than did the FSC-P&C, but its 2010 revisions added new requirements for eight key issues that it previously did not address, making the two generally aligned in the scope of issues covered. As of 2015, the FSC P&C covered three potentially "costly" issues that the PEFC still did not: carbon emissions, restrictions on conversion to plantations, and worker wage requirements (see the middle panel of Figure 2). PEFC covers two issues relating to public perceptions that FSC-P&C do not: managing the aesthetic impacts of forestry and allowing public access.

Prescriptiveness: Overall, the FSC maintained more prescriptive requirements in its Principles & Criteria than the PEFC benchmarks (the top panel of Figure 2), but the PEFC moved closer to the FSC-P&C in some key areas (the middle panel of Figure 2). These include additional requirements on issues including indigenous rights, community benefits, and public reporting and consultation (see the online appendix for specific language). The PEFC became at least as prescriptive as the FSC–P&C on over half of key issues. In absolute terms, the PEFC increased prescriptiveness on 19 key issues and decreased on none, whereas the FSC–P&C increased on 13 and decreased on four. Yet significant differences remain. The FSC-P&C contain more prescriptive language on most ecological criteria, including protected areas and restrictions on conversion to plantations. On the whole, both programs have more procedural requirements than substantive requirements (i.e. they are more focused on process than outcomes). Despite convergence in the PEFC's revised requirements, the FSC-P&C remained more prescriptive than PEFC requirements on 17 of the 48 key issues whereas PEFC requirements are more prescriptive on nine issues, with both programs being equally prescriptive on 19 issues. Because the PEFC started at a lower level but increased prescriptiveness on more issues than did FSC P&C, the resulting pattern is an "upward convergence" (the bottom panel of Figure 2).

Policy settings: One particularly controversial issue is the conversion of natural forests to

Figure 2: Scope and Prescriptiveness of FSC P&C and PEFC 2008-2015

23
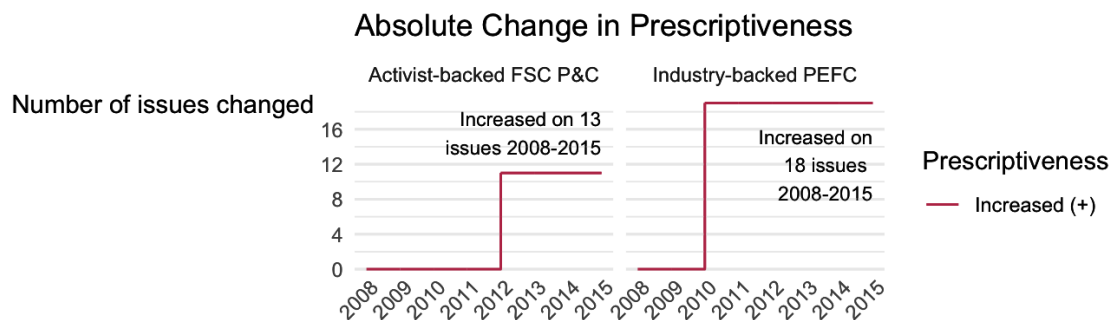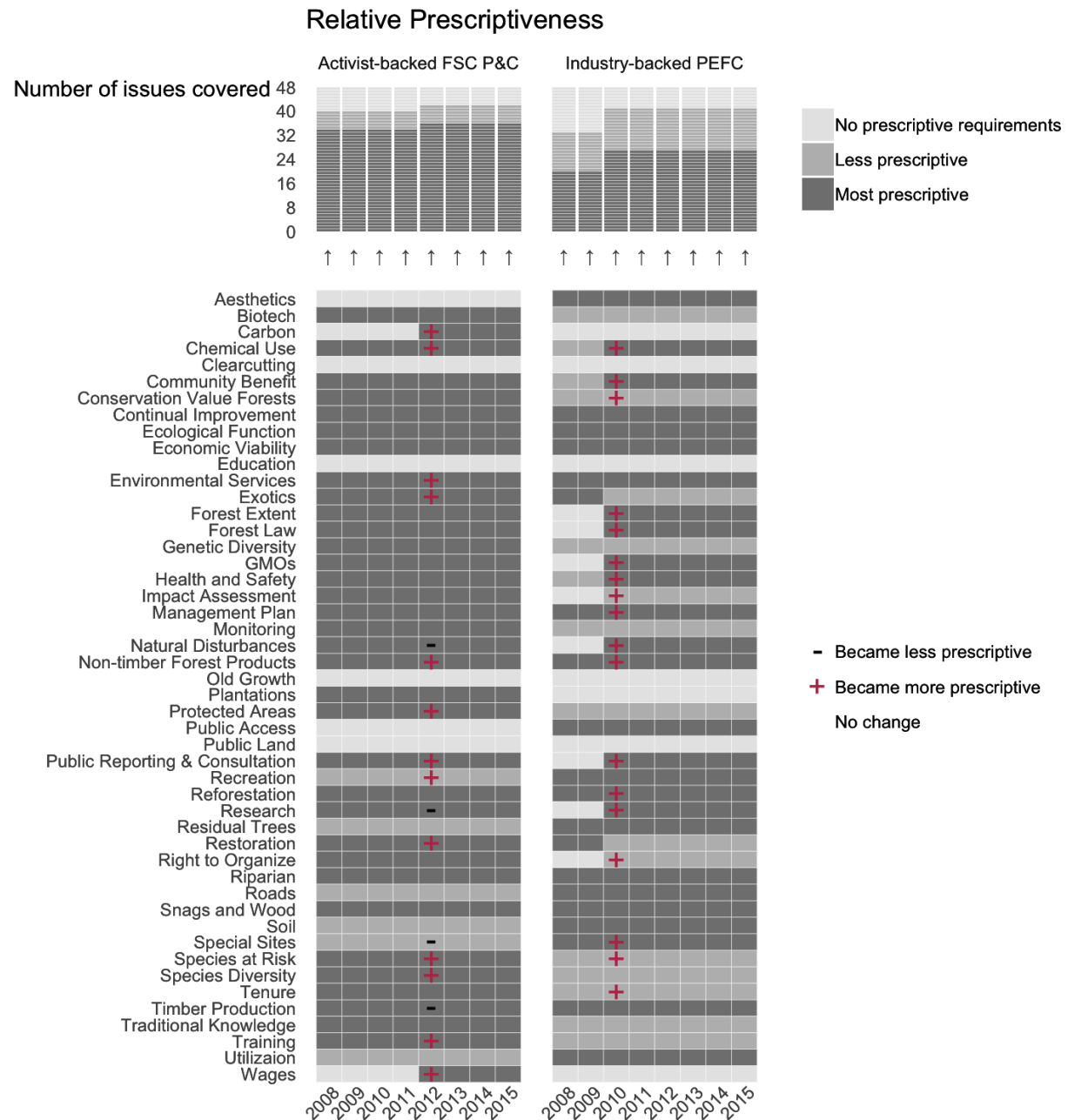
timber plantations. Both programs only allow conversion of natural forest to plantation under "justifiable circumstances," which differ qualitatively between the two programs. For the FSC, this means that conversion has "clear, substantial, additional, secure, long-term conservation benefits." For the PEFC it means that conversion must have "long-term conservation, economic, and social benefits." They also differ regarding the extent of forest conversion allowed. The FSC-P&C allow companies to convert "limited areas" while the PEFC allows "small proportions of forest types." Both standards specify that conversion must not damage culturally or socially significant areas, but whereas PEFC suggests that forests should only be certified if the conversion occurred before 2011, the FSC-P&C require that conversion occurred before 1994, significantly different thresholds.

