From metaphors to measures: observable indicators of gradual institutional change

PHILIP ROCCO
Travers Department of Political Science, University of California, Berkeley, USA
E-mail: procco@berkeley.edu

CHLOE THURSTON
Department of Political Science, Johns Hopkins University, USA
E-mail: cthurst1@jhu.edu

Abstract: Scholarship on social policy has recently emphasised the importance of gradual processes of institutional change. However, conceptual work on the identification of processes such as drift, conversion and layering has not produced clear empirical indicators that distinguish these processes from one another, posing major problems for empirical research. We argue that, in order to improve the validity of its empirical findings, scholarship on gradual change should – and can – pay more attention to issues of measurement and detection. We then contribute to this goal by clearly articulating observable indicators for several mechanisms of gradual institutional change and validating them against extant empirical work on political economy.

Key words: gradual change, historical institutionalism, process tracing, social policy

Introduction

In the last decade, studies of institutional change have increasingly moved away from examining highly visible political processes and events. Rather than focusing on major policy reforms or regime changes, there is increasing attention to how "subterranean political processes" have shifted social policies in the United States and other advanced industrialized democracies (Hacker 2004; Mahoney and Thelen 2010; Streeck and Thelen 2005). When strong financial regulations fail to adapt to revolutionary changes in the market, for instance, their practical effect is significantly weakened. Similarly, when an existing public education system is forced to compete with private charter schools financed by the state, it may cease to carry out the function it did before its competitors

arose. However important they are, these processes of "policy drift" and "layering", respectively, cannot be readily tracked with the standard tools for observing institutional change.

In this article, we argue that, in order to better identify, explain and distinguish gradual-change processes from one another, scholars should place a greater emphasis on the observable implications of these processes, developing valid indicators out of systematised concepts (Adcock and Collier 2001). Since the mechanisms posited by gradual-change scholarship are not easily identifiable, the absence of clear observable implications creates challenges for empirical research (King et al. 1994). The utility of metaphors for institutional change is further challenged if one wishes to adjudicate between multiple possible narratives to describe an event or to test more general hypotheses about why a class of events tends to occur. In spelling out empirical implications of gradual-change processes, we do not claim to have provided a definitive guide to these processes, nor do we claim that current empirical research on these processes is invalid. However, by nominating clear observable implications for several commonly invoked processes of gradual institutional change (drift, conversion and layering), we hope to promote "visual acuity and intellectual responsibility" (Tilly 2004, 597) in this field of study, allowing scholars to carry out better empirical research as well as to identify where the gaps in the gradual-change framework are.

Our argument proceeds in four parts. First, we discuss what we term the Gradual Change Framework (GCF) and its contribution to historically rooted scholarship on policy change, as well as the absence of empirical indicators for processes of change identified by this scholarship. Second, we synthesise extant empirical research on four dimensions of gradual institutional change (initial ambiguity, change-agent resources, and institutional structures and outcomes) to produce a set of empirical indicators that allow us to distinguish descriptively between processes of gradual institutional change. Third, we validate our indicators by testing them against two landmark works in the study of gradual change, Kathleen Thelen's *How Institutions Evolve* (2004), as well as Jacob Hacker and Paul Pierson's *Winner-Take-All Politics* (2010a). The empirical research done in both of these works, we argue, generally provides support for the claim that processes of gradual institutional change can be distinguished from one another using the indicators we develop here. Where the evidence provided by this research is more ambiguous, our framework reveals possibilities for future empirical research to substantiate arguments about gradual institutional change. We conclude with a discussion of the stakes and implications of our argument, contending not that metaphor and measurement are at odds,

but rather that a clear and careful use of measurement will make for a more compelling discussion of metaphors and mechanisms.

The challenge of detecting gradual institutional change

Though long a topic of interest in institutionalist research, the problem of explaining institutional change developed into a coherent agenda through the 2000s (North 1990; Blyth 2002; Campbell 2004; Streeck and Thelen 2005). Early scholarship on institutional change often assumed that social and political institutions were generally stable, changing mainly under conditions of major exogenous pressure in sharp, discontinuous and visible breaks with the past. Responding to this claim, major works by Thelen (2004), Streeck and Thelen (2005), Mahoney and Thelen (2010), Pierson (2004, 2005) and Hacker (2004) have demonstrated that institutions may also shift without exogenous pressure, and do so gradually – sometimes undetected – over hundreds of years.

As a result of these discussions, scholars have developed a multitude of labels to explain the variety of channels through which institutions may change gradually. Falleti and Lynch (2009, 8), for example, point to 14 such causal mechanisms, including conversion (redirecting old institutions to serve new goals), layering (the construction of new institutions alongside similar existing institutions), drift (institutional change through policy inaction) and displacement (the dislodgment of an old institution with a stable new one) (see also Streeck and Thelen 2005). These labels – many of them metaphors – have improved the GCF in several ways. First, they have drawn researchers' attention to lesser-noticed features of the political, social and economic landscape. Rather than focusing only on overt policy action, concepts like drift have highlighted the way that policy inaction can change policy outcomes by virtue of policymakers' inability to respond to underlying shifts in the social and economic context (Hacker 2004; Hacker and Pierson 2010a, 2010b). Second, since these metaphors are broadly framed (rather than specific to each setting), they can travel to other contexts to help explain processes of change in multiple contexts. For instance, gradual-change metaphors can be invoked to characterise the development of both legislatures through the layering of multiple institutional designs over time (Schickler 2001), as well as changes in the implementation of public policy regimes over time, such as the layering of public policy institutions that encourage the development of private markets for insurance and gradually pull away support for pre-existing public programmes (Hacker 2002). More recently, scholars have drawn from the GCF to explain the development of an internal research market in the European Union (Chou 2012), as

well as the Eurozone's institutional response to the series of crises beginning in 2007 (Salines et al. 2012). Yet, despite several calls for greater attention to the operationalisation of these mechanisms across contexts (Immergut and Anderson 2008; Falleti and Lynch 2009), the utility of these approaches has largely remained heuristic.

Embracing these metaphors as labels for complex processes of change poses a practical challenge: the mechanisms they refer to are largely invisible. In contrast to the major changes associated with exogenous shocks, such as the zeroing-out of institutional budgets or the creation of novel packages of policy reform, we cannot view processes of layering, drift and conversion so directly. None of these processes often invokes direct budgetary shifts, and few involve formal changes to policy that are visible in public legislation, administrative rules or judicial decisions. In the absence of easily identifiable empirical indicators, how can one make an argument about gradual institutional change?

