When do institutions “bite”? Historical institutionalism and the politics of institutional change.

Giovanni Capoccia
University of Oxford
giovanni.capoccia@politics.ox.ac.uk

Abstract

Historical institutionalist theories of endogenous change have enhanced our understanding of institutional development by providing a theoretical vocabulary for analyzing how institutions may be renegotiated over the long run by social and political actors. In these theories, however, the causal impact of institutions themselves on political outcomes, including their own change and reform, is less developed—a significant problem for an institutional research program. This paper addresses this problem by proposing strategies that integrate historical institutionalism’s insights into endogenous institutional change with a systematic analysis of the institutional conditions under which “bottom-up” processes of gradual change are likely to be counteracted. In particular, the institutionalization of cultural categories and the allocation of power over the timing of reform within institutional and policy configurations are important variables for understanding how pre-existing institutions may enable institutional incumbents to channel, delay, or prevent institutional change.

Forthcoming, Comparative Political Studies
Introduction

Over the past quarter-century, historical institutionalism (hereinafter HI) has established itself as an influential research program in comparative politics (e.g. Thelen & Steinmo, 1992; Hall & Taylor, 1996; Thelen, 1999; Pierson & Skocpol, 2002; Pierson, 2004; Streeck & Thelen, 2005a; Mahoney & Thelen, 2010a; A. Campbell, 2012; Fioretos, Falleti & Sheingate, forthcoming). After a period in which the focus was on the question of how political and social behavior is structured by institutions and public policies, the emphasis of recent work in HI has shifted to the analysis of institutional change. In particular, theories of gradual endogenous institutional change have been highly influential and represent one of the most important theoretical frontiers in HI research (Hacker, 2004; 2005; Thelen, 2004; Streeck & Thelen, 2005b; Mahoney & Thelen, 2010b; Hacker, Pierson & Thelen, 2015; Conran & Thelen, forthcoming). This literature has provided two important correctives to accounts of institutional development based on path dependency and metaphors of punctuated equilibrium. First, it has theorized how endogenous institutional change might take place, correcting views that relegated significant change to exogenous shocks (e.g. Krasner, 1984). Second, it has emphasized the role of agency in institutional development, thus correcting abstract views of institutional persistence that relied on images of “inertia” and “stickiness”. At the same time, however, theories of gradual institutional change have focused more on how social and political interactions
transform institutions than on how institutions themselves structure those interactions. This theoretical emphasis may result in a conception of institutions as overly plastic (e.g. Pierson, 2006: 116; Weyland, 2008: 284; Hall, 2010b), one that ultimately under-theorizes their causal role in politics, and in particular their role in the processes by which they themselves change or persist.

At this stage of theoretical development in the HI research program, therefore, a central task confronting scholars is to understand the conditions under which institutions and policies\(^1\) structure social behavior or become themselves “the object of strategic action” (Hall, 2010a: 204). The overarching theoretical goal is to identify systematically the ways in which an institutional configuration at a certain point in time \((t_i)\) influences the interaction of social and political actors so that institutional change, of varying scope and intensity, or institutional stability is achieved at a later point in time \((t_z)\). This paper argues that significant advances towards this goal can be made by integrating theories of endogenous change with a sharper theoretical focus on two variables: the institutionalization of cultural categories and the allocation of power over the timing of institutional reform within an institutional configuration. As I discuss below, the focus on these two variables is not arbitrary, but builds directly on two key micro-foundational processes that are posited, but whose implications are not fully articulated, in theories of endogenous institutional change: widespread non-
compliance with formal rules “on the ground”, which can trigger incremental processes of defection and reinterpretation that ultimately change how an institution operates; and shifts in the social coalitions that underpin institutions leading to the renegotiation of institutional arrangements. On the one hand, a strong institutionalization of cultural categories can sanction negatively defection and reinterpretation of formal rules, thus containing the institutional transformation triggered by such processes. On the other hand, when institutional incumbents control the timing of the institutional reform agenda, they can resist bottom-up pressure for change by delaying action until the saliency of reform among the target population has waned, thus making it harder for political entrepreneurs to assemble and sustain a coalition for change.

The purpose of this paper is to advance the theoretical conversation on how institutions develop over time by building on theories of endogenous institutional change. These theories rest on multiple and analytically distinct micro-foundational processes, and do not claim to offer an exhaustive account of all possible ways in which institutions can change gradually. Similar to these theories, therefore, the paper does not propose an all-encompassing framework either: institutionalization of cultural categories and control over the timing of reform are not the only ways in which institutions can acquire stability and causal force. Future theorization on endogenous institutional change is likely to identify micro-foundational processes
besides rule defection, reinterpretation and coalition shifts. With respect to these future advances on the causes of endogenous change, this paper offers a strategy for theorizing the mirror outcome of stability, following the general principle that theorizing stability by building on theories of change is generally easier than the reverse (Sewell, 2005: 126; see also Abbott, 2001: 24, 240-260): identify institutional characteristics that empower incumbents to counter the micro-foundational, bottom-up processes of change.

At the same time, the integrated theory of endogenous institutional development discussed in the paper also complements theories of institutional path dependency in two respects. First, an explicit theoretical focus on the variable capability of institutions to entrench cultural categories that constrain institutional change addresses an aspect that research on “interpretive feedback effects” (Pierson, 1993) has left mostly unexplored. Second, the approach suggested in this paper concurs with existing HI efforts to go beyond metaphors of stickiness and “institutional reproduction,” but, in contrast with the current emphasis on change, it theorizes important institutional conditions under which institutional incumbents are most likely to achieve stability or at least delay or channel change—thus complementing the prevalent focus on reformers in the literature (e.g. Gingrich, 2015). More generally, the discussion contributes, within the context of HI, to a broader
emerging research agenda in comparative-historical analysis on the pathways through which power can be self-reinforcing (Pierson, 2015).

The paper articulates these theoretical issues and proposes strategies for making them analytically tractable in HI empirical work. It is organized as follows: The next section clarifies the stakes of this discussion, namely the theorization and empirical assessment of the causal force of institutions. The following section briefly teases out the two micro-foundational processes—non-compliance with rules and shifts in the social coalitions underpinning an institutional arrangement—that are central to the logic of endogenous institutional change. The subsequent two sections discuss and illustrate how these two bottom-up processes that drive endogenous gradual change can be countered by, respectively, strong institutionalization of cultural categories and concentration of power over the timing of institutional reform in the hands of institutional incumbents. The conclusion summarizes the discussion and articulates its implications for the HI research program.