Both FSC-P&C and PEFC added new requirements on socio-economic issues, land tenure rights, and stakeholder consultations. In addition to citing the UN Declaration on Rights of Indigenous Peoples, both included criteria that require free, prior and informed consent of indigenous peoples and local communities. The FSC-P&C referenced "free and informed consent" concerning control over forest operations and compensation for the use of traditional knowledge. Both standards also recognized legal, traditional, and customary rights. However, the FSC-P&C are more prescriptive, defining the topics that require consultation with indigenous peoples, while the PEFC standards are more procedural, requiring only that engagement takes place. The FSC's criteria regarding public consultation include special obligations to "affected stakeholders" compared to "interested stakeholders" while PEFC requirements for "local people and other stakeholders" are the same.

Both programs cover similar ecological issues, with some qualitative differences. Both FSC–P&C and PEFC requirements prohibited the use of GMOs in the area being certified, with some possible flexibility should scientific evidence affirm the safety of GMO trees. FSC–P&C allow documented and monitored use of biological control methods but prohibited a specific list of "Highly Hazardous Chemicals." The PEFC added prohibitions on pesticides that remain biologically active and highly toxic pesticides where viable alternatives are

available. The PEFC explicitly stated that managers should avoid chemicals where they threaten water quality, while FSC–P&C water protection criteria were less explicit. The FSC–P&C and PEFC had similar requirements for sustainable production of timber and non-timber forest products (NTFPs), but the FSC-P&C set a higher level of protection for animal habitat. While the FSC–P&C required protection of rare and threatened species and their habitats, the PEFC only required that protected and endangered species not be exploited for commercial purposes and that managers take measures for their protection "where necessary," without defining these conditions.

Summary: Overall, while the PEFC added more requirements concerning indigenous rights and labor standards and came to cover a similar scope of issues to the FSC P&C, the FSC-P&C remained more prescriptive on social issues and significantly more prescriptive on ecological issues. Compared to the prescriptiveness of the FSC-US and SFI described below, the FSC–P&C and PEFC requirements exhibit more convergence on both scope and prescriptiveness (compare Figures 3 and 4) though many differences in policy settings remain.

4.2.2 Comparing the FSC-US and SFI Scope: Consistent with the international level, the activist-backed FSC-US program and industry-backed SFI program in the United States address a similar scope of issues, but the FSC-US is more prescriptive on most (the top panel of Figure 3). As of 2016, the FSC-US did cover six potentially "costly" issues that the SFI did not; community benefit requirements, forest extent restrictions, required impact assessments, protected area restrictions, restoration requirements, and indigenous tenure protections (the middle panel of Figure 3). The SFI, in turn, covered one issue contributing to forestry research, that the FSC-US did not. Both programs added requirements on greenhouse gasses in 2010. SFI allows for the conversion of natural forests to plantations if ecological impacts are not significant and the converted forest type is not rare, but in 2015, SFI added a requirement to conduct an assessment of these impacts. Yet, the FSC-US still maintained more prescriptive requirements, only allowing certification of plantation forests if they were converted from natural forests before 1994. FSC-US also requires a portion of

these plantations to be maintained as, or restored to, natural conditions.

Prescriptiveness: In 2008 the FSC-US was more prescriptive on 36 of 48 key issues, and the SFI was more prescriptive on five issues. These five are some of the most "business-friendly" issues: continual improvement of management planning, educating the public about forestry, contributions to forestry research, worker training, and efficient material utilization. In 2016 the FSC-US was more prescriptive on 37 key issues, and the SFI was more prescriptive on the same five issues. The two standards were equally prescriptive on five issues. This means that the FSC-US had the "most prescriptive" requirements—those as prescriptive or more than any other program—on 42 issues and the SFI had the most prescriptive requirements on 10 (the top panel of Figure 3).

Counting changes made to the FSC-US and SFI standards between 2008 and 2016 reveals an "upward diverging" pattern, where the FSC-US became more prescriptive than did the SFI (the bottom panel of Figure 3). Of 48 key issues, the FSC-US became more prescriptive in 20, whereas SFI became more prescriptive in 12, eight in 2010, one more in 2013, and three more in 2015.

Policy Settings: Issues such as clearcut size limits and limits on harvesting near streams clearly illustrate enduring differences between the SFI and the FSC-US because policy settings on these issues can be compared both qualitatively and quantitatively. Qualitatively, the FSC-US increasingly restricts the size and shape of clearcuts to reflect "natural disturbance" and maintain ecological functions regardless of how it looks, whereas the SFI emphasizes "the visual impacts of forestry" and requires rapid site "green-up." Quantitatively, the SFI limited clearcuts for all forest types to an average of 120 acres with no maximum and no limits for harvesting with 20 percent tree retention (i.e. intensive but not clearcut harvesting). In contrast, the FSC-US limits clearcuts to a 40-acre average and 80-acre maximum, with additional restrictions based on region and forest type. FSC-US also limits harvesting with 20 percent tree retention to a 100-acre average and 80-acre maximum, with additional restrictions based on region and forest type (Figure 4). For harvesting near streams, the