An additional challenge is in the portability of these concepts, a central goal of the GCF, and yet one that poses persistent problems. Mahoney and Thelen (2010) attempt to address these problems in their edited volume by further specifying the conditions that promote each type of gradual institutional change. The project contributes to the GCF agenda by discussing previously overlooked issues such as agency and the role of structural features such as policy ambiguity or institutional veto points in promoting certain types of outcomes over others. However, while specifying in greater detail the conditions of gradual institution change, Mahoney and Thelen (2010) and their contributors do not tackle the underlying challenge of identifying institutional change; rather, their framework merely adds layers of labels. The fundamental invisibility of these processes remains unaddressed.[1]

The consequence of vague metaphors for processes of change, then, appears sizable: even if we have well-developed theories of the channels through which these invisible processes operate, we will not be able to confirm or deny these causal claims with any confidence, nor will we be able to say how certain we are about our descriptive analytical choices (for instance, the choice to characterise the development of a particular

---

[1] We acknowledge that Mahoney and Thelen (2010, 30) go to great lengths to describe the processes of gradual institutional change, such as the identification of agents of change (insurrectionaries, symbionts, subversives and opportunists). Even so, the authors provide little in the way of clear empirical indicators to distinguish these types of agents from one another. Rather, the elaboration of these types of actors remain at a heavily theoretical level, describing actors as *concepts*, without providing additional detail about how to use observable behaviours to identify agents of change.

public institution as drift rather than conversion). The latter of these problems – the problem of *identifying* social processes – is especially immediate, because it concerns a basic challenge of social science research: to produce useful analytical operationalisations of concepts that can describe social processes.

Operationalising the GCF: observable indicators

Following a variety of scholarship on qualitative methodology (King et al. 1994; George and Bennett 2005; Collier et al. 2011), we argue that invisible mechanisms can produce visible implications.[2] In fact, by creating an initial catalogue of the visible signs of these processes, we think it is possible to begin to see the outlines of the mechanisms more clearly, and eventually, to provide qualitative categorisations for the universe of known cases of gradual institutional change (Adcock and Collier 2001). Additionally, developing a set of observable indicators of gradual institutional change might allow for a more fruitful discussion of any remaining gaps in the GCF itself.

To locate the visible implications of gradual processes of institutional change that would allow us to distinguish between them, we performed a close reading of canonical literature describing processes of conversion, layering and drift.[3] Based on the authors' descriptions of each process, we were able to contrast these mechanisms with respect to four indicators: (1) initial institutional ambiguity or malleability, (2) the way in which pre-existing institutions structure change, (3) the particular resources of change agents and (4) the outcome of institutional change. Table 1 lays out four indicators for distinguishing between processes of gradual institutional change.

After developing these indicators, we evaluated them against extant empirical scholarship on gradual institutional change, establishing an initial sense of the content validity of our approach (Adcock and Collier 2001). It is worth emphasising that these indicators do not imply strict rules of necessity or sufficiency. Rather, they are meant as an incremental and additive set of indicators to guide empirical classification of processes and to establish the level of certainty in a given analytical choice.

---

[2] We follow Elster (1989, 3–10) in suggesting that most mechanisms in the social sciences tend to be invisible but recoverable with empirical indicators.

[3] Though we reviewed a variety of literature, Mahoney and Thelen's (2010) broad-ranging literature review became central to our analysis. For the sake of space, and because it is a residual category in their study, we leave out a discussion of the process of displacement.

Table 1. Empirical indicators for processes of gradual institutional change

	Drift	Conversion	Layering
Initial institutional ambiguity or malleability:	Low	High	High
Institutional structures induce change through:	Path dependence, creating *status quo* bias	Granting actors discretionary capacities to alter institutional meanings	Coupling of multiple institutions; differential growth
Change agents must have access to discretionary, intellectual or material capacities to:	Keep institutional updating off of the policy agenda	Manipulate interpretation of institution's rules	Sponsor and carry out modest changes or propose marginal amendments
Final institutional outcome:	Institution must not change, but institutional outcome post-drift must not be similar to institutional outcomes pre-drift	Institution post-conversion must be functionally different than institution pre-conversion	Initial institution must change and should become similar to the institution that was layered on top of it

As incrementally more evidence on a case seems to confirm the existence of one process versus another, then, our framework allows researchers to classify that process as the one that "best" describes the observed phenomena and to do so with a high level of certainty. As the evidence becomes less confirmatory of the existence of one process, however, it may lead researchers to be more refined in their attempts at classification and more cautious in the level of certainty they report. We illustrate how our indicators allow analysts to distinguish more definitively between processes below.

Initial institutional ambiguity or malleability

One major difference between processes in the GCF is whether institutions change despite or because of the clarity with which they were designed. One way to assess the clarity of initial institutions is to examine the constraints imposed by the rules that comprise an institution. Compare, for example, a policy with a number of statutory constraints (say a categorical grant from the federal government to cities to establish highly specialised mental health facilities) to one those with few (say, a block grant, which provides funds to cities to "improve mental health"). Though the first may contain some ambiguities as Sheingate (2010) claims exist in all formal rules, it is clear that the second embeds within it significantly more ambiguity, in part because of the lack of rules that constrain action. The GCF claims that varying degrees of ambiguity embedded in institutional rules lead to different types of institutional change; highly ambiguous institutional rules elicit change through layering and conversion, while less ambiguous institutions elicit change through drift.

Ambiguous institutions – with few constraining rules – provide an opportunity structure for groups that suffer losses in a given institutional settlement to recast institutional goals. As extant research on social policy confirms, information about initial institutional ambiguity itself can help us distinguish between two groups of processes rather than each process individually. In processes of conversion and layering, institutions with highly mutable or ambiguous rules allow losing groups to reinterpret policy goals either because of their discretionary resources (conversion) or because of their ability to marshal other resources to propose new institutional solutions alongside old ones (layering). In particular, research shows that ambiguous statutory language can allow groups to reshape institutions to fit their preferences. Howell (1992), for instance, recovers evidence of conversion (without specifically labelling it as such) in French industrial relations reform through a close reading of laws and a legislative history of these enactments, which demonstrates that ambiguous

statutory language permitted political parties of the right and left to implement industrial relations policy differently without formally changing the rules.[4]

Research on social policy development also shows that layering is marked by initial institutional ambiguity. As Henig's (2008) work on the emergence of conservative strategies to retrench US secondary education programmes created in the 1960s shows, actors hoping to reform education depended upon ambiguities in the initial design of the programme, which both created funds for public education as well as tools to assess and hold the public education system accountable, placing it in tension with privatised alternatives from the beginning. Conservatives – Henig colourfully reveals using internal strategy memos – referred to the coalition built by these statutory ambiguities as the "Education Blob" (p. 298). As with the reforms discussed by Howell (1992), the ambiguity contained in the Elementary and Secondary Education Act is not absolute; it was clear to all observers that the purpose of the reforms was to create new funding opportunities for public schools. However, the reforms' provisions on evaluation were vague, permitting some uncertainty about whether or not the laws' authors intended public schools to compete directly with private alternatives. These tensions between funding and performance evaluation permitted conservatives to propose the viability (and even superiority) of private alternatives to the public system in the 1990s, using performance metrics to buttress their critiques of public schools and their advocacy on behalf of voucher programmes.