The Problem: Endogenous institutional change and the causal force of institutions

The seminal works by Thelen (2004), Hacker (2004; 2005), Streeck and Thelen (2005b) and Mahoney and Thelen (2010b) propose a novel theoretical vocabulary to study endogenous, gradual, and transformative institutional change. The main foil of this
influential approach is path dependency models of institutional development, which emphasize that institutions are typically stable features of a political landscape.² In path dependency models, the stability of institutions is attributed to their influence on the resources and incentives of actors and to the development of institution-specific assets such as skills, privileges, knowledge of procedures, and networks with other actors (Pierson, 2004: 21, 30-35, 149). Against this view of institutions as equilibria, the agenda-setting contributions mentioned above raise significant theoretical and empirical objections. Theoretically, they consider the path dependency view of institutional development as incapable of incorporating institutional change, which in path-dependency theories is typically attributed to exogenous shocks (e.g. Mahoney & Thelen, 2010b: 4; see also Peters, Pierre, & King 2005; cf. Pierson, 2004: 52-53, 135, 157). Empirically, they argue that in many circumstances, gradual institutional change is transformative rather than adaptive and bounded, as is instead often implied by path-dependency accounts of institutional development (Streeck & Thelen, 2005b: 6; Mahoney & Thelen, 2010b: 2; see also Thelen, 2004; cf. Pierson, 2000: 265).

Institutions, according to this scholarship, are “distributional instruments laden with power implications” (Streeck & Thelen, 2005b: 9). As such, they are constantly beset by contestation from below, and for this reason “there is nothing automatic, self-perpetuating, or self-reinforcing about institutional arrangements.” Instead, “...a

² I refer to path dependency models in political science based on conceptions of increasing returns (Pierson, 2000; 2004). For a broader discussion of path dependency, see e.g. Mahoney (2000).
dynamic component is built in” (Mahoney & Thelen, 2010b: 8). Therefore, rather than as equilibria, institutions are best conceived as *arenas of conflict*, as “regimes” in which both “rule-makers,” defined as the actors that set and modify, “often in conflict and competition” the formal rules that constitute an institution, and “rule-takers,” the actors that are expected to comply with such rules, struggle to adapt the institution to their needs and agendas (Streeck & Thelen, 2005b: 13). Institutional change often reflects underlying processes of imperfect compliance, rule reinterpretation, and coalition-building among social and political actors. Thus, without excluding the possibility of abrupt institutional breakdown and replacement due to an exogenous shock, scholars in this tradition of analysis maintain that institutional change will more typically be endogenous, gradual, and transformative. They identify several types of such change: layering indicates a process by which new formal rules are added onto old ones to change their effects (Schickler, 2001); conversion refers to the reinterpretation of existing rules to serve new purposes (Thelen, 2004); drift points to situations in which old rules are intentionally not adapted to changing social conditions, with the purpose of gradually changing their effects (Hacker, 2004); exhaustion is the process “by which an institution ‘withers away’ through … overextension diminishing its capacity to do what it was originally invented to do” (Streeck, 2009: 125).
It is difficult to overestimate the impact that theories of endogenous institutional change have had on HI scholarship, in particular on analyses of the development of the institutions of modern capitalism and welfare states (e.g. Béland & Waddan, 2012; Ornston, 2012; Hemerijck, 2013; Thelen, 2014). Furthermore, the insights provided by these theories are increasingly being applied to domestic political institutions and policies (e.g. Sheingate, 2010; Broschek, 2013; King & Smith, 2014), to international law and regulations (e.g. Rixen, 2011; Alter, Helfer, & McAllister, 2013), as well as to interdependence between the two (Farrell & Newman, 2014). Overall, this literature makes a convincing case for the ubiquitousness of gradual, endogenous, and transformative institutional change. As some have recognized, however, the debate over whether significant (i.e. not merely adaptive) institutional change is more likely to be abrupt and exogenous, or gradual and endogenous, is somewhat beside the point. On the one hand, scholars in different traditions have provided ample empirical evidence of both types of institutional change. On the other hand, although it is true that institutions typically do not endure by standing still (Thelen, 2004: 217), the issue of whether gradual change has attained a “transformative” threshold or whether it instead remains “bounded” depends ultimately on the conceptualization of the inner “political logic” of the particular institution (e.g. Weir, 2006: 177; see also J. Campbell, 2004: 27; Streeck & Thelen, 2005b: 18). When, as is often the case, there is no theoretical agreement on this point, classifying institutional change as “bounded” or
“transformative” runs the risk of becoming simply a semantic issue of “whether the glass...is still half full or already half empty” (Streeck, 2009: 17; see also Peters, Pierre & King 2005: 1287; Thelen, 2014: 4).

The real issue in this debate is how the causal force of institutions should be theorized. The concern is that by theorizing, as theories of endogenous institutional change do, institutional development as driven fundamentally by underlying social processes, institutions might lose their independent causal force over political action—including political action aimed at changing the institutions themselves (e.g. Pierson, 2006: 116; Weyland, 2008: 284; Hall, 2010b). This is a serious matter for an institutionalist research program, which, to avoid casting institutions as an epiphenomenal intermediary between the strategies of powerful actors and broader political outcomes, requires a robust theorization of the role of institutions in causing political outcomes, including their own development. Indeed, institutional analysis (of any theoretical flavor) must strike a middle ground between characterizing institutions as absolutely constraining or as malleable tools in the hands of powerful actors. Depending on the circumstances, institutions can possess one or the other characteristic: it is undeniable that institutions are often shaped and redesigned by political actors but it is also undeniable that institutions may constrain both “rule-makers” and “rule-takers”. As has been correctly remarked, institutional analysis should attempt to theorize this “partial bite” of institutions, by specifying the
conditions under which institutions can be malleable or constraining (Lohmann, 2003: 97; see also Hall, 2010a: 204).³

The micro-foundations of endogenous institutional change

The current state of HI theory is characterized by a bifurcation that does not satisfactorily address the problem of the “partial bite” of institutions. On the one hand, path dependency theories, by relegating change to exogenous factors, are unable to theorize when institutions will be durable and when they will be plastic. On the other hand, theories of endogenous institutional change offer only an incomplete theorization of the conditions under which institutions “bite.” On the one hand, institutional stability and continuity is generally theorized in conjunction with processes of drift or conversion, by which institutions (or their effects) change endogenously under the surface of stable formal rules. Strictly speaking, an outcome in which formal institutions influence strategies, incentives and values of actors consistently over time, shaping political outcomes and favoring their own stability remains untheorizable in this framework. On the other hand, this work typically associates the emergence of different types of institutional change with two variables

³ An insightful line of scholarship focuses on informal institutions, defined as socially shared, unwritten rules that give rise to regularities above and beyond those prescribed by formal rules, which therefore have little or no “bite” (Helmke & Levitsky, 2004: 727). This literature captures much of what happens in the developing world. Although there have been some attempts to apply theories of endogenous institutional change to weakly enforced formal institutions (e.g. Onoma 2010), the theories in question typically focus on developed democracies and assume that power-holders have sufficient capacity to enforce formal rules (see, most explicitly, Pierson, Hacker & Thelen 2015: 183, fn.2; see also Levitsky and Murillo 2009: 127-128).
that capture only incompletely what drives gradual endogenous change: the (high or low) amount of discretion that the design of formal rules vests in the hands of decision-makers, and the (high or low) number of “veto capabilities” in the institutional arrangement and the surrounding political environment (e.g. Hacker, 2005:48; Mahoney & Thelen, 2010b: 3, 19). Even though labels and definitions vary slightly, essentially these variables shape the preferences of elites to either “work within” or “work outside” an existing institutional framework to achieve their goals (Hacker, 2004: 246-248).