Figure 3: Scope and Prescriptiveness of FSC-US and SFI 2008-2016

Relative Prescriptiveness

Absolute Change in Prescriptiveness

27
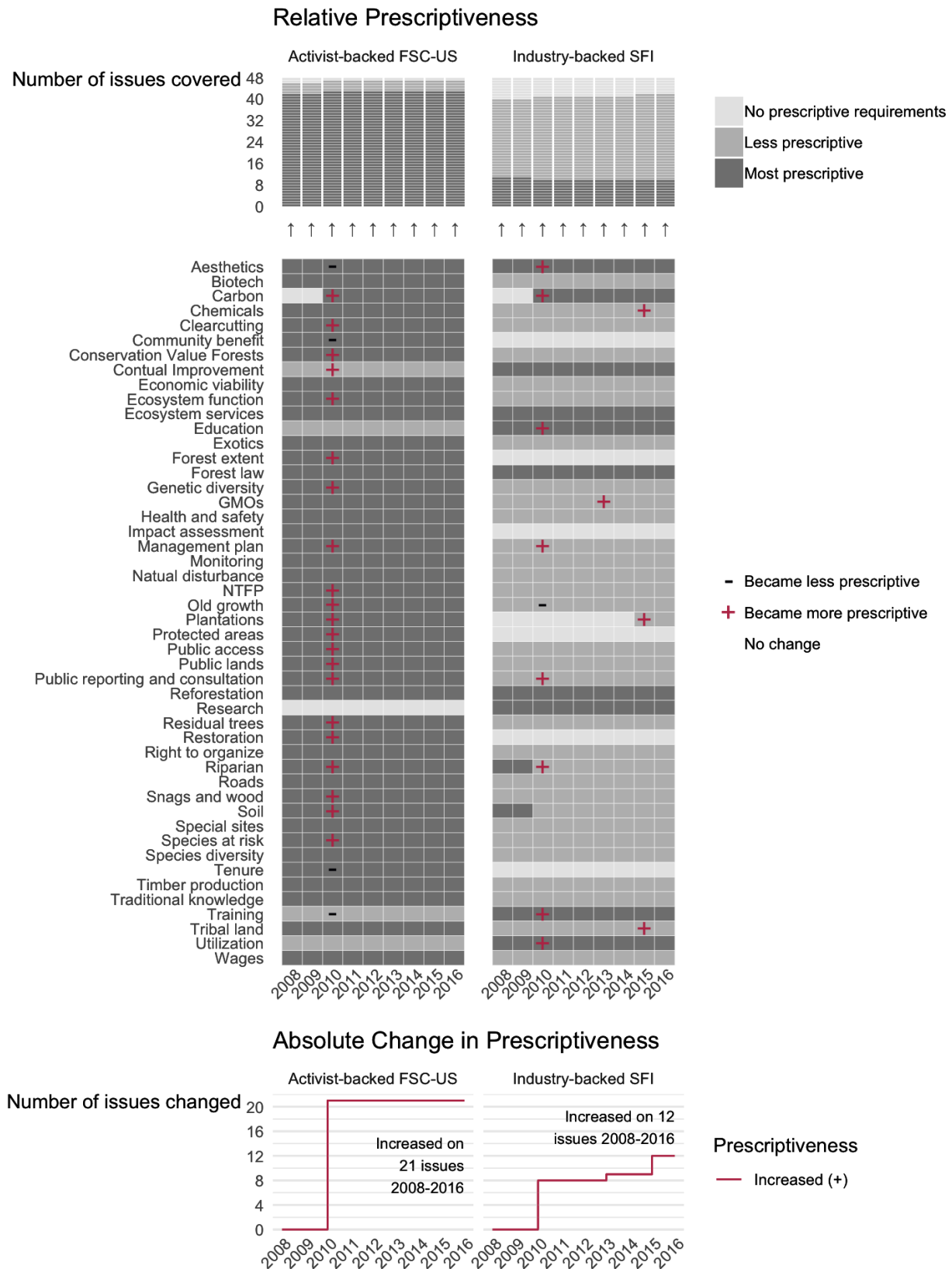
FSC-US lists specific requirements for water quality, habitat, and other objectives with a focus on restoration. Additionally, most FSC-US regions have numeric minimum riparian buffer zones (Figure 5). In 2015, SFI expanded its definitions of riparian areas but continued to allow more discretion regarding what managers include in plans to protect water resources with no numerical minimums beyond those in state laws and best management practices. While other policy settings can only be compared qualitatively, the FSC-US clearly requires higher levels of performance on many social and ecological issues (Table 4)

Figure 4: Limits on Clearcut Size
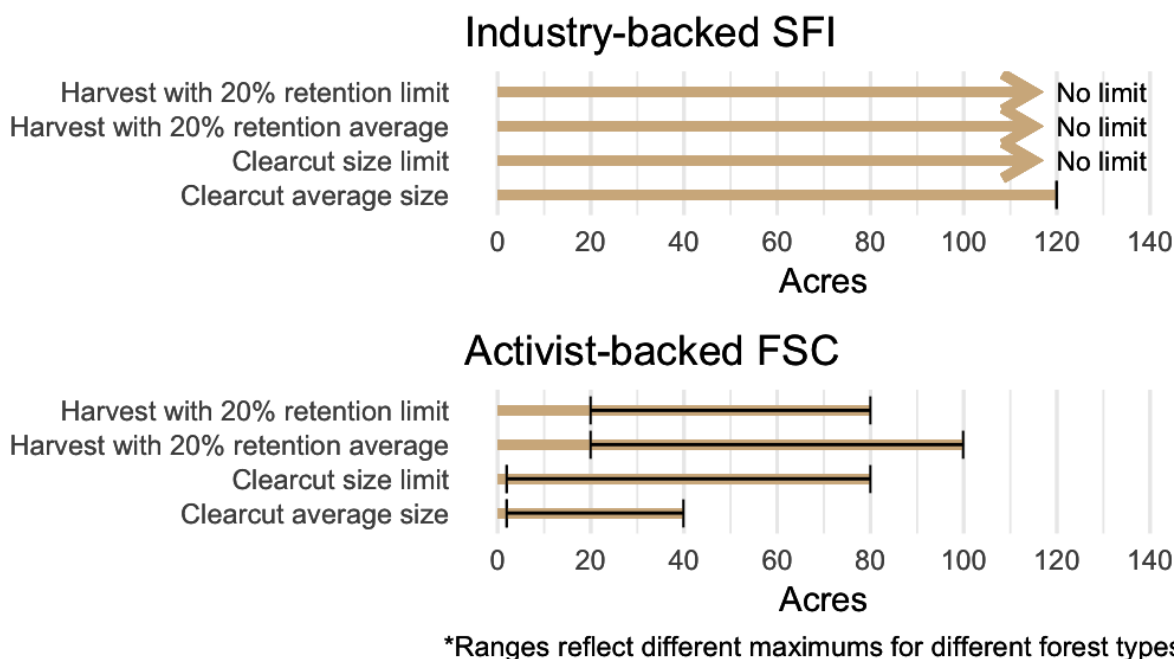
While both the FSC-US and SFI became more prescriptive, they did so to different degrees and in different areas. The SFI's changes in 2010 emphasized issues related to industrial capacity (e.g. worker training requirements) and reputation (e.g. managing the visual impact of harvesting, communicating with stakeholders about logging, and educating the public about forestry), issues where SFI already had the most prescriptive requirements. Changes made the same year by the FSC-US emphasized conservation-oriented forestry while removing a training requirement.