Drift, the remaining process of institutional change that we elaborate on, is distinguished from conversion and layering by having more explicit institutional formulations, which make it more difficult for agents of change to remould rules internally to fit their goals. When rules are firmer – either due to plainer, less malleable language or to fewer internal contradictions – they may be less vulnerable to structural mutation, and agents of change may have a difficult time manipulating them to their advantage. Unambiguous initial institutions can thus result in change through drift, in which precise institutional rules do not adapt to changing social and economic contexts.

Marmor (2000), for instance, recovers evidence of drift in retiree health insurance policy in the United States by demonstrating a gradual shift away from specific goals of broad-based coverage enshrined in Medicare's enacting legislation. To demonstrate the specificity of Medicare policy at

---

[4] See also Dodds and Kodate (2012) for a discussion of *vague* initial rules, as opposed to multiple co-extant rules.

the outset, Marmor not only details its statutory basis but also illustrates differences between understandings of the law possessed by Medicare's enacting coalition and political coalitions that ultimately refused to update the policy to adapt to changing economic circumstances (pp. 151–157). Without a doubt, providing evidence to evaluate an institution's initial ambiguity is difficult. Legislative history is, however, one of many ways to reveal whether or not multiple audiences perceived a given institution in the same way. As the research presented here suggests, understanding which processes of institutional change are likely to be at work depends critically on our ability to discover how an institution appeared at the time of design.

How institutional structures induce change

Existing institutions are critical to processes of institutional change, because they create differential opportunities for defenders of the *status quo* to veto proposed alterations and afford actors opportunities to exercise discretion in initiating those changes. The way these structures work to shape actor motivations and resources (and the evidence required to satisfy them) differ, however, across change processes, permitting us to use evidence about how institutions work as a way of demarcating processes from one another.

Drift is induced when path-dependent features of institutional design – whether the result of collective action problems associated with updating an institution, dense interest group environments that may make coordination on change difficult to accomplish or the adapted expectations of policy makers – prevent actors from easily updating an institution (Pierson 2004). We can establish that drift has occurred in a given case by looking at the difficulties faced by institutional reformers. Updating institutions to reflect current exogenous environments is most likely to be difficult in institutions where high numbers of veto points make assembling a coalition for institutional change difficult (Immergut 1992; Cox and McCubbins 2001; Tsebelis 2002) or when actors with a high level of discretion over the administration of a given policy oppose updating it to reflect current exogenous environments. Thus, when institutional change occurs in environments that are dominated by such identifiable veto points, it is reasonable to suspect that drift may be at work. By contrast, when few veto points are identifiable, we may be able to rule out the process of drift as a characterisation of the case.

Conversion, by contrast, implies that the institution in question is "open" or "susceptible" to manipulation by change agents who can alter its meaning. As with drift, the visible signs of conversion are the institutional designs that generate it. As noted in Table 1, institutions generate conversion by containing ambiguous meanings for essential rules. In

practice, institutional ambiguity is present in a variety of forms. One of the more common forms, however, is that of an institution with multiple, non-complementary logics or goals. Howell (1992), for example, finds evidence of how institutional structures induce conversion by highlighting two competing logics of France's Auroux labour reforms, which both sought to induce collective bargaining processes *and* self-management by workers (*autogestion*) supported by right-leaning parties. Since these two goals co-existed in the language of the law, the implementation of the reform fluctuated between the two depending on which party was in power. Howell's argument is strengthened by detailed, case study evidence of how shifts in party led to shifts in on-the-ground policy implementation without formal updates to the law.

Layering can be distinguished from the above processes, because it occurs as the product of institutional interaction. As Schickler (2001) highlights, layering occurs when one institution directly challenges another parallel institution, wearing away support for that institution over time. A good example of how layered institutional structures induce change is again provided by Henig's (2008) work on the transformation of public education. Henig shows that layering took the form of new institutional alternatives to public education, such as voucher programmes at the state and local level, which developed *after* the crafting of secondary education reforms but in the absence of formal revisions to federal policy. Importantly, the existing institutional structure did not dismantle public education wholesale. Rather, it placed it into direct confrontation and competition with private alternatives, much to the system's detriment. This type of process clearly distinguishes layering from conversion and drift. In Henig's case, institutional change occurred through the development of clear alternatives to existing public policies rather than by the failure of reformers to update those policies or via a gradual shift in the meaning of those policies over time, which he illustrates vividly with evidence of the timing of institutional changes in both public and private education sectors.

How agents initiate change: motivations and resources

Another theoretical advance in the GCF is its recognition of "change agents", individuals motivated by their desire to alter institutional structures and facilitated by the resources they possess to change those structures. Generally speaking, agents should initiate institutional change when they possess enabling resources. However, as a reading of the empirical literature makes apparent, these resources differ substantially across processes of institutional change.

With drift, change agents try to preserve an existing institution by blocking attempts to adapt it to changes in the external social or economic environment. To do this, agents must possess discretionary, intellectual or material resources to keep updates off the agenda. Evidence of drift, then, can be recovered in several additional different ways. Most importantly, drift presumes the existence of agents who occupy formal institutional positions with decision-making discretion, empowering them to act to alter an institution's trajectory. Moreover, in this capacity, they must keep off the agenda updates to the existing institution that are proposed by "institutional defenders" who want to preserve the existing institutional design against erosion. If such agents do not possess formal institutional discretion, however, other resources must be in place to permit them to enact drift. For instance, such change agents may possess material resources (e.g. interest group mobilisation capacity) that give them access to and influence over those with formal discretion. Alternatively, change agents may possess intellectual resources such as scientific authority that allow them to rationally persuade those with discretion that updating a given institution is not technically feasible (Espeland 1998).

A good example of how change-agent behaviour can reveal the process of drift is provided by Carpenter's (2010) work on financial reform during the Obama administration. As Carpenter's evidence on legislative and executive decision-making shows, a number of reform proposals that plotted course corrections for regulation of complex financial products (those excluded from the category of "plain vanilla" products) were kept off the agenda by conservative Democrats whose pivotal votes dictated the shape of reform bills during committee markup in both houses of Congress. Committee staffers, seeking to appease these members, deleted new regulations from evolving legislation. In addition to these actors with formal discretion in Congress, Carpenter's evidence on actions taken by economists at the Federal Reserve also shows that actors in control of reputation-conditioned legitimacy in the policy realm can also keep issues off the agenda. In the case of financial reform, economists in government agencies – with a disciplinary agenda that precluded imposing regulations in places where clear market failures could not be identified – helped to seal the fate of a number of financial reform proposals by failing to consider them at all.