These typologies capture an important way in which existing institutions can shape their own change by generating different sets of incentives for decision-makers, but are based on the tacit assumption that institutional change is the outcome of deliberate political strategies of “policy subversion” on the part of powerful elites (Culpepper, 2011: 194). In addition to such elite-based processes, however, theories of endogenous institutional change are premised on two, not fully theorized, micro-foundational processes that drive endogenous change “from the bottom up”: noncompliance with formal rules (in the form of reinterpretation or outright defection) and shifts in the social coalitions underlying an institutional arrangement. The first of the two is driven by the existence of gaps between the letter and the actual meaning of formal rules. As Streeck and Thelen argue, due to factors including the cognitive limits of “rule-makers” and their limited capability to control rule implementation, as well as
inconsistencies in rules, unforeseen changes in external conditions, or open contestation on the part of “rule-takers”, such gaps will always be present. By fueling political contestation over the form, functions, and salience of specific institutions, the presence of gaps between what a rule says and how it is applied allows incremental processes of rule defection and reinterpretation that lead to gradual but transformative institutional change over the long run (Streeck & Thelen, 2005b: 19; see also Streeck, 2009: ch. 17; Mahoney & Thelen, 2010b: 10). The second process emphasizes how gradual institutional change can result from changing coalitions among social groups, often spearheaded by the “losers” of previous institutional battles (Thelen, 2004: 33, 295; Mahoney & Thelen, 2010b: 9; Thelen, 2014). Mahoney and Thelen, aiming to provide a general theory of endogenous institutional change that applies beyond the realms of political economy and social policy, trace the connections between actors, their preferences, the probable coalitional alignments deriving from such preferences, and the type of institutional change that is likely to result (Mahoney & Thelen, 2010b: 24-27; 29-31).

The theorization of these micro-foundational social processes of endogenous institutional change has identified ways in which institutions can be transformed and renegotiated “from the bottom up”. This framework, however, does not clarify the role that existing institutions may play in influencing the direction, intensity, and even the possibility of change in direct opposition to these social processes. How do existing
institutional configurations influence non-compliance “on the ground” and the formation of social coalitions in support of institutional change so as to empower supporters of the institutional status quo to channel, delay, or prevent undesired change? How should variation in this capability of pre-existing institutions be conceptualized and measured? How can this conceptualization be systematically incorporated into analyses of endogenous institutional development? The remainder of the paper addresses these questions, identifying and theorizing two important dimensions of institutions that can counter bottom-up processes of change, and illustrating the theory with examples.

**Countering non-compliance: Institutionalization of cultural categories**

In their work on liberalization in contemporary capitalist systems, Streeck and Thelen argue that bottom-up rule reinterpretation and non-compliance can be driven by the material interests of the actors affected by an institutional configuration (Streeck & Thelen, 2005b: 33; see also Streeck, 2009). Ideational factors, however, would seem to play at least as important a role. To be sure, it may be difficult to disentangle the role of interests from that of norms, values and identities in processes of institutional change (e.g. Hall, 2005: 133; 2010a: 211-212; Weir, 2006: 175). But ideational factors remain analytically distinguishable from material interests, and, even in Streeck and Thelen’s framework, their independent causal role is visible. For example, they
maintain that third-party rule enforcement is an important *definitional* trait of institutions. Third parties intervene “not necessarily because they identify with the interests” of any of the parties in a dispute, but as “an expression of moral disapproval” to protect actors whose “legitimate, normative expectations have been disappointed” (Streeck & Thelen, 2005b: 10-11). However, moral disapproval and the legitimacy of third-party sanctions can vary substantially across different institutional contexts depending on the level of social entrenchment of the norms and values supporting a certain institutional configuration (e.g. Greif, 2006: 30). Other examples are possible (Streeck & Thelen, 2005b: 14; Mahoney & Thelen, 2010b: 13), but more generally, the conditions facilitating endogenous change via bottom-up defection and reinterpretation are best conceived as *variable* across institutional contexts rather than as a characteristic of institutions as such. Analyses of liberalization in capitalist economies that use this theoretical approach (Thelen, 2004; Streeck & Thelen, 2005a; Streeck, 2009) may have identified an institutional configuration and a historical period in which the normative underpinnings of institutions were vulnerable to defection and reinterpretation. These circumstances may certainly not be unique. However, in other policy sectors—and other historical contexts— the norms supporting an institutional configuration may be more strongly entrenched and therefore less open to the processes of reinterpretation and non-compliance that lead to gradual institutional change. By incorporating this variation in theories of institutional
development, we avoid the risks connected with generalizing based on a specific episode and context, however important these may be.

Cultural categories and institutional stability

One of the causes of the “degree of embeddedness” (Blyth, 2001: 25) of the social norms supporting an institutional and policy configuration are the institutions themselves. To be sure, the effect of institutions on their ideational underpinnings is likely to be partial. Perhaps the main reason that compliance with institutions and rules is never perfect is that the target population is always exposed to social and ideational forces that are broader than any single institution or set of connected institutions. At the same time, the literature on feedback effects has convincingly argued that “clusters” of connected policies (defined as a set of policies and institutions that have the same, or largely overlapping, target populations;4 Pierson, 2006: 121), far from being simply the reflection of underlying social structures, can influence the resources and incentives of collective and individual actors (e.g. Patashnik, 2008), and can have “interpretive” feedback effects, providing the mass public with general interpretive

---

4 In the literature on U.S. public policy, the concept of “target population” refers to those persons and groups whose behavior a public policy seeks to change in a certain direction through coercion or inducements (e.g. Ingram and Schneider 1991: 334). The specification of coercion-based and inducement-based policy instruments (e.g. Peters 2000: 39) is less central here, so I use the term interchangeably with Streeck and Thelen’s “rule takers”. In the context of the discussion of “cluster effects”, I refer to “target populations” because the term makes clear that different public policies and institutions can affect the same or broadly overlapping social groups.
schemes through which to see the world (Pierson, 1993; 2004; Béland, 2010). Empirical analyses of interpretive feedback effects have focused on levels of participation, feelings of political efficacy, and levels of interpersonal trust among different social groups (for reviews, see A. Campbell, 2012: 336-338; Hacker & Pierson, 2014: 664-666). The scholarship on ideas and institutions (e.g. Blyth, 1997) usefully complements these analyses by showing how norms and identities can become institutionalized in organizations (e.g. Berman, 2001: 238), routine practices (e.g. J. Campbell, 2004: 113) and policies (e.g. Blyth, 2001: 13, 25), outliving the conditions of their initial rise to prominence and exerting independent causal effects on norms, values and identities of the actors involved (Béland & Cox, 2011: 9-10).