The bulk of the divergence occurred on ecological requirements like protecting habitat,

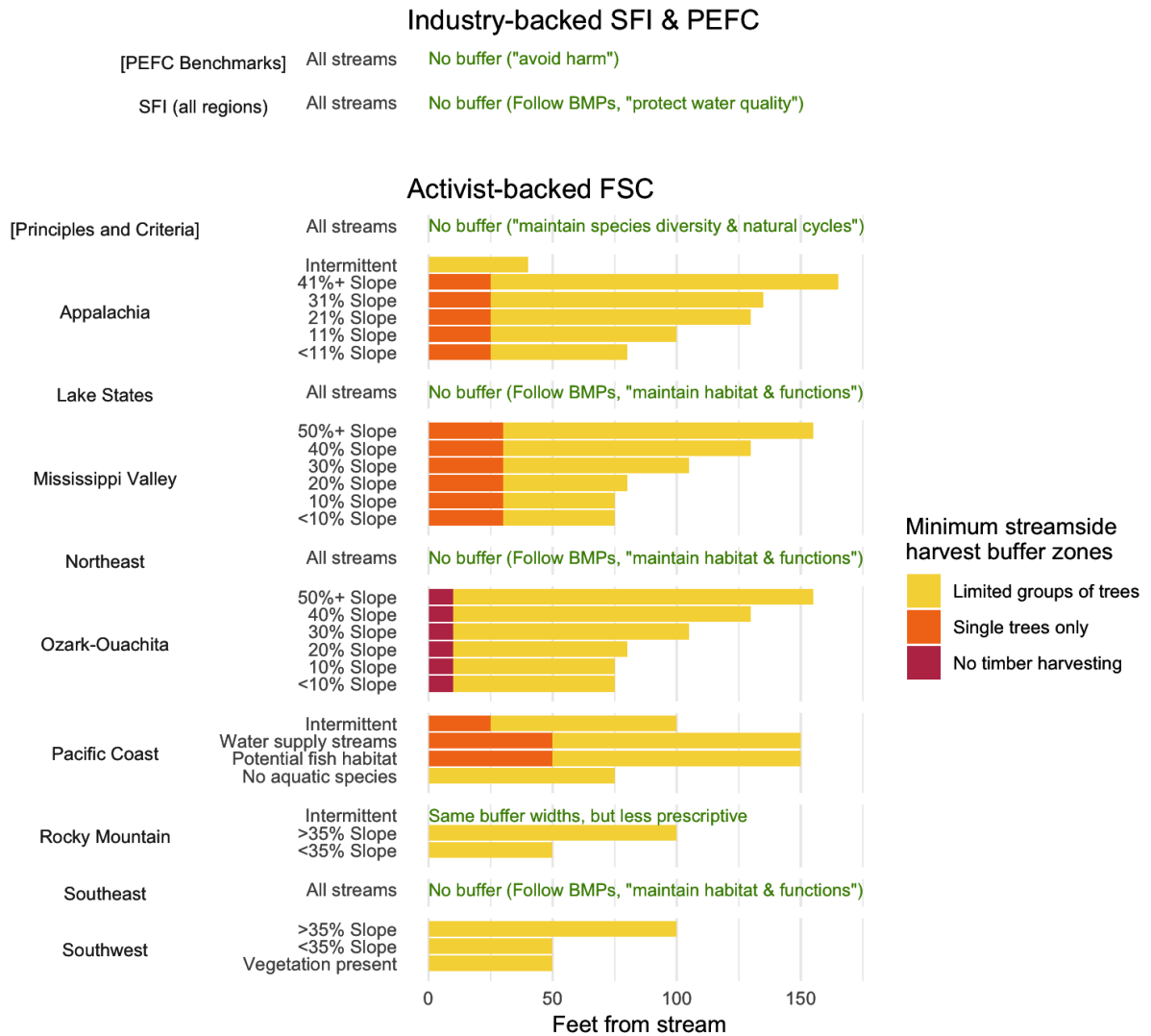Figure 5: Limits on Harvesting Near Streams

Table 4: Qualitative levels of performance required

Issue	Activist-backed FSC-US	Industry-backed SFI
Indigenous peoples' rights	Recognize and uphold rights, customs, culture including UNDRIP. No threat to rights or resources. Free, prior, and informed consent on public and private lands. Engage indigenous peoples and consult with affected groups. Cooperate to identify and protect significant sites. Compensate for indigenous knowledge and utilize as requested.	A written policy acknowledging a commitment to recognize and respect rights.
Public Reporting and Consultation	Required on public and private lands.	Required on public lands.
Forest conversion to non-forest	Prohibited except limited areas where clear, substantial, additional, secure, long-term conservation benefits.	No specific policy.
Old growth forest	Old growth is normally mapped as conservation forest. Only restoration management on public land. Legacy trees not harvested. Maintain structure, composition, and processes. A portion of the forest is restored where old growth would naturally occur.	Support and participate in programs for old growth conservation in the region — no identification or restoration requirements.
Protected areas	Conserve or restore a representative area of natural ecosystems. Assess and maintain environmental values and necessary conservation measures.	No specific policy.
Threatened and Endangered Species	Survey and report or assume the presence of vulnerable, imperiled, and critically imperiled. Maintain habitat & viable populations.	Program to protect threatened and endangered species and known sites. Protect viable occurrences of critically imperiled or imperiled species.

where the FSC-US became more prescriptive while the SFI stayed constant or, in the case of preserving old-growth forests, decreased in prescriptiveness. Regarding protected areas, the FSC-US continued to require that managers preserve representative samples of habitats, but, since 2010, also requires an assessment of the adequacy of permanent protections. SFI's requirements for protected areas continue to be encompassed mainly by its requirements to protect imperiled species. SFI continues to require plans to identify and protect moderately to highly valuable known populations of imperiled or critically imperiled species (designations G1-G2). In contrast, the FSC-US expanded the scope of species requiring protection in 2010 to include natural heritage species and candidate species (designations G1-G3, S1-S3, N1-N3). The FSC-US added requirements to conduct surveys for any at-risk species potentially present or presume that listed or candidate species are present if the forest is in a species' range. For old-growth forests, in 2010, the FSC-US added prescriptive requirements to restore a portion of old-growth forests where they would naturally occur, and it continues to demand protection measures that prohibit harvesting in most cases. In 2010, SFI removed a requirement to maintain sufficient old-growth acreage to maintain biodiversity, but in 2015 added a requirement to participate in conservation planning.