During the process of conversion, agents of change also tend to be motivated to refashion an institution, albeit not through formal revisions, but instead through the manipulation of mutable institutional meanings. To do this, agents must either possess discretionary resources to change the meaning of a given institution or have material or intellectual capacities that give them access to discretionary agents with the ability to

manipulate the meaning of a given institution. As with drift, arguments about conversion usually depend on agents with power to change institutions after their enactment. However, conversion can be distinguished from drift because we are not looking for agents with the ability to keep issues off the agenda. Rather, we are looking for agents who can change the meaning of institutions themselves.

Teles' (2008) work on the mobilisation of conservative lawyers and legal scholars provides an example of how agents are mobilised to convert existing institutional meanings. Conservatives, Teles argues, faced a major challenge to altering liberal legal institutions: law schools and the organised bar possessed a "still-powerful liberal understanding of rights" (p. 2). To combat this understanding, conservative entrepreneurs poured their resources into the development of institutions for the study of law (such as academic programmes in Law and Economics) that shifted the understanding of rights among legal intellectuals and Washington lawyers. In doing so, these entrepreneurs hoped to gradually mutate long-established legal doctrines until they were more in line with conservative philosophy. As is consistent with Mahoney and Thelen's description of conversion, these agents seek not to push new policy alternatives, nor to keep policy issues off the agenda, but to remake existing institutions (in this case legal institutions such as canons of construction and interpretation) from the inside out.

Layering, by contrast, is distinguished by the activation of change agents who have the capacity to carry out modest programmatic changes, which peel away public support from the extant institution. Their reasons for creating new institutions may or may not be related to a desire of a "destructive" nature. Indeed, their efforts may not initially be perceived as contributing to institutional opposition, though they often are. To engage in layering, agents must have the resources to sponsor and carry out modest changes at the margins of existing policies. This can include discretionary resources to propose and implement new institutions, material resources to lobby on behalf of a new institution and intellectual resources to rationally persuade discretionary agents that adoption and new implementation of a new institution is necessary.

The mobilisation of agents in the service of institutional change through layering is described especially well in Henig's (2008) research on the conservative reform of public education. Similar to work on layering by Béland (2007), Henig provides evidence about the institutional position of reform's proponents. He effectively shows that the conservatives who transformed public education institutions were already inside the educational establishment, serving as directors of state departments of education, as part of the education bureaucracy during the Reagan presidency or as

intellectuals in burgeoning conservative think tanks. These agents were thus well disposed to propose and carry out alternative approaches to education policy, using performance metrics to tell a story about a "failing" system of public schools and approving experimental waiver programmes that permitted private schools to compete more directly with the public system. Rather than altering institutions from the inside out, change agents had the discretion to make modest investments in new institutions in order to peel away support from the extant system (Butler and Germanis 1983).

Institutional outcome

A final difference between the processes of institutional change highlighted by the GCF is the eventual outcome of the process in question. Though the implication of all of these theoretical processes is that a given institution appears highly dissimilar in its logic and operations than it did before a process of change was undertaken, these outcomes differ in important ways that require distinctive empirical indicators. We characterise those indicators here.

The outcome of drift implied by the GCF is that, after social or economic changes have occurred, an institution must fail to adapt to changing circumstances. As such, the process of drift is recoverable through a few different types of evidence. First, it could be signalled when actors within a given institution (or those observing it) acknowledge that the state of the world has changed, but the institution has not "kept pace" with changes. Though these statements are often the product of political strategy, public rhetoric can still tell us quite a bit about how policy makers and the public perceive the status of a given institution (Mayhew 2002). Second, drift is clearly signalled when new patterns of social or economic behaviour (e.g. rising inequality) occur and would not have occurred if the counterfactual condition of a policy reform had been met. In other words, since we can assume that policies have effects on social or economic behaviour (Pierson 2004), changes in this behaviour that would be unlikely under a counterfactual reform scenario are valid indicators of drift. Particularly extreme forms of this behaviour include regulatory failures (e.g. the 2008 financial crisis) resulting from a failure to take into account new risks (Goldin and Vogel 2010).

By contrast, the process of conversion can result in an institution being implemented differently "on the ground" after its meaning has been manipulated by actors with discretion. Thus, there are several types of evidence that are confirmatory of conversion. First, conversion could be signalled by the differential application of policies or procedures in a given institution over time. For instance, if frontline bureaucrats, over

time, come to understand their responsibility to client populations that were not originally included in statutory or administrative rules, this signals some evidence of conversion. An example of this process can be found in the work of Pressman and Wildavsky (1973), which describes how city officials involved in implementing federal urban policy in the 1960s shifted the pool of intended policy beneficiaries away from those discussed in statute (the impoverished) and towards other client bases (powerful political interests in urban areas).

Second, conversion might be signalled through the differential definitions of policy goals expressed by political and administrative actors. For instance, when political and administrative actors no longer saw the Tennessee Valley Authority as the agent of the Roosevelt Administration but of local economic interests, it is fair to say that a conversion of its goals or purposes had occurred (Selznick 1949). Finally, conversion might be signalled by the development of policy innovations (such as monetary or normative incentives or technological instruments) that are directed at different purposes over time. For instance, as social welfare programmes began being deployed as a means of discipline in the 1990s, Soss et al. (2011) tracked substantial changes in the techniques used to sanction clients.

Layering changes institutions by creating viable alternatives to the original institution. Its effects are distinctive from those of conversion or drift, because it results in specific social, economic or political adjustments that correspond with the alternative institutions that develop later in time, rather than the changed implementation of a policy that corresponds with new understandings of its initially ambiguous meaning. In particular, layering results in changes in public support for a given institution, differential growth of the old and new institutions and, finally, the obsolescence or adaptation of the first. There are thus a variety of distinctive observable implications of this process. First, as Esping-Andersen (1990) and Hacker (2005) highlight, layering could cause shifts in membership to a new institution, such as the decision by citizens to supplement their public insurance with private insurance. Second, layering could cause shifts in elite support across layered institutions, as in Esping-Andersen's (1990) discussion of welfare-state dualism or Rothstein's (1998) study of customised private alternatives to social insurance in Sweden, both of which demonstrate the feedback that can occur when private alternatives are layered on public services. Third, initial institutions could become obsolete, with gradual reductions in funding or administrative support over time, as in Palier (2005) or Levy's (2005) studies of the liberalisation of French social policy. Otherwise, public institutions may have to adapt to become like private institutions,

as Henig (2008) suggests occurred when public schools began to borrow techniques seen as "successful" in their private counterparts.