In the case of clusters of institutions and policies, this effect of institutions on their social underlay is best conceptualized in terms of institutionalization of cultural categories (Steensland, 2006). The insight that institutions categorize and thereby

---

5 An influential literature that aims to build theoretical bridges between rational choice institutionalism and HI focuses on how institutional arrangements affect the preferences of actors (Katznelson & Weingast, 2005; see also Levi, 2005; Ahlquist & Levi, 2013). Although the terminology of “preferences” serves well the theoretical goal in question, it also has costs. In particular, it blurs distinctions between the various components of preferences, distinctions which have traditionally been emphasized in institutionalist scholarship. These include the difference between normative orientations and cognitive structures (e.g. J. Campbell, 2004) and the difference between internalized beliefs (which can be normative or cognitive, and constitute identities), instrumental beliefs (beliefs about cause-effect, e.g. Hall, 1989; 2010a: 207-208), and behavioral beliefs (expectations of what other actors involved in interactions will do, e.g. Greif, 2006: 31).

6 I use Steensland’s expression “cultural categories” as a more precise label than “ideas”. As discussed below, cultural categories are at the root of different types of ideas (e.g. J. Campbell, 2004: 90-107). They influence both the cognitive and normative sphere and have an impact on the framing strategies of elites as well as on the more diffuse orientations of the public and elites (Steensland, 2006: 1281). In sociology,
influence their social constituency—and that categorization is the result of a policy choice rather than a mechanistic translation of pre-existing characteristics of an institution’s social environment—has a long pedigree in the work of historians and social scientists (e.g. Wahrman, 1995; Scott, 1998). As mentioned, the task at hand is to identify variation: namely, which characteristics of an institutional cluster, by being more or less present, influence the degree of entrenchment of cultural categories among the target population, thus making institutions more or less vulnerable to endogenous change via defection and reinterpretation. To this aim, consider the following passage from Rogers Smith’s account of the institutionalization of racial identities in the US:

“...American racial identities have actually been created by relatively autonomous governmental institutions that labeled some as ‘white’, some as ‘black’, some as different races [...]. Those institutional arrangements—antebellum state civil and criminal laws, late-nineteenth- and early-twentieth-century state Jim Crow laws, federal census rules, immigration and naturalization statutes, judicial rulings on all of the above, police enforcement of all of the above [...] have been more the direct source of racial constructions than socially widespread racial ideas” (Smith, 2006: 94).

Even though Smith’s main contention is that different “racial political-institutional orders” have competed over the course of American political development,⁷ his

---

⁷ See also King and Smith (2005). King and Smith (2014) trace the incremental reformulation and instrumentalization of color-blind ideas by a powerful conservative coalition, which provided the coalition in question with an important communicative advantage over the liberal advocates of race-
description of the historical construction of racial identities captures a real set of processes by which a cluster of connected institutions and policies can influence ideational orientations among rule-takers. Implicit in the passage above is that the *degree of consistency across connected institutional and policy realms* in the use of cultural categories expressed in formal rules as well as the judicial and bureaucratic interpretation of the rules correlates positively with the degree to which such categories are entrenched in the target population.

What mechanisms underlie this correlation? Besides scholars of race and ethnicity in American politics (such as Smith) and comparative politics (e.g. Lieberman and Singh 2012a), scholars of gender have also shown how the consistent and protracted use of cultural categories in policy and institutional clusters is likely to amplify the importance of one facet of identity (which is always multidimensional; e.g. Hall, 2005: 133; 2010a: 211-212) at the expense of other facets, shaping the way individuals perceive themselves and others (Starr, 1992: 278). Jenson, for example, highlights these dynamics in her comparison of early policies on women’s working hours, maternal leave and infant protection in France and the US. French policies were based on a conception of women as workers as well as childbearers, while analogous policies in the US were designed around an idea of women as active in the household conscious redistributive policies. They attribute the lack of prominence of race in the American political debate to this framing strategy, notwithstanding the persistence of profound racial divisions in American society,
rather than in the labor market. These policies institutionalized different conception of women’s social and economic status and were highly consequential for subsequent debates on women’s suffrage and equal rights (Jenson, 1989).

These scholars operationalize the institutionalization of cultural categories by first determining the relevant institutional and policy “cluster”, then identifying the relevant categories used in the design of formal rules, and finally assessing how consistently those categories are used across the policy realms constituting the cluster. This strategy has been developed perhaps most thoroughly by Lieberman and Singh (2012b: 268-274). They establish criteria for what constitutes use of ethnic categories (religious, tribal, racial, and linguistic) in each of nine connected policy and institutional realms relevant to the ethnic-based classification of the population, including the design of censuses and ID cards, rules on marriage, employment, and education, and the allocation of political autonomy or special privileges. Each realm is then scored dichotomously. This strategy allows them to build an additive index of “institutionalized ethnicity” (ranging from 0 to 9) that measures the degree of consistency in the use of the same categories across connected policy realms.

The dynamics documented by these scholars apply beyond the realms of ethnicity and gender politics, and potentially counter endogenous institutional change in two ways. First, the more consistent and protracted the use of cultural categories in
policy clusters, the harder it will be to introduce institutional reforms that challenge those categories. This effect, which operates on the framing strategies available to reformers (e.g. Schön & Rein, 1984), is indirectly demonstrated by the sociological literature on cultural resonance and policy debates (for a review, see Benford & Snow, 2000). This literature shows that framing strategies that “resonate” with established social values often bring popular support and elite allies to political entrepreneurs seeking policy reform. By contrast, opponents of entrenched values are typically at a rhetorical disadvantage. They cannot ignore socially entrenched values because doing so will make their proposals difficult to communicate to the relevant constituencies and will render them vulnerable to being labelled “radical” by their political adversaries (and therefore easily defeasible). To succeed, they have to spend substantial amounts of time and resources to do “definitional work” in order to recast the meaning of these categories (Steensland, 2008: 12; see also Pedriana & Stryker, 1997: 679).  

8 Steensland, for example, shows how the strong institutionalization of the cultural categories of “deserving” and “undeserving” poor had a crucial impact on the framing strategies of reformers and was at the root of the failure to introduce a
guaranteed annual income (GAI) policy in the US in the 1960s and the 1970s (Steensland, 2006; 2008).9

Second, and even more importantly for countering rule reinterpretation and defection, the consistent use of cultural categories in an institutional and policy configuration is likely to generate powerful incentives for courts and bureaucracies to interpret rules in line with such categories—thus limiting bottom-up gradual change. Accounts of endogenous institutional change emphasize that, as actors endowed with significant discretion in interpreting and implementing formal rules, judges and bureaucrats can reinterpret rules to the point of transforming how these operate in practice (e.g. Hall & Thelen, 2009: 19; Béland & Waddan, 2012: 29-30).10 In this view, “losers” of prior legislative battles may focus their efforts on courts and bureaucracies to shift the interpretation and patterns of implementation of formal rules to their advantage (e.g. Thatcher & Stone Sweet, 2003: 17). This strategy may be particularly attractive in a situation in which powerful veto players make reforming formal rules difficult (Hacker, Pierson, & Thelen, 2015). However, when clusters of formal rules are supported by strongly institutionalized cultural categories, reputational incentives are likely to make courts and bureaucracies more inclined to resist pressures to radically alter how rules work “on the ground”. As a number of important lines of research