The FSC-US and SFI's changing requirements to designate and preserve conservation areas exemplify their overall upwardly diverging prescriptiveness, with the SFI adding some prescriptive requirements but the FSC-US adding even more prescriptive requirements. In 2010, the SFI added new requirements to collect data on "Forests of Exceptional Conservation Value" (FECV), which we compare to the FSC's requirements for "High Conservation Value Forests" (HCVF). Also, in 2010, the FSC-US added language regarding monitoring and adaptive management of HCVFs. While the acronyms and even the additional language appear similar, the FSC-US added more prescriptive requirements requiring certain areas to be designated HCVFs and prescriptive accountability mechanisms for HCVF management. SFI allows more flexibility in FECV management. HCVFs under the FSC-US require significantly more than baseline practices (Newsom and Hewitt, 2005), while SFI's FECV requirements

have been criticized as not significantly exceeding legal baselines which already protect threatened and endangered species.

Summary: Overall, each program had distinct areas in which its requirements were more prescriptive. The FSC's requirements tend to demand that forest operations "resemble natural processes" and "maintain ecosystem function." This language appears more frequently and forcefully in the 2010 standard concerning issues including clearcutting, riparian management, HCVFs, protected areas, old-growth forests, snags and downed wood, residual trees, genetic diversity, plantations, restoration, natural disturbance, non-timber forest products, soil protection, road building, and management planning. In contrast, the SFI was most prescriptive on issues such as material utilization, research, training, education, and public reporting and consultation. The eight key issues on which the SFI increased prescriptiveness in 2010 also reflect the SFI's focus on industry capacity and reputation. These included aesthetics, public reporting, education, training, and utilization.

The 2015 changes to the SFI standard reflect a different tack. In contrast to the previous focus on more "business-friendly" issues related to industry capacity and reputation, the three issues on which the SFI increased prescriptiveness in 2015 reflect social and ecological goals. These include prohibiting the use of certain toxic chemicals, restricting the circumstances under which natural forest can be converted to plantation, and requiring a written policy to recognize and respect indigenous rights.

5   Discussion

5.1   Overall comparison

By distinguishing different types of stringency on a comprehensive set of issues, our framework improves upon blunt measures of "high" or "low" based on generalizations or only a few issues. We now aggregate our issue-specific results to reflect on more general trends:

One way to group issues is by ecological goals, social goals, and business-oriented goals.

For the purposes of evaluating our hypotheses, we consider business-oriented goals to be efforts that provide net utility to the industry while social and ecological goals generally inflict costs for broader aims.

On ecological goals, the FSC-US standard was significantly more stringent than the SFI standard on both scope and prescriptiveness dimensions. On social goals, results are more mixed. On scope, the FSC-US standard protects land tenure and requires that local communities benefit from harvesting in ways that were unmatched by SFI's standard, but the SFI requires contributions for forestry research, which the FSC does not. Numerically, one could say that FSC-US had a broader scope of social benefits (depending on what issues one considers "social"), but the programs do present tradeoffs between conceptions of the public good. On prescriptiveness, the contrast is again more stark, with the FSC-US standard having significantly more prescriptive requirements on most social issues. On policy settings, the two standards have significant differences. On labor standards and indigenous rights, the FSC-US standard required higher wages and more attention to rights than the SFI standard did. In short, by conventional definitions of what counts as a social issue, by most qualitative comparisons, and certainly in terms of prescriptiveness, the FSC-US standard is more stringent than the SFI standard on social issues.

On more business-oriented goals such as efficiency (e.g. levels of cut tree utilization), industry capacity (e.g. workforce training and research), and industry reputation (e.g. public education and aesthetics), the conclusions are largely reversed. SFI is slightly broader in scope, requiring contributions to research where FSC does not, is more prescriptive, and requires increasingly demanding performance levels on many business-friendly issues.

5.2 PATTERNS OF CHANGE

The dominant patterns were equilibrium and upward divergence. In most years, neither program changed on any issue (the center cell in Table 3, "equilibrium"). Most changes for both programs occurred in 2010 where the overall pattern was divergence (also called

differentiation), rather than convergence or stability. The vast majority of changes (twenty-one of twenty-seven issues changed) fit a pattern where one program increased prescriptiveness while the other did not (or in one case, did increase to a lesser degree) and the program that increased stringency already had the more prescriptive requirements. On eighteen issues, the other program stayed the same, leading to upward divergence. On three issues, the less prescriptive program decreased prescriptiveness, leading to opposing divergence (see Table 5).

For all sixteen issues on which only the FSC-US added requirements, it already had the more prescriptive requirements, and almost all of these additions address ecological problems. Similarly, for three out of the four issues on which only the SFI added requirements, the SFI already had more prescriptive requirements. Qualitatively, these three issues—maximizing the utilization of cut trees, public education, and worker training—reflect concerns for the efficiency, reputation, and capacity of the forest products industry. Educating the public about forestry practices and products and training workers may not exclusively benefit individual firms. Because of the broad adoption of SFI standards, such requirements may provide collective benefits for the sector, in the form of a positive public image and skilled workforce.

Table 5: 2010 and 2015 Patterns of Change in Prescriptiveness among U.S. Forestry Certification Programs (FSC-US and SFI) on 48 Key Issues

2010

	Converging	Parallell	Diverging
Increasing	1	3	18
Opposing or Eqilibrium	0	21	3
Decreasing	2	0	0

2015

	Converging	Parallell	Diverging
Increasing	3	0	0
Opposing or Eqilibrium	0	45	0
Decreasing	0	0	0

Convergence and parallel change were rare. An upward parallel change occurred on only three issues in 2010: forest management planning, controlling carbon emissions, and reporting and consultation, where both programs added requirements. We classify the addition of protections for riparian zones by both SFI and FSC-US as another case of upward divergence rather than upward parallel change because the requirements for riparian protection added by the FSC-US are more prescriptive. Upward convergence only occurred where FSC-US added requirements on the issue of "continual improvement" of harvesting operations, an issue usually associated more with the SFI. This outcome is interesting because scholars generally predict that private regulations that are less stringent overall will converge toward "benchmark" standards like FSC's (Overdevest, 2005, 2010). Instead, we find the FSC ratcheting up on an issue where its industry-backed competitor had more stringent requirements. Indeed, most studies overlook the possibility that industry-backed standards like the SFI may be more stringent on some issues. We see downward convergence only on Community Benefits and Tenure Rights where the more prescriptive FSC-US removed requirements, thus moving closer to SFI. No issues exhibited downward parallel or downward diverging trajectories.

Since the significant revisions of both programs in 2010, only SFI has updated its national-level requirements, mostly in 2015. In contrast to the 2010 changes, the pattern in 2015 is a moderate upward convergence. SFI increased prescriptiveness on three issues where it did not already have the most prescriptive requirements. While a smaller scale of change than 2010, this upward convergence is notable because it focuses on regulating toxic chemicals, plantations, and harvesting on tribal lands, which likely have net costs rather than benefits for the industry.