Assessing narratives of institutional change

The empirical indicators outlined above are only useful to the extent that they can be verified against extant research on gradual-change processes. Following the agenda of Adcock and Collier (2001), we validate our framework by applying it to studies, invoking mechanisms of gradual institutional change. In this section, we examine three cases that make use of the GCF, drawing two from Thelen's landmark 2004 study of gradual institutional change in the German system of skills training, as well as examining a more recent case of gradual institutional change identified by Hacker and Pierson in *Winner-Take-All Politics* (2010). Since these texts represent some of the most promising applications of gradual-change concepts, examining them has the benefits of both providing a "critical test" of the usefulness of our scope conditions and indicators and opening up new lines of research on these subjects by students of gradual institutional change.

Studying these cases also illustrates the value of our method by highlighting where the research could go further in providing empirical evidence for its claims about gradual institutional change. As we illustrate below, at least one of the empirical implications we discuss here (initial institutional ambiguity) has proved vexing for scholars to substantiate with data, especially in the case of drift. By clarifying the empirical indicators of gradual institutional change and showing where scholarship is at pains to make the case for gradual institutional change, our approach has the advantage of guiding future investigations of social policy rather than merely critiquing existing work.

Conversion and layering in the development of German skill-formation institutions

In her 2004 book, *How Institutions Evolve*, Thelen examines the origins and development over time of skills training in Germany, Britain, the United States and Japan. Thelen examines in detail the development of skills-training institutions in Germany over a century, allowing her to identify several mechanisms that have gradually shifted the institution away from its original intention, actors and meanings. A central puzzle of Thelen's involves the German vocational training system: how did a system that was initially anti-labour transform into a symbol of the German social-democratic model? And how did a system originally characterised by decentralisation develop into the German training system

as we know it today – one known for its uniformity and standardisation? Thelen identifies two mechanisms that explain the development of the original Handicraft Protection Law into what was considered the German model by the mid-twentieth century: layering and conversion (pp. 35–36).[5] We examine each process separately below.

The layering of an industrial training system atop an artisan one.   The first major narrative presented in Thelen's case study of the German vocational system is that of the *layering* of an industrial skills training regime atop the existing artisan (Handwerk) system developed in the nineteenth century. Rather than create a training system from scratch, industrial employers and organised labour created a system of skills certification that ran parallel to the artisan system based in the nineteenth century. However, as Thelen writes, the German training model ultimately shifted "away from the decentralized artisanal system toward the centralization, standardization and uniformity that are not considered its core and defining features" (pp. 35–36).

How well do Thelen's details of the change process correspond to the observable implications that we argue suggest a process of layering? First, Thelen's discussion of the history surrounding the 1897 Handicraft Protection Act depicts a law with clear statutory constraints endowing the artisan sector with para-public authority to regulate and certify apprentices (see the first row of Table 1). While clear with respect to authority over skills certification, the Act also unearthed a "negative space" with respect to alternative skills-based organisations, enabling those organisations to develop their own capacities so long as they did not encroach on the rights granted to the handicraft chambers (pp. 50–52). The initial omission reflected the initial mission of the Act whose designers, as Thelen shows, were concerned not with skill training as a whole, but rather with destabilising the growing political power of labour by placing the authority to certify skills in the hands of artisans (p. 44). In short, the empirical evidence Thelen provides in the form of legislative history roughly fits the observable implications that we argue are likely to correspond either to layering or to conversion.

As the second row of Table 1 suggests, for layering to be a likely driver of change, we should also expect to observe conflicts occurring between the original institution and the institution on which it was allegedly "layered" atop. Thelen's account of the early years of development of the industrial skills system shows that the parallel institutions often came into

---

[5] Though we do not focus on it in our evaluation here, Thelen also highlights a process of drift to explain recent challenges to the German system.

contact, but does not provide evidence that the interaction was antagonistic. Rather, between unions' desire to preserve their members' skills certification and competitive pressures between the machine industry and artisan sector, there was a clear incentive for actors within the industrial sector to "accommodate and in a sense work around that system" rather than try to create an entirely new system (p. 41). Thelen does, however, provide evidence of differential growth occurring between the artisan and industrial sectors as a broader pattern of industrialisation takes form in the twentieth century. Thus, without having to come into direct conflict, broader contextual factors appear to have gradually eroded the power of the artisan sector while simultaneously increasing the prominence of industry. In a way, then, the change we observe corresponds to drift as much as it does to layering.

Moving onto a discussion of the resources of change agents (in this case, industrial employers and unions), however, pushes the balance of evidence back in favour of layering (see the third row of Table 1). Thelen's evidence does not suggest that either of these agents sought to accomplish institutional change through inaction or that they did not have the resources to alter the institutions more directly (both of these corresponding to drift). Instead, Thelen shows that, once unions were granted the entry (during the Weimar era) into the German skill system debate, they sponsored modest amendments and attempted to preserve many existing elements of the system (pp. 53, 62).

Finally, as the fourth row of Table 1 suggests, post-layering institutions should not be similar to the pre-existing institutions. As of the 1920s, the industrial training regime appeared layered alongside the traditional craft system. However, in the post-war period, the interaction of these two layers finally produced the institutional outcome Thelen alludes to: one that was no longer decentralised in the handicraft tradition, but rather characterised by "a much higher degree of standardization and uniformity" (p. 41). While the evidence conforms to what we should expect to observe for a process of layering to be likely, there is some ambiguity as to whether that last step to move towards a more standardised and uniform skills training system was accomplished through layering or required an agent to actively convert the institution in its final moments. Here, Thelen's account would be strengthened by considering whether additional information might instead support the logic of conversion, at least in completing the process of institutional transformation.

Vocational training from labour foe to labour friend.   In contrast with the evidentiary ambiguities of the first example, Thelen's discussion of the conversion of the training system into one that embraced labour as a key

social partner is much clearer with respect to the observable implications we presented. Thelen argues that, in addition to the layering of new industrial training institutions alongside old ones, actors – once incorporated into the formal system – also engaged in the process of institutional *conversion*, redefining the institution away from its aims of labour exclusion towards a model that incorporated both industry and labour as full partners. By 1969, these partners were able to shift the purpose of the protections in place since the 1890s from a system aimed at labour suppression to one that considered labour a vital administrative partner in industrial skill formation. As with the previous example of layering, we can also examine whether indicators for conversion would be met based on the evidence provided in the narrative.