---

9 Blyth analyzes similar dynamics in the reform of the Swedish welfare state (e.g. Blyth, 2001:13; 2002: 202-222).

10 These are key themes of the literatures on the “judicialization of politics” and the “politics of delegation” respectively. For general discussions see Stone Sweet (2000) and Huber and Shipan (2000).
have shown, judges and bureaucrats are typically attentive to their individual reputations and that of their organizations, which they see as closely linked. Courts typically refer (implicitly or explicitly) to prevailing social norms to frame their interpretation of a formal rule (e.g. Lin, 1999; Kotler, 2000), and there is substantial empirical evidence that they take their broad social and political legitimacy very seriously (e.g. Gibson, Caldeira & Baird, 1998; Gibson, 2007). These judicial legitimacy concerns can act as a powerful factor in limiting the courts’ propensity for “evolutionary interpretations” of formal rules (e.g. Garrett, Kelemen & Schulz, 1998; Kelemen, 2001: 624). Regarding bureaucrats, a nascent but already empirically robust literature offers powerful evidence that in most situations bureaucratic decision-making can only be explained by reputational incentives and not by the desire to maximize wealth and power, contrary to what is commonly assumed in formal analyses of the “politics of delegation” (e.g. Carpenter, 2010: 57-58; Carpenter and Krause, 2012).

Moreover, the reputational advantages derived from respecting entrenched categories can favor the emergence of autonomous increasing-return processes that reinforce the incentives of judges and bureaucrats to “stay the course”. In courts, judicial precedent is likely to develop.11 By developing precedent, judges can reduce

11 The important function of precedent for judges is further demonstrated by the widespread reliance of courts on previous decisions in civil law systems, which, unlike common law systems, do not possess a formal system of precedent (e.g. Shapiro, 1981: 126).
their workload by deferring to prior rulings (Hathaway, 2001: 126), enhance their legitimacy by casting themselves as decision-makers bound by the “objective” current state of the law (Shapiro, 1981), and increase the overall saliency of judicial modes of reasoning, “packaging policy as doctrine” (Stone Sweet, 2002: 128). In general, bureaucrats also value the workload-reducing effects of routinization (e.g. Page, 2012). More importantly, bureaucrats are likely to perceive their broad reputation in society as a necessary condition for building power and prestige (both agency and personal), which in turn are essential to obtaining budgetary funds and recruiting high-quality personnel, which feeds back into increased status and autonomy (Carpenter, 2010: 34, 49, 54; see also Carpenter, 2001).

In sum, the central hypothesis that emerges from this discussion is that gradual endogenous institutional change via rule defection and reinterpretation is more likely when connected policies and institutions are based on contradictory or inconsistent cultural categories. By contrast, the consistent use of cultural categories in an institutional cluster is likely to entrench such categories among the relevant target populations, making substantial reform of the formal rules more difficult and countering non-compliance “on the ground”—a process that is likely to be self-reinforcing over time. Of course, even under these conditions, formal rules are not

---

12 The degree of consistency in the use of cultural categories differs from the degree of “discretion” available in interpreting formal rules used in the typologies of institutional change mentioned earlier (Hacker, 2005: 48; Mahoney & Thelen, 2010b: 3, 19) as well as from the related concept of “precision” in
deterministically immutable, nor will courts and bureaucracies interpret them identically across time and space. Pressure for change can always build up in society: as mentioned, “definitional work” by stakeholders can in principle transform entrenched cultural categories or the way in which these categories are conventionally understood in society at large, and can create space for institutional change either through formal reform, or via judicial or bureaucratic reinterpretation. Indeed, history is replete with examples of gradual but ultimately radical change of institutional configurations that were once based on strongly entrenched cultural categories repeated across several sets of rules and policies. Take suffrage rules for women or civil rights for ethnic minorities. Such definitional work, however, is likely to be costly and to take a substantial amount of resources and time, during which formal institutions and rules, as well as their interpretation and implementation, will be stable and influence the attitudes, expectations, and strategies of the actors involved.

Illustration: The stability of “militant democracy” in post-war Germany

The system of “militant democracy” established in the Federal Republic of Germany after WWII (e.g. Thiel, 2009) illustrates how a highly consistent use of cultural categories in several connected policy realms can shield an institutional configuration rule design (Hacker, Pierson & Thelen 2015: 189). The consistent use of cultural categories across a policy cluster does not drive different forms of change but affects the stability of the core elements of the cluster. Moreover, measuring discretion is difficult beyond cases in which a formal rule sets a precise quantitative standard (e.g. the level of the minimum wage), while the consistency in the use of cultural categories is in most cases amenable to reliable operationalization (Lieberman & Singh, 2012b: 268-274).
from transformational change via “bottom-up” defection and reinterpretation. Traditionally interpreted as a reaction to the failure of the Weimar Republic to resist the Nazi takeover, the legal system of the Federal Republic includes constitutional rules, parliamentary statutes, and early decisions of the Federal Constitutional Court, that have defined strict boundaries between the categories of legitimate and illegitimate political action which, in the latter case, can be targeted with repressive state action. Illegitimate political activities are those which threaten the “liberal-democratic fundamental order”, an expression included in several articles of the 1949 Basic Law and relevant statutes, and articulated further in a 1952 decision of the Federal Constitutional Court (Decisions of the Federal Constitutional Court, Vol. 2, pp. 12-13). Legislative activity and court decisions reaffirming these categories span several connected policy realms, including the rules governing the activities of political parties and associations; those defining the exercise of the constitutional rights to free expression and free assembly; and those establishing the requisites for employment of certain classes of public officials (e.g. Brinkmann, 1983; Jaschke, 1991). Since 1949, consistent trajectories of legislation and jurisprudence have reaffirmed the core categories, and implementation, although uneven, has not substantially changed the way rules operate on the ground (e.g. Thiel, 2009). Furthermore, the national and regional “Offices for the Protection of the Constitution”, tasked since 1950 with gathering information on extremist individuals and groups, have reported publicly on
“extremist” activities, thus further entrenching in the public the categories of legitimate and illegitimate political action (e.g. Jaschke, 1991).

Given the political sensitivity of these rules, there have been a number of occasions on which they have not been applied fully or on which their implementation has been accompanied by heated public controversy and contestation. These episodes, however, have not substantially affected how the rules have been interpreted over time. For example the Berufsverbot—the political vetting of civil servants—has been the subject of domestic and international controversy. Notwithstanding the polemic, which peaked during the 1970s, all the main German political parties took an essentially favorable view of this practice, and the Constitutional Court found it to be constitutional in a well-known 1975 ruling (Frisch, 1977: 13). The rules that allow the forced dissolution of extremist parties afford another example of consistent rule interpretation in the face of controversy. In 2001, the federal government and parliament requested the Constitutional Court to consider the forced dissolution of the extreme right-wing National-Democratic Party of Germany (NPD). The Court rejected the request on the grounds that the evidence for the case had been acquired through the objectionable police practice of infiltrating the party cadres with numerous government spies. Even in these circumstances, which induced several commentators to question the fundamental fairness of the procedures, a majority of