Overall, the dominant pattern of change from 2008 to 2016 is upwardly diverging prescriptiveness and no change in scopes. Qualitatively, the upward diverging pattern results from the activist-backed FSC increasing prescriptiveness on ecological protection and the industry-backed SFI on issues that provide net benefits to the forestry sector. These results in hand, we can compare them to the hypotheses presented in section 2.3.

5.3 Implications for theory

Applying our framework to the case of forestry certification reveals how one could reach different conclusions by looking at different dimensions of change. If focusing only on program scope, one would find little support for any theory predicting change—either convergence or divergence. If focusing just on prescriptiveness on ecological issues, one would find divergence, with the activist-backed FSC-US becoming more prescriptive at a faster rate than the industry-backed SFI. But if focusing only on prescriptiveness of issues of industry capacity and reputation, one would find the opposite, with the SFI becoming more prescriptive at a faster rate than the FSC-US.

Anticipating such possibilities (indeed the literature is rife with such seemingly contradictory results), we restated several leading hypotheses in ways that distinguish scope and prescriptiveness in section 2.3 (Example Hypotheses 1.1 and 1.2) and distinguish costly and business-friendly issues— (Example Hypotheses 2.1 and 2.2). While fully testing any causal explanations of policy change is beyond the scope of this paper (our focus is measuring the dependent variable), we can reflect on whether our measurements are consistent with these restated hypotheses.

Assuming that changes in scope are less costly than changes in prescriptiveness, our findings are somewhat consistent with Example Hypothesis 1.1. Across social, ecological, and business issues, the activist-backed program and industry-backed program slightly converged on scope and clearly did not converge on prescriptiveness. If "talk is cheap," but prescriptive requirements are costly, it makes sense that an industry-backed program would add language similar to that of an activist-driven standard without adopting costly mandatory performance thresholds. We found such a pattern on many issues. This result suggests that studies aiming to test theories rooted in the cost of compliance must distinguish measurements of their dependent variable based on the dimensions of scope or prescriptiveness.

Our findings are also somewhat consistent with Hypothesis 1.2, again more clearly for prescriptiveness than for the scope of issues addressed. We observe overall divergence on

prescriptiveness, most clearly on ecological issues. As neither program changed significantly on overall scope, we cannot tell whether changes in scope are more likely to be matched by competing programs. Both programs did add requirements regulating carbon emissions in 2010, but it is unclear if this change in scope is one program reacting to the other or both programs responding to a third causal factor.

Our findings most clearly align with hypotheses 2.1 and 2.2. The activist-backed FSC-US has a slightly more comprehensive scope and much more prescriptive requirements on activist-driven issues, while the industry-backed SFI standard has a more comprehensive scope and more prescriptive requirements on issues more related to industry collective action problems. Additional research should further test these and other hypotheses, using similarly precise and comprehensive measures of regulatory stringency.

5.4 Industry-backed certification programs as a form of collective action.

Our finding that the SFI was more prescriptive and continued to become more prescriptive than the FSC-US on certain issues highlights how industry-backed certification programs can serve their industry in two ways. First, they provide individual firms with a service—market signals of "social responsibility" that require a credible third party. These signals would be more expensive to send by complying with an activist-backed regulation. Second, they provide a mechanism for the industry to improve its collective reputation and capacity by coordinating contributions to collective goods, a common function of industry associations. While change on costly issues likely driven by competition among standards, on business-friendly issues, activist pressure is not necessary. Indeed, firms often collaborate on the latter type of issues (e.g., industry standards, reputation, and capacity) through trade associations without pressure from activists.

Regarding the first, Cashore et al. (2004) point out that industry-backed alternatives aim to save firms money by offering a label that sends a "green" or "socially responsible"

signal in the market without some of the more costly demands of activist-backed programs. Credible signals are often based, in part, on perceived stringency, but because perceived and actual stringency may differ, industry-backed programs can send "credible" signals while maintaining lower compliance costs. Labels like SFI are not necessarily "meaningless" or pure "greenwashing"—indeed a certain level of stringency is often required to maintain legitimacy—but exceeding this "floor" imposes costs on firms. On many issues, industry-backed programs may succeed in creating the necessary impression of equivalence to the stringency of activist-backed standards with substantially less prescriptive requirements. For example, the SFI requirements for "Forests of Exceptional Conservation Value" (FECV) are much less prescriptive than the FSC-US requirements for "High Conservation Value Forests" (HCVF), despite the similar name. By supporting alternative programs, firms collectively create institutions that help them maximize the impression of stringency while minimizing the costs of doing so. This dynamic describes most key issues in our study.

Regarding the second, the fact that SFI developed more prescriptive standards than did FSC-US on several issues is inconsistent with the predictions that competition between industry-backed and activist-backed competition will lead to a "race to the bottom" on all issues. It is also inconsistent with the prediction that activist-backed standards be more prescriptive on all issues. However, the substance of these issues suggests that these requirements are unrelated to competition with the FSC. SFI has the most prescriptive requirements for actions that firms may take anyway—like training and efficiency—or that may be driven by their own collective action problems—like managing the visual impact of harvesting and sector-level reputation.

In sum, where the SFI has more prescriptive requirements than the FSC, it requires things that many firms may do anyway (e.g., train workers or educate the public), but have added collective benefits the more widely they are adopted. While unforeseen by existing theories, the fact that the SFI is more prescriptive on some issues is not surprising when we understand these issues as providing net benefits to the sector regardless of activist pressures

or consumer demands.

6   CONCLUSION

Scholars have made substantial progress in developing theories of how economic and political forces shape the substance of private regulations, and how these different requirements then affect levels of adoption and compliance. We have argued that testing these theories requires more precise statements of the types of policy substance to which they apply and research that measures change across programs and over time. Our framework for measuring policy substance, and for using longitudinal data to classify patterns of change, offers a foundation for further research about how competing private regulations compare, how they evolve, and why. There is no perfect way to compare incommensurate policies. We have nonetheless made our best effort to offer a method to do so. By applying this method, we have quantified differences that can be quantified and described as richly as possible those comparisons that one can only make qualitatively.

Through the case of forestry standards in the U.S., we show what can be gained by careful measurement of policy change across a comprehensive scope of issues in a specific domain. Our results show different patterns depending on whether one looks at policy scope, prescriptiveness or specific policy settings. Careful measurement uncovered trends that previous scholarship has missed and which contradict the predictions of several dominant theories. It also reveals that apparent empirical debates in the literature can be attributed to research design choices. Some scholars chose a few key issues and found convergence. Others looked broadly and did not see it. We have looked both precisely and broadly and found both conclusions are correct. Activist-backed and industry-backed programs converged in policy scope on a few issues, but their scopes have seen little change. Furthermore, we found these programs to have diverged overall on prescriptiveness, because, while both standards "ratcheted up," they did so at different rates and on different policy issues. We hope that this deep dive into defining policy change in one domain not only enables scholarship on the

causes of public and private regulation in forestry but that it also offers a model for similar research in other policy domains.