The first evidence we look for are indicators of how open the initial institution was to renegotiation (refer to the first row of Table 1). In addition to the contextual features of the 1897 laws discussed in the previous section, Thelen draws on employer newsletters to illustrate employers' recognition of the "lack of specific legal grounding for chambers in the area of vocational education", leaving the training system highly mutable (p. 248). This evidence that the institution was open to manipulation by change agents lends additional support to the argument that conversion could be a likely driver of change, though as in the previous case, these conditions also lend themselves to layering.

For conversion to be the driver of change, we would also expect the agents themselves to have the resources and interest in manipulating these multiple institutional meanings (as suggested in row three of Table 1). As Thelen points out, employers capitalised on the legal ambiguities, incorporating them into their rhetoric "in conflicts with unions over regulation of firm based training" (p. 248). During the 1950s and 1960s, labour and employers clashed about the meaning of skills training and the purpose of employer self-governance, further suggesting that change agents had the capacity to redefine existing institutions (pp. 251–253). Clashes over the purpose of the skills training regime persisted through the debates leading up to the 1969 Vocational Training Act, with employers and labour maintaining different interpretations of who should regulate skills, as Thelen demonstrates in her coverage of the different rhetoric presented by each side (pp. 259–260). However, as Thelen's account suggests, the ability of the actors to reframe the meaning of skills-certification institutions was as important as the opening of a window in 1966 with the formation of a "Grand Coalition" between the CDU and the SPD (p. 260). It was in this context that labour and employers gained the discretionary capacities to alter institutional meanings (our third indicator), helping to produce institutional conversion. Through the 1969 law, labour was converted into a "key pillar of social

partnership" with management (p. 218), helping to solidify what we view today as the key features of the German employment system (our fourth indicator, as suggested in the last row of Table 1). The process by which this happened largely corresponds to the observable indicators in Table 1 that correspond to conversion.

Drift and the politics of increasing inequality

In *Winner-Take-All Politics*, Hacker and Pierson (2010a) explain the rise of increasing economic inequality in the United States including the hyper-concentration of wealth at the top of the income distribution as the result of changes in public policy as well as changes in the organisational basis of politics. Where policy is concerned, the theoretical core of Hacker and Pierson's argument is centred on the mechanism of policy *drift*, which helps to explain the shifting political fortunes of both organised labour and business in the latter half of the twentieth century. To be sure, drift is not the only mechanism at work in Hacker and Pierson's account, but it is the dominant mechanism that connects public policy to the rise of winner-take-all economics. Here, we review one major narrative of drift used by Hacker and Pierson and assess the extent to which the arguments put forward by the authors correspond to the evidentiary standards we outlined in the previous section.

Hacker and Pierson argue that changes in the organisation of the economy alone cannot account for the extraordinary weakening of American labour unions in the late twentieth century (pp. 56–62). In the 1960s, labour's power resources (measured by union membership) were particularly vulnerable to the geographical relocation of industries to states with more repressive labour laws (so-called "right to work" states). If policy makers had updated industrial relations policies to adapt to this pattern of relocation, it is likely that the propensity of an individual in the United States to join a union would not have declined as precipitously as it did, and union power would have remained a strong countervailing force against an increasingly aggressive business lobby (a comparison to Canada is utilised here). Below, we briefly review the extent to which Hacker and Pierson's evidence in this narrative meets the empirical standards indicated for drift outlined in Table 1.

First, where the lack of institutional ambiguity is concerned, it is not exactly clear that the *ex ante* industrial relations policies like the National Labour Relations Act were meant to protect union power. In fact, within little more than a decade of its passage, this policy was formally updated by the Taft-Hartley Act in ways that limited the interests of organised labour, permitting states to enact right-to-work statutes that severely curtailed union power (Gross 2010).

Although it might be plausible that labour relations policy "drifted" away from a settlement that unambiguously favored labour, the authors do not provide enough evidence to settle this score. If the authors were to deploy legislative history to demonstrate that labour relations policy, even in the wake of Taft-Hartley, was understood by a diversity of actors as protecting the capacity of labour to organise politically and to show that, at least at the outset, opponents of the act did not conceive of passing state-level right-to-work laws, the claim that labour relations policy was an unambiguous settlement in favour of unions – from which the United States drifted away over time – would be a more credible one. Thus, though the evidentiary standard for the claim of initial specificity as a criterion of drift is perhaps only partially met, the standards we highlight here help to suggest a path for future research on this instance of institutional change.

Passing over the way in which policy designs induced drift, we turn to the third indicator – change agents' resources to keep updates to existing policy off the agenda, as highlighted in the third row of Table 1. In the case of industrial relations policy, the resources of actors who wanted to forestall updates to the policy are clear: Hacker and Pierson (pp. 131–132) effectively describe how, in 1978, a powerful coalition of newly mobilised business groups (the Business Roundtable, the Chamber of Commerce and the National Association of Manufacturers) poured their resources into defeating updates to labour relations bills in the Senate. Moreover, Hacker and Pierson discuss how deliberate choices on the part of the President – who chose to focus his efforts on the Panama Canal treaty – left updates to industrial relations policy in the dust. Though it is apparent that labour's allies in Congress (namely in the House) recognised changes in the social and economic environment and proposed labour law reforms, no changes were ultimately adopted. Where resources are concerned, our evidentiary standards seem to be met in Hacker and Pierson's elaboration. The mobilisation of business interests is a key part of Hacker and Pierson's story. They vividly detail how the manoeuvres of the newly mobilised business community – flush with resources and outfitted by conservative think tanks – mobilised against reforms that would have enhanced the National Labour Relations Board's ability to penalise corporations that violated labour laws, outspending labour 3-to-1 on lobbying. To make this case, Hacker and Pierson present evidence from participants in policy-making discussions, one of whom claims that the approach of the business community had become more aggressive. As one congressional aide claims, "I don't think [business] missed a single possible opponent of that bill in our state" (p. 131).

Finally, we must consider the institutional outcome that accompanies drift as a true change in institutional performance compared to the

pre-drift period (as suggested in the fourth row of Table 1). Shifts in social or economic behaviour should, in other words, be traceable to the absence of policy change and difficult to imagine in a counterfactual scenario in which that policy had been adopted. Although in Hacker and Pierson's narrative of labour relations policy, the social and economic outcome in question – a decline in union representation in the workforce – is certainly a partial result of the absence of legislative change in the United States, it is also hard to discount the contribution of a weak protective labour policy *ex ante*, as well as the prevalence of state-level right-to-work statutes (pp. 56–61). However, Hacker and Pierson present no confirmatory evidence that the absence of new labour relations laws is the most important contributor to the outcome of declining union representation that they see, even if they do present compelling evidence that decreasing unionisation cannot be explained by economic changes alone (pp. 60–61). Though assessing the effect of non-events is difficult in a statistical sense, a more compelling account of drift would include, at the very least, a consideration of a counterfactual scenario in which labour relations policies that were updated would produce a more compelling portrait of drift (George and Bennett 2005, 181–204).