13 This was the fifth such request to the Court since 1951.
judges of the Court were in favor of proceeding with the trial (but they fell short of the supermajority necessary in such cases; Flemming, 2003). Furthermore, the dismissal of the case and the ensuing press scandal did not stop a renewed request to ban the NPD, which enjoyed broad political backing and which the Court is considering at the time of this writing.\textsuperscript{14}

To be sure, not even an explicit use of crisp cultural categories in several connected institutional realms, a strong judicial tradition clarifying these categories, and dependable bureaucratic implementation and monitoring can ensure full compliance. Streeck and Thelen are correct when they maintain that no rule can guarantee such an outcome. The case of German “militant democracy”, however, powerfully illustrates not only that strongly entrenched categories can sanction and therefore discourage non-compliance, but also that non-compliance does not necessarily go together with the erosion and ultimate transformation of institutions. In post-war Germany, a certain degree of non-compliance has certainly existed, as evidenced by recurring episodes of neo-Nazi public activity, and even, in some cases, political violence --but this has hardly led to transformational institutional change. To be sure, anti-extremist rules have been the object of recurring public debate between legal and political commentators in favor of the current order and those who criticize

such rules as contrary to liberal values of free expression and political association. Notwithstanding these debates, the normative boundaries defining legitimate and illegitimate political activity (especially right-wing extremism) have remained relatively stable over the past sixty-five years (Capoccia, forthcoming). And because of the highly institutionalized cultural categories that inform the rules of several connected policies and the consolidated judicial interpretation of those rules, rule implementation always remains a possibility, even though non-compliance might be tolerated, generally for reasons of political pragmatism, during certain periods.

**Influencing coalition dynamics: Agenda control and the timing of reform**

As mentioned above, endogenous institutional change can occur not only through rule defection and reinterpretation, but also through shifts in the social coalitions underlying an institutional arrangement (Mahoney and Thelen 2010a). For example, Thelen (2014) shows how shifting coalitions of employers’ and employees’ organizations in different economic sectors have led to distinct forms of liberalization in advanced industrial economies. As with processes of rule defection and reinterpretation, however, pre-existing institutions can, in principle, influence coalitional dynamics and therefore can operate as one factor that shapes—at times decisively— their own development. In particular, *institutional power* may influence the dynamics of social coalitions in favor or against institutional reform. Even though in
theories of endogenous institutional change power is a defining characteristic of institutions (Streeck & Thelen, 2005b: 9-12; Mahoney & Thelen, 2010b: 7-8), the impact of power asymmetries generated by institutions on their own development remains under-theorized in this tradition of analysis. To be sure, emphasizing the importance of shifts in social coalitions in driving endogenous institutional change captures the fact that rule makers, by virtue of formal institutional power alone, are generally not in a position to change those institutions by fiat, or to resist any pressure for change coming from rule-takers. However, a more complete account of the politics of institutional change should recognize that, under certain circumstances, the power asymmetries built into institutional arrangements influence the dynamics of social coalitions in favor or against institutional stability. Coalitions are generally not formed on a level playing field: ceteris paribus, rule makers, i.e. individuals who hold institutional power, are likely to enjoy significant advantages over rule-takers in shaping the coalitional field. Compared with rule-takers, they generally are endowed with greater information and resources, and with a more central position in organizational and cultural networks. Thus, institutional power asymmetries, even though challenged and contested, may still play a significant role in the formation of social coalitions aimed at promoting institutional change, as well as, importantly, resisting transformative change— a role that needs to be theorized more fully. In some circumstances, the social coalition sustaining an institutional arrangement may be
partially endogenous to that institutional arrangement. When institutional incumbents are able to use their power to prevent or disrupt the formation of alternative, pro-reform coalitions, they will limit, delay or stop transformative change. In other words, the playing field for the formation of social coalitions is often not level: power asymmetries can shape which coalitions emerge to support what proposals for institutional change—and sometimes might even determine whether reformist coalitions emerge in the first place.

Saliency of reform and agenda control

One way to theorize the variable capacity of institutions to shape coalitional dynamics and thus influence their own change is to consider the agenda-setting power that is sometimes conferred upon incumbents (Moe, 2006: 62; Pierson, 2015) —in particular the power to control the agenda of institutional reform. This issue has certainly not been neglected in HI analyses. For example, emphasizing that individual interests and identities are always multidimensional, Hall argues that the strategic use, in proposals for institutional reform, of “rhetorical frames” (Schön & Rein, 1984: 32) tailored to appeal to certain facets of political identity, and not others, can substantially influence the types of reformist coalitions that emerge and the likely direction and intensity of institutional change (Hall, 2005: 133; 2010a: 211-212). Hausermann builds on Riker’s heresthetics (Riker, 1986) to show how governments can exploit the multidimensional nature of pension policies —in which benefit levels, type of financing, and eligibility
criteria vary largely independently from each other— and take advantage of the fragmentation of social and political organizations to “package” reform proposals that draw the support of particular social coalitions (Hausermann, 2011). Through such strategies of “coalitional engineering”, rule-makers can steer institutional change by using their agenda-setting power to influence coalitional dynamics.

These HI scholars focus on the importance of agenda-setting power in directing institutional change. However, in some circumstances agenda control and coalitional engineering can be extended to delaying, containing or preventing undesired change altogether, thus achieving institutional stability. This outcome can be the result of strategies of agenda control that have been extensively studied in a long-standing literature on social choice (e.g. Plott & Levine, 1978, Ordeshook & Schwartz, 1987). The insights of this scholarly tradition can be usefully deployed to understand how institutional incumbents may preserve stability by keeping institutional reform off the agenda. Current theories of endogenous institutional change, however, emphasize the possibility that political entrepreneurs and “losers” of previous institutional battles may assemble coalitions exactly to put reform on the agenda. When this happens, incumbents that control the timing of decision-making on institutional reform are likely to be in a position to preserve the institutional status quo.

This claim rests on the expectation that it will be easier to defend the institutional status quo when the issue of reform has low saliency among rule-takers. In
such cases, political entrepreneurs seeking institutional change will typically face higher collective action costs to gather sufficient support to renegotiate the existing institutional configuration. This can occur when rules are highly technical, generating large information asymmetries between rule-makers and rule-takers (e.g. Culpepper, 2011); or when the highly complex nature of the institutional structure deters rule-takers from investing the time and resources necessary to actively oppose the status quo; or when more pressing issues absorb the attention of rule-takers (e.g. Kingdon, 1995: 184; Jones & Baumgartner, 2005: 259). Even so, institutional reform can become highly salient due to “triggering events” such as a crisis or an external shock (e.g. Baumgartner & Jones, 1993: 129-130; Jacobs & Weaver, forthcoming: 10). However, as a considerable literature on policy-making has demonstrated, the reaction of rule-takers to such events is generally short-lived (e.g. Kingdon, 1995: 198). Although this may be less true of specialized institutions or organizations with a small target population, generally speaking individuals have limited cognitive capabilities, time, and resources. Moreover, especially in advanced societies, individuals are typically members of the target population of several institutional, policy and organizational configurations, which makes it even less likely that they will support a specific project of institutional reform over a sustained period (Jones & Baumgartner, 2005).