This approach also has practical value. First, the power dynamics among groups that promote programs like the FSC or the SFI have created an environment in which competing claims about policy substance and how it has changed confuse buyers. The politics of private regulation revolve around "public comparisons that would resolve the debate about whose standards were higher" (Overdevest, 2010). We offer concepts to clarify what "higher" standards may mean. Second, it is impossible to measure the impact of a set of regulatory requirements without disentangling their component parts. Thus, our textual analysis is a necessary first step for efforts to assess impacts on the ground.

Most importantly, our framework and analysis offer a model for careful measurement of policy change as a variable. It is tempting to take shortcuts by making broad generalizations or by selecting what is easy to measure or what others have highlighted. However, if private governance scholars are to build testable theories or collect the kinds of evidence needed to test existing theories, our study illustrates the benefits of defining policy change in ways that can be applied across programs and over time. Doing so will not only improve the quality of research and theory, it may also uncover entirely new puzzles and insights.

References

Abbott, K. W. and Snidal, D. (2009). The Governance Triangle: Regulatory Standards Institutions and The Shadow of the State. In Mattli, W. and Woods, N., editors, *The Politics of Global Regulation*. Princeton University Press, Princeton, N.J.

Abbott, K. W. and Snidal, D. (2010). International regulation without international government: Improving IO performance through orchestration. *The Review of International Organizations*, 5(3):315–344.

Abderrazak, C. and Youssef, A. B. (2009). Multiplicity of eco-labels, competition, and the environment. *Journal of Agricultural and Food Industrial Organization*, 7(2):1–22.

Atkinson, L. and Rosenthal, S. (2014). Signaling the Green Sell: The Influence of Eco-Label Source, Argument Specificity, and Product Involvement on Consumer Trust. *Journal of Advertising*, 43(1):33–45.

Auld, G. (2014). *Constructing private governance: The rise and evolution of forest, coffee, and fisheries certification.* Yale University Press, New Haven.

Bartley, T. (2003). Certifying Forests and Factories: States, Social Movements, and the Rise of Private Regulation in the Apparel and Forest Products Fields. *Politics & Society*, 31(3):433–464.

Bartley, T. (2007). Institutional Emergence in an Era of Globalization: The Rise of Transnational Private Regulation of Labor and Environmental Conditions. *American Journal of Sociology*, 113(2):297–351.

Baumgartner, F. R. (2002). Positive and negative feedback in politics. In Baumgartner, F. R. and Jones, B. D., editors, *Policy Dynamics*. University of Chicago Press, Chicago.

Berger, S. and Dore, R. (1996). *National diversity and global capitalism.* Cornell University Press.

Bodansky, D. (1999). The Legitimacy of International Governance: A Coming Challenge for International Environmental Law? *The American Journal of International Law*, 93(3):596.

Botzem, S. and Dobusch, L. (2012). Standardization Cycles: A Process Perspective on the Formation and Diffusion of Transnational Standards. *Organization Studies*, 33(5-6):737–762.

Bozzi, L., Cashore, B., Levin, K., and McDermott, C. (2012). The Role of Private Voluntary Climate Programs Affecting Forests: Assessing their Direct and Intersecting Effects. In Ronit, K., editor, *Business and Climate Policy: The Potentials and Pitfalls of Private Voluntary Programs*. United Nations University Press.

Brunel, C. and Levinson, A. (2016). Measuring the Stringency of Environmental Regulations. *Review of Environmental Economics and Policy*, 10(1):47–67.

Cashore, B. (1997). *Governing Forestry: Environmental Group Influence in British Columbia and the US Pacific Northwest.* PhD thesis, University of Toronto.

Cashore, B. (2002). Legitimacy and the Privatization of Environmental Governance: How Non-State Market-Driven (NSMD) Governance Systems Gain Rule-Making Authority. *Governance*, 15(4):503–529.

Cashore, B., Auld, G., and Newsom, D. (2004). *Governing Through Markets: Forest Certification and the Emergence of Non-state Authority*. Yale University Press.

Cashore, B. and Howlett, M. (2007). Punctuating Which Equilibrium? Understanding Thermostatic Policy Dynamics In Pacific Northwest Forestry. *American Journal of Political Science*, 51(3):532–551.

Darnall, N., Potoski, M., and Prakash, A. (2010). Sponsorship Matters: Assessing Business Participation in Government- and Industry-Sponsored Voluntary Environmental Programs. *Journal of Public Administration Research and Theory*, 20(2):283–307.

DeLeon, P. and Rivera, J. E. (2009). *Voluntary Environmental Programs : a Policy Perspective.* Lexington Books, Lanham.

Delmas, M. and Montiel, I. (2008). The Diffusion of Voluntary International Management Standards: Responsible Care, ISO 9000, and ISO 14001 in the Chemical Industry. *Policy Studies Journal*, 36(1):65–93.

Djelic, M.-L. and Etchanchu, H. (2017). Contextualizing Corporate Political Responsibilities: Neoliberal CSR in Historical Perspective. *Journal of Business Ethics*, 142(4):641–661.

Eberlein, B., Abbott, K. W., Black, J., Meidinger, E., and Wood, S. (2014). Transnational business governance interactions: Conceptualization and framework for analysis. *Regulation & Governance*, 8(1):1–21.

Fischer, C. and Lyon, T. P. (2014). Competing Environmental Labels. *Journal of Economics & Management Strategy*, 23(3):692–716.

Fransen, L. (2011). Why Do Private Governance Organizations Not Converge? A Political-Institutional Analysis of Transnational Labor Standards Regulation. *Governance*, 24(2):359–387.

Grabosky, P. (2013). Beyond Responsive Regulation: The expanding role of non-state actors in the regulatory process. *Regulation & Governance*, 7(1):114–123.

Grabs, J., Auld, G., and Cashore, B. (2018). Switching between lenses for a better view: A meta-assessment of private regulatory governance and the role of public policy.

Green, J. F. (2013). *Rethinking private authority: Agents and entrepreneurs in global environmental governance.* Princeton University Press, Princeton, NJ.

Green, J. F. (2017). Blurred Lines: Public-Private Interactions in Carbon Regulations. *International Interactions*, 43(1):103–128.

Green-Pedersen, C. (2007). The Dependent Variable Problem within the Study of Welfare State Retrenchment: Defining the Problem and Looking for Solutions. *Journal of Comparative Policy Analysis: Research and Practice*, 6(1):3–14.