Discussion

The analysis of the above empirical work provides us with four insights into the quality of our approach to operationalising gradual-change concepts. Our first finding is that the empirical research here hews fairly close to the empirical standards we have set forth. For instance, despite the authors' uncertainty as to the initial ambiguity of the National Labour Relations Act, there is little doubt that what Hacker and Pierson are describing is drift and *not* conversion, layering or even displacement (which we have excised from our discussion here). Similarly, the patterns Thelen identifies with processes of layering and conversion are each distinctive, although elements of drift do appear in the layering narrative. More broadly, all of the forms of institutional change here can be described as *gradual*: they occur over long swaths of time as the result of initial institutional structures, the actions of change agents or the combination of these factors with new exogenous pressures. Thus, we can conclude that the general process characteristics developed by the GCF can operate as analytical heuristics that help to simplify complex social phenomena. However, the inconsistencies we found force us to suggest that applying these heuristics demands clear operational strategies and creative thinking about how to deploy research techniques to meet what seem like challenging evidentiary standards.

Second, we find a few inconsistencies between the empirical work and our indicators. For instance, as has been clear throughout each of the case studies, initial institutional ambiguity is difficult to identify, not only because ambiguity itself is difficult to measure, but also because, in complex environments, it is sometimes difficult to identify which institution is the relevant one. This is a problem that we believe can be rectified by using techniques to better identify institutional ambiguity, which we identified above, and by the empirical researchers' creation of an inventory of initial institutional designs to which eventual institutional change can be traced. Here, we think the technique of process tracing discussed by George and Bennett (2005) may be helpful. Again, however, only by developing clear empirical standards against which to test claims of gradual institutional change can we better differentiate between processes like drift, conversion and layering.

Third, with respect to the inconsistencies we recovered in the identification of change-agent resources, finding evidence on these points may be somewhat more difficult than researchers realise initially. Discovering evidence of institutional meanings and change-agent resources, for example, may require delving into private documents rather than public statements and, even then, may require the additional development of indicators that should be found in these statements if these hypotheses are true. As such, the task of developing or adopting techniques of actor-centred history is one that researchers of gradual institutional change should undertake in the future (Bensel 2004; Canaday 2009). With respect to resources, finding evidence may be somewhat less challenging, but no less time-consuming, especially if discretionary, material and intellectual resources are to be investigated. This may include probing more archival sources as well as publicly available documents for evidence of each of these forms of resources.

Finally, it is worth highlighting that change processes discussed in the GCF can be the result of structures, agents or both. Though substantial analytical problems seem not to have resulted from the absence of clarity on this point, we would note that, as the complexity of causal narratives increases, the absence of strong accounting procedures could become highly problematic. Since differences in the importance of structure and agency may create different empirical indicators of change, it is imperative that, when writing up research, analysts specify the relative importance of each of these factors to their overall causal narrative.

Conclusion

Our framework advances discussion of gradual-change processes – which presently focuses on conceptualisation – to a conversation that includes

issues of observation and measurement.[6] A research agenda that oper-
ationalises concepts of gradual institutional change, we argue, has practical
applications for empirical analysis in terms of description, explanation
and the mundane practice of research.

By deploying a set of clear observable indicators for processes of
institutional change, scholars can begin to see how common processes
operate across different institutional contexts. For example, using the
observable indicators named here, we may be able to find that drift is very
common at some levels of institutional analysis (public policy, for
instance), but less common in others (political parties or legislatures).
In addition to providing the tools for discovering where processes of
institutional change tend to take place, our framework (and other frame-
works that consider the observable implications of processes in the GCF)
allows scholars to better adjudicate between competing interpretations of
data and to falsify claims about those processes. Relatedly, our method of
classification permits us to be transparent about the degree of uncertainty
involved in our choice of how to classify theoretical processes, making it
possible to see where remaining evidentiary gaps lie. As our discussion of
the gaps in the narratives of institutional change by Thelen and Hacker
and Pierson illustrates, our approach can allow qualitative scholars to be
honest about the extent to which their classification of a given process
of change is certain. Our framework may also allow scholars to detail
scenarios in which multiple processes of change are at work in a given
institution, a possibility not contemplated by the GCF itself.[7] By attending
to qualitative indicators of processes and explicitly addressing whether or
not evidence meets the benchmarks we have laid out, qualitative scholars

---

[6] We do not wish to challenge the usefulness of gradual-change metaphors as heuristics for
identifying previously unnoticed processes of gradual change, or for drawing attention to
features of institutions (e.g. ambiguity and agency) that empiricists had previously overlooked
in favour of smoking guns. Moreover, we do not argue that our set of distinguishing criteria or
confirmatory evidence for processes of institutional change is definitive, as we feel there is much
room for debate and discussion of the relative importance of these criteria. We certainly
recognise the necessity for flexibility when describing change, as processes cannot and should
not fit perfectly into these categories all of the time. Furthermore, while our focus is on public
policies, there is good reason to believe that different processes may have different indicators
depending on the type of institutions under consideration. Understood as institutions, political
parties, courts and public policy regimes may exhibit significantly different observables for
similar processes of change. What our analysis underlines, however, is the utility of considering
the observable implications of the processes that have been advanced by the GCF.

[7] While the theoretical work on gradual institutional change tacitly emphasises an "either/
or" approach to labelling cases of gradual institutional change, empirical scholarship recog-
nises the possibility of multiple processes of institutional change occurring either simulta-
neously or within the same time horizon (see Thelen 2004; Jacobs 2010).

can make rigorous, valid arguments that expose the political and social forces behind institutional change.

Our framework also provides an alternative to quantitative work on institutional change that operationalises key concepts with data on shifts in policy budgets. Though budgets for public policies are an important indicator of changes in government's goal commitments, studying budgets as indicators of institutional change can obscure alterations in the content of policy administration (Baumgartner 2013). What may look like a stable budget may actually, for instance, be failing to meet new demands in the social or economic context of public policy. As Hacker (2004) suggests, studying changes in the budget of one programme alone may even prevent us from seeing some forms of institutional change like conversion in which change occurs through alterations in the meaning of a piece of policy. Our approach to operationalising forms of institutional change thus adds a necessary complement to quantitative studies by providing a rigorous technique to qualitatively identify different types of institutional change as they occur. This may add to the fruitful exchange of ideas across quantitative and qualitative institutional change scholarship (Baumgartner 2013).