To elaborate briefly, imagine a situation in which two collective actors, A (the rule-makers) and B (the rule-takers), interact over institutional reform. At $t_1$, B can
force A to put institutional reform on the agenda. However, it is difficult for B to sustain mobilization for reform over time, and A can decide when institutional reform will be enacted. In this situation, A can simply wait until $t_2$—when B’s mobilization is low again—to shelve reform. In this situation, even though B may provisionally set the agenda at $t_1$, A does so at $t_2$, which is decisive for the outcome. If, therefore, the institutional configuration confers upon power-holders control over the timing of the agenda of institutional reform, they will likely be in a strong position to resist “bottom-up” pressures for change via shifts in social coalitions. Through control over the timing of institutional reform, institutional incumbents can respond to crisis in the short term by “gesturing” towards reform through various rhetorical expedients (e.g. Hacker & Pierson, 2014: 651), but delay action until the saliency of reform among rule-takers is again low, at which point they can either adopt cosmetic changes that do not challenge the existing power configuration, or shelve reform altogether. In these conditions, even though a majority among the rule-takers may in principle be in favor of change, change may not in fact happen.\footnote{This strategy presents common aspects with processes of drift (Hacker, 2004; 2005). Controlling the timing of reform points to a way in which institutional incumbents can enjoy the advantages of “playing defense” (Hacker & Pierson, 2014) when strategic inaction may be insufficient.}

**Illustration: Resisting reform of corporate governance in Europe**

An illustration of how control over timing of decision-making on reform affects the scope and prospects of institutional change can be found in recent comparative
analyses of corporate governance reform (e.g. Büthe & Mattli, 2011: 60-98; Schnyder 2012). For instance, in his recent comparative analysis of rules on takeovers, mergers and executive remuneration, Culpepper asks why recent public debates on corporate governance reform in advanced economies invariably led to the institutional outcome preferred by interest organizations of managers of large firms, “often against substantial political opposition” (Culpepper, 2011: 3). This outcome —which in some cases has involved reform, while in others has been characterized by successful resistance to reform, i.e. institutional stability— is due to what Culpepper calls the structure of “quiet politics”, which arises when highly technical formal rules, such as those of corporate governance, are the potential object of institutional reform (Culpepper, 2011: 179-185). As explained earlier, even technical rules can sometimes become the object of broad-based social mobilization, but since industry organizations are typically in the institutional position to strategically delay intervention, they can often shape, delay, or shelve any subsequent reforms. Their institutional power over rule-making is due to the tendency of politicians to defer to the expertise of managers: on the one hand, legislators de facto delegate most formal rulemaking to ad hoc expert committees and working groups, which advise on or negotiate legislation, and on which managerial interests are dominant; on the other hand, legislators leave considerable policy-making space to self-regulation (Jones & Pollitt, 2004; Mattli &

---

16 Grzymala-Busse (2015: 337-338) identifies similar patterns in her study of how Churches influence policy in democracies.
Politicians have incentives to delegate because electors are typically uninterested in institutional reform in the field of corporate governance, due to its technical nature.\footnote{This dynamic is well documented in the specialized literatures. In corporate law, Bebchuk and collaborators have shown how the main limit on rent extraction on the part of company executives is “public outrage”, which is intermittent and often absent (e.g. Bebchuk, Fried & Walker 2002, Bebchuk & Fried 2003). In economics, studies have found that media attention—which again is typically intermittent—leads to better protection of minority stakeholders (e.g. Dyck & Zingales 2002). In general, even work on corporate governance reform that highlights the important causal role of political-electoral dynamics, acknowledges that such dynamics only matter when reform is salient in the public eye (e.g. Roe & Vatiero 2015: 21).}

Most of the time, the structure of decision-making outlined above keeps corporate governance reform off the broader political agenda, thus avoiding the attention of rule-takers—such as, for example, workers with pension income invested in companies, and, for broader issues such as executive pay, the general electorate. During more exceptional moments, when reform captures the attention of rule-takers (due to a crisis, a scandal, or to initiatives of political entrepreneurs) this institutional framework often gives managers the ability to control the timing of corporate governance reform and thus substantially limit the subsequent reform outcome. In other words, the creation of a private interest committee allows industry actors “to appear to relent to calls for greater regulation without transferring such regulation to an unpredictable forum like a legislature”, while allowing actors represented on the committee to time deliberations in such a way to reach a decision “at a moment when
the temporary rise in public salience has dissipated” (Culpepper, 2011: 9, italics added).

The literature offers several examples of this dynamic. In France, on several occasions during the 2000s, politicians responded to financial scandals and increased salience among the population of the issue of executive pay by calling for tighter regulation. The government, however, repeatedly delegated the issue to industry committees and self-regulation. This enabled the sector to weather each scandal by delaying action and thus avoid any substantial and obligatory restrictions on executive pay (Gomez, 2010; Culpepper, 2011: 167-174).19 Practices of consultation and negotiation can also have similar effects. In the Netherlands in the mid-1990s, at a time when the issue of transparency in corporate governance and minority shareholder rights had unusually high salience in the press, the government responded by calling for less protection for large shareholders. Negotiations on reform between the government and the managers’ associations lasted two years, after which a compromise, much closer to the position of the industry than to that of the government, was reached. In the meantime, the salience of the issue in the press and

19 Laws passed in 2005 and 2007 only imposed relatively mild transparency requirements (Gomez, 2010). A 2009 law limited benefits for senior executives, but only for companies that received public money (Culpepper, 2011). The principle of shareholder influence over executive pay has been introduced in a new self-regulatory code adopted in 2013, but it is not mandatory.

among the public had waned, allowing managerial interests to launch a lobbying offensive which prevented even that compromise from becoming law (Culpepper, 2011: 86, 108; Schnyder, 2012: 1442-1444; see also Corhay and Tourani, 2000).

To sum up, the mere fact that “losers do not disappear in politics”, often invoked to point to an important mechanism of gradual institutional change through shifts in the social coalitions that underpin an institutional configuration, is per se uninformative. Barring perhaps a few extreme situations, the losers never entirely disappear, and therefore this constant feature of politics cannot by itself explain variation in whether and how institutions change endogenously. The particular attributes of specific institutions, however, can help address this issue. Although institutional incumbents will generally have a strong incentive to perpetuate the status quo, since it guarantees their power, different institutional configurations may afford them different levels of opportunity to make it costly for reformers to mobilize sectors of rule-takers in pro-reform coalitions. Given the typically cyclical dynamics of issue salience among rule-takers, institutional configurations that give power holders control over the timing of institutional reform place them in a very strong position to influence the shape, the sustainability and even the possibility of reform coalitions, and to resist undesired institutional change.