Gulbrandsen, L. H. (2004). Overlapping Public and Private Governance: Can Forest Certification Fill the Gaps in the Global Forest Regime? *Global Environmental Politics*, 4(2):75–99.

Gulbrandsen, L. H. (2014). Dynamic governance interactions: Evolutionary effects of state responses to non-state certification programs. *Regulation & Governance*, 8(1):74–92.

Hall, P. A. (1993). Policy Paradigms, Social Learning, and the State: The Case of Economic Policymaking in Britain. *Comparative Politics*, 25(3):275.

Hassel, A. (2008). The Evolution of a Global Labor Governance Regime. *Governance*, 21(2):231–251.

Heyes, A. and Martin, S. (2017). Social Labeling by Competing NGOs: A Model with Multiple Issues and Entry. *Management Science*, 63(6):1800–1813.

Howard-Grenville, J., Nash, J., and Coglianese, C. (2008). Constructing the License to Operate: Internal Factors and Their Influence on Corporate Environmental Decisions. *Law & Policy*, 30(1):73–107.

Howlett, M. and Cashore, B. (2007). Re-Visiting the New Orthodoxy of Policy Dynamics: The Dependent Variable and Re-Aggregation Problems in the Study of Policy Change. *Canadian Political Science Review*, 1(2):1–14.

Howlett, M. and Cashore, B. (2014). Conceptualizing Public Policy. *Comparative Policy Studies : Conceptual and Methodological Challenges*, pages 17–33.

Hsueh, L. and Prakash, A. (2012). Incentivizing self-regulation: Federal vs. state-level voluntary programs in US climate change policies. *Regulation & Governance*, 6(4):445–473.

Hudson, I. and Hudson, M. (2003). Removing the Veil?: Commodity Fetishism, Fair Trade, and the Environment. *Organization & Environment*, 16(4):413–430.

Judge-Lord, D. (2013). *Mechanisms of Policy Feedback: Interactions among Regulations and Public Investments in US Farm and Forest Politics.* PhD thesis, Yale University.

Kollman, K. and Prakash, A. (2001). Green by Choice? Cross-National Variations in Firms' Responses to EMS-Based Environmental Regimes. *World Politics*, 53:399–430.

LeBaron, G. and Burgoon, B. (2018). Explaining Patterns of Labor Protection and Exploitation in the Global Tea Supply Chain: Evidence from India.

Li, Y. and van 't Veld, K. (2015). Green, greener, greenest: Eco-label gradation and competition. *Journal of Environmental Economics and Management*, 72(C):164–176.

Loconto, A. and Fouilleux, E. (2014). Politics of private regulation: ISEAL and the shaping of transnational sustainability governance. *Regulation & Governance*, 8(2):166–185.

Lyon, T. P. and Maxwell, J. W. (2008). Corporate Social Responsibility and the Environment: A Theoretical Perspective. *Review of Environmental Economics and Policy*, 2(2):240–260.

Maxwell, J., Lyon, T., and Hackett, S. (2000). Self-Regulation and Social Welfare: The Political Economy of Corporate Environmentalism. *The Journal of Law and Economics*, 43(2):583–618.

McDermott, C. (2012). Trust, Legitimacy and Power in Forest Certification: A case study of the FSC in British Columbia. *Geoforum*, 43:643–644.

McDermott, C., Cashore, B., and Kanowski, P. (2010). *Global Environmental Forest Policies: An International Comparison*. Earthscan, London.

McDermott, C., Noah, E., and Cashore, B. (2008). Differences That "Matter"? Identifying Analytical Challenges in the Comparison of Forest Certification Standards. *Journal of Environmental Policy and Planning*, 10(1):47 – 70.

Meidinger, E. (2003). Forest Certification as Environmental Law Making by Global Civil Society. In Meidinger, E., Elliot, C., and G, O., editors, *Social and Political Dimensions of Forest Certification*, pages 293–329. Forstbuch: Remegen-Oberwinter.

Meidinger, E. (2006). The Administrative Law of Global Private-Public Regulation: the Case of Forestry. *The European Journal of International Law*, 17(1):47–87.

Mills, R. W. (2016). The interaction of private and public regulatory governance: The case of association-led voluntary aviation safety programs. *Policy and Society*, 35(1):43–55.

Newsom, D. and Hewitt, D. (2005). The Global Impacts of SmartWood Certification. Technical report, Rainforest Alliance.

Overdevest, C. (2005). Treadmill Politics, Information Politics, and Public Policy. *Organization & Environment*, 18(1):72–90.

Overdevest, C. (2010). Comparing forest certification schemes: the case of ratcheting standards in the forest sector. *Socio-Economic Review*, 8(1):47–76.

Overdevest, C. and Zeitlin, J. (2014). Assembling an experimentalist regime: Transnational governance interactions in the forest sector. *Regulation & Governance*, 8(1):22–48.

Poret, S. (2016). Label Battles: Competition among NGOs as Standard Setters.

Prado, A. M. (2013). Competition Among Self-Regulatory Institutions. *Business & Society*, 52(4):686–707.

Prakash, A. (2000). *Greening of the Firm: The Politics of Corporate Environmentalism*. Cambridge University Press, Cambridge.

Rodrik, D., Subramanian, A., and Trebbi, F. (2004). Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development. *Journal of Economic Growth*, 9(2):131–165.

Smith, T. M. and Fischlein, M. (2010). Rival private governance networks: Competing to define the rules of sustainability performance. *Global Environmental Change*, 20(3):511–522.

Tigar, J. S. (2017). Resolute Forest Products, Inc. et al. v. Greenpeace International et al.

van der Ven, H. (2015). Correlates of rigorous and credible transnational governance: A cross-sectoral analysis of best practice compliance in eco-labeling. *Regulation & Governance*, 9(3):276–293.

van der Ven, H., Rothacker, C., and Cashore, B. (2018). More Eco-Labels but Fewer Forests: Lessons from Non-State Market Driven Governance in the Soy, Palm Oil, and Cocoa Sectors.

Vince, J. and Haward, M. (2017). Hybrid governance of aquaculture: Opportunities and challenges. *Journal of Environmental Management*, 201:138–144.

Vogel, D. (1995). *Trading Up: Consumer and environmental regulation in a global economy.* Harvard University Press, Cambridge, MA.

Weimer, D. L. (2006). The Puzzle of Private Rulemaking: Expertise, Flexibility, and Blame Avoidance in U.S. Regulation. *Public Administration Review*, 66(4):569–582.

Weimer, D. L. and Vining, A. R. (2005). *Policy Analysis: Concepts and Practice.* Routledge, New York, 6 edition.