---

20 To be sure, losers may have different capacities to reverse or chip away at unfavorable institutional settlements by virtue of the political and institutional context. See the discussion in Patashnik (2008, esp. 25-31) in relation to “general interest” legislative reforms in the U.S. over the past three decades.
Conclusion: Reactionaries and the politics of institutional stability

Theories of gradual and endogenous institutional change have yielded important progress in the conceptualization of how institutions evolve over time. Their focus, however, on how institutions can be the object of strategic action aimed at various forms of change, if taken exclusively, ultimately risks exposing the theorization of institutional development to a paradox. Any institutional theory is ultimately premised on the gap between social change and institutional change. The two differ in intensity, extent, and, most importantly, in their temporalities. At one level, social change is continuous (e.g. Abbott, 2001: 255) and institutions are necessarily affected, at least in the basic sense that the individuals that occupy institutional roles move on to different roles, retire, or die, and are replaced by other individuals with different ideas and social backgrounds. But even though actors and context change continuously, the constant stream of social change does not necessarily, automatically, and fully translate into institutional change. Ultimately, the primary task of any institutional theory is to theorize why this is the case. Yet the actors that take center stage in HI theories and narratives of institutional development are reformers; our inquiries are typically geared towards understanding how and why they attain their goals in full or in part. Hence the paradox: if the primary goal of institutional theory is to explain the gap between social change and institutional change, then we should be at least as concerned about theorizing the conditions under which institutions enable or
constrain social and political actors to advance institutional stability—how institutions can empower the reactionaries, not in the ideological sense of the term but in the more neutral sense of “opponents of institutional change”. In the end, it is their agency that inserts the wedge between social and institutional change.

To understand when institutional power-holders will be successful in opposing institutional change, this paper has proposed to integrate existing theories of endogenous institutional change with a new theorization of how incumbents may or may not be empowered under existing institutional arrangements. Theories of endogenous change posit multiple micro-foundational processes as drivers of bottom-up institutional change: rule non-compliance and coalitional shifts. This paper theorizes those aspects of the existing institutional configuration that might limit or counter such drivers of change. As powerful as these processes may be in some circumstances, institutional change is not ineluctable, nor is the pace and direction of change entirely the product of social forces interacting on a level, institution-less playing field. Pre-existing institutions can influence the extent and even the possibility of gradual transformational change in two distinct ways. First, they can institutionalize cultural categories, which makes it more likely that rule defection will be sanctioned, consolidates understandings of formal rules, and inserts a wedge between imperfect compliance (likely to exist everywhere) and gradual institutional transformation. Second, institutional constellations may empower incumbents to steer or stymie
transformative change by allowing them to time decisions on institutional reform, which can be used strategically to make it more difficult for “losers” of previous institutional fights (again likely to exist virtually everywhere) to challenge the coalitional equilibrium underpinning an institutional arrangement.

The discussion has implications not only for theories of endogenous institutional change but also for other much-used conceptual tools in HI analyses such as critical junctures and path dependency. With respect to critical junctures, typically conceived as moments of fluidity and uncertainty in which path-dependent institutions are created (Capoccia & Kelemen, 2007), the present discussion offers a handle on the oft-bemoaned problem that identifying a critical juncture can only be done ex post, i.e. once the new institutional arrangement has displayed path-dependent characteristics over the medium to long run (e.g. Taylor, 2009: 493; Peters, Pierre & King 2005: 1289). Some have argued that the durability of new institutions is directly correlated to the level of conflict characterizing the juncture. The argument is that losers of the initial institutional battle, in those cases where the battle was hard-fought, are more likely to continue to oppose and ultimately reverse the institutional outcome (e.g. Arneil, 2010). This argument, however, has never been thoroughly tested. Indeed, substantial evidence exists that institutions established after intense conflict may prove enduring (e.g. Levitsky & Way, 2012). The present discussion shifts attention to two aspects of an institutional configuration created during a critical juncture that allow for
probabilistic predictions on how likely it is that the institution will resist bottom-up processes of endogenous change. If the institution emerging from the juncture incorporates cultural categories that are already used in a consistent manner across a number of connected institutions and policy areas, or if it concentrates control over the timing of reform in the hands of incumbents, then it might be possible to say relatively early on that the critical juncture resulted in a path-dependent institution that is relatively resistant to bottom-up processes of endogenous change.

Regarding path dependency, the paper builds on the insights of the literature on ideas and institutions, and proposes to integrate the original theorization of “interpretive” feedback effects with a more explicit focus on the capacity of institutions to entrench cultural categories. Furthermore, elaborating on the criticism of metaphors of institutional “reproduction” (Thelen, 2004: 293), the discussion lays down the foundations for theorizing, classifying and studying empirically the proactive strategies that institutional incumbents may pursue to safeguard their power. It may not always be easy to identify such strategies, since they may be less visible than reformers’ strategies, but this lack of visibility is at least partially attributable to the failure of current theories and research designs to make room for them.

Defenders of the institutional status quo are not just reactive “veto points” standing in the way of reform; they typically take initiatives to slow down, channel or stop change. To produce a compelling theoretical image of the politics of institutional
change, we should theorize these strategies more explicitly. By theorizing the types of institutional arrangements that help incumbents block institutional change, this paper has suggested two priorities in this research agenda. On the one hand, we should study systematically the initiatives that can be employed by institutional power-holders to promote or further entrench cultural categories that legitimize the institutional status quo. These might include education or propaganda initiatives, and symbolic policies aimed at reinforcing corporate or collective identities (e.g. Ahlquist & Levi, 2013). The cost and effectiveness of such strategies in defending given cultural categories are likely to differ substantially depending on whether the same cultural categories already inform the design and logic of other connected institutions and policies. Institutionalist scholars are well equipped to study these processes: a well-developed literature emphasizes the role of cultural production in triggering institutional change (e.g. Blyth, 2002), and many of its core insights can be extended to understanding how cultural production can be used to support institutional stability.

A second priority is to complement our extensive and sophisticated knowledge on “cycles of attention” in the mass public (e.g. Jones & Baumgartner, 2005) with further theorization and empirical research on two issues. First, understanding how cycles of attention vary across institutions with target populations that differ in size, ideological intensity, and involvement in other institutional arrangements is likely to offer insights into the ability of incumbents to maintain the status quo by strategically timing
decision-making on reform. Second, we should seek to identify the strategies that incumbents can intentionally (and instrumentally) enact to “crowd” the agenda, and how such strategies may be more or less successful depending on the institutional and political environment.

To pursue these theoretical and empirical goals, we should not abandon current theories of endogenous institutional change but rather extend them to understand more fully the institutional conditions and the strategies that favor endogenous institutional stability. This paper has endeavored to advance the theorization of the conditions under which institutions are best conceptualized as equilibria—because their impact on the resources, incentives, strategies and attitudes of rule-makers and rule-takers makes them likely to endure—and the conditions under which institutions are best conceptualized as arenas—relatively open to change and renegotiation. By shedding light on two important factors that contribute to institutional stability, it has made the case that the conditions that favor the defenders of the institutional status quo are as critical to understanding institutional development as the conditions that allow political actors to push for change. Reactionaries should populate our narratives of the politics of institutional change as much as reformers.
Bibliography


Jacobs, A. & Weaver, K. Forthcoming. When Policies undo Themselves, *Governance*.


Smith, R. (2006). "Which Comes First, the Ideas or the Institutions?" In Shapiro, I., Skowronek, S. & Galvin, D. (Eds.), *Rethinking Political Institutions*, (pp. 91-113), New York: NYU Press.