In addition, our framework provides a crucial tool for scholars interested in theory testing. Identifying causal relationships between political variables and institutional change – for instance, Mahoney and Thelen's conjecture about the importance of the durability of the *status quo ante* in contributing to different processes of change – relies heavily on our ability to properly identify such processes. Our framework and others like it, we argue, can move the GCF towards this goal by allowing scholars interested in describing processes of institutional change to validate their claims against evidentiary standards and to identify the level of analytical uncertainty in these claims. Our approach strengthens explanations and arguments by holding them to appropriate standards of evidence commensurate with qualitative causal inference rather than forcing historical research into the box of quantitative methods (King et al. 1994). The theoretical project of the GCF – to identify and verify the conditions under which gradual processes of institutional change occur – is still in its embryonic stage. By clearly outlining observable indicators of the invisible social processes the GCF identifies, we hope to allow this theoretical project to flourish.

Finally, our framework can also provide a useful solution to several practical problems of historically informed research in the social sciences. Many, if not most, major contributions to this type of research have involved not just novel interpretations of historical phenomena, but have also used original archival work to uncover narratives missed by the

secondary literature. However, as any historical researcher knows, going into the archives can be daunting without a clear set of *ex ante* standards for the kinds of confirmatory evidence one needs. Though scholars often claim that some of this knowledge can only be developed by "soaking and poking" in the archive itself (George and Bennett 2005), our framework provides a time-saving field guide to the kinds of evidentiary demands set out by the GCF. Having a sense of these demands before conducting fieldwork, we argue, could prove useful "on the ground".

Though ours is certainly not the only reasonable elaboration of the GCF's concepts, and it is certainly true that operationalisation strategies may vary conditionally on the domain of empirical research, the framework advanced in this article is a step in the direction of a stronger programme of empirical research on gradual institutional change in political science. Our efforts at outlining the empirical indicators are thus part of a larger (and still incomplete) process of elaborating the GCF in institutional theory. The analytical approach we outline here stands to match the GCF's theoretical promise with the empirical tools necessary to test, extend, and if necessary, revise the framework's core claims.

Acknowledgements

References

Adcock R. and Collier D. (2001) Measurement Validity: A Shared Standard for Qualitative and Quantitative Research. *American Political Science Review* 95(3): 529–546.

Baumgartner F. R. (2013) Ideas and Policy Change. *Governance* 26 (April): 239–258.

Béland D. (2007) Ideas and Institutional Change in Social Security: Conversion, Layering, and Policy Drift. *Social Science Quarterly* 88(1): 20–38.

Bensel R. (2004) *The American Ballot-Box in the Nineteenth Century*. New York, NY: Cambridge University Press.

Blyth M. (2002) *Great Transformations: Economic Ideas and Institutional Change*. Cambridge, UK: Cambridge University Press.

Butler S. and Germanis P. (1983) Achieving Social Security Reform: A "Leninist" Strategy. *Cato Journal* 3(2): 547–556.

Campbell J. (2004) *Institutional Change and Globalization*. Princeton, NJ: Princeton University Press.

Canaday M. (2009) *The Straight State: Sexuality and Citizenship in Twentieth-Century America*. Princeton, NJ: Princeton University Press.

Carpenter D. (2010) Institutional Strangulation: Bureaucratic Politics and Financial Reform in the Obama Administration. *Perspectives on Politics* 8 (September): 825–846.

Chou M. (2012) Constructing an Internal Market for Research Through Sectoral and Lateral Strategies: Layering, the European Commission, and the Fifth Freedom. *Journal of European Public Policy* 19(7): 1052–1070.

Collier D., Brady H. and Seawright J. (2011) Outdated Views on Qualitative Methods: Time to Move On. *Political Analysis* 18: 506–513.

Cox G. and McCubbins M. (2001) The Institutional Determinants of Policy Outcomes. In Haggard S. and McCubbins M. (eds.), *Presidents, Parliaments, and Policy*. New York, NY: Cambridge University Press, 21–63.

Dodds A. and Kodate N. (2012) Understanding Institutional Conversion: The Case of the National Reporting and Learning System. *Journal of Public Policy* 32(2): 117–139.

Elster J. (1989) *Nuts and Bolts for the Social Sciences*. New York, NY: Cambridge University Press.

Espeland W. (1998) *The Struggle for Water: Politics, Rationality, and Identity in the American Southwest*. Chicago, IL: University of Chicago Press.

Esping-Andersen G. (1990) *The Three Worlds of Welfare Capitalism*. Princeton, NJ: Princeton University Press.

Falleti T. and Lynch J. (2009) Context and Causal Mechanisms in Political Analysis. *Comparative Political Studies* 42(9): 1143–1166.

George A. and Bennett A. (2005) *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.

Goldin I. and Vogel T. (2010) Global Governance and Systemic Risk in the 21st Century: Lessons from the Financial Crisis. *Global Policy* 1(1): 4–15.

Gross J. (2010) *Broken Promises: The Subversion of U.S. Labor Relations, 1947–1994*. Philadelphia, PA: Temple University Press.

Hacker J. (2002) *The Divided Welfare State: The Battle Over Public and Private Social Benefits in the United States*. New York, NY: Cambridge University Press.

—— (2004) Privatizing Risk Without Privatizing the Welfare State: The Hidden Politics of Social Policy Retrenchment in the United States. *American Political Science Review* 98(2): 243–260.

—— (2005) Policy Drift: The Hidden Politics of US Welfare State Retrenchment. In Streeck W. and Thelen K. (eds.), *Beyond Continuity: Institutional Change in Advanced Political Economies*. Oxford: Oxford University Press, 40–82.

Hacker J. and Pierson P. (2010a) *Winner-Take-All Politics: How Washington Made the Rich Richer—And Turned its Back on the Middle Class*. New York, NY: Simon & Schuster.

—— (2010b). Drift and Democracy. Paper presented at the 2010 Annual Meeting of the American Political Science Association, Washington, D.C., September 2–5, 2013.

Henig J. (2008) Education Policy from 1980 to the Present: The Politics of Privatization. In Glenn B. and Teles S. (eds.), *Conservatism and American Political Development*. New York, NY: Oxford University Press, 291–323.

Howell C. (1992) The Contradictions of French Industrial Relations Reform. *Comparative Politics* 24 (January): 181–197.

Immergut E. (1992) *Health Politics: Interests and Institutions in Western Europe*. New York, NY: Cambridge University Press.

Immergut E. and Anderson K. (2008) Historical Institutionalism and West European Politics. *West European Politics* 31(1–2): 345–369.

Jacobs A. (2010) Policymaking as Political Constraint: Institutional Development in the U.S. Social Security Program. In Mahoney J. and Thelen K. (eds.), *Explaining Institutional Change: Ambiguity, Agency and Power*. New York, NY: Cambridge University Press, 94–131.

King G., Keohane R. and Verba S. (1994) *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press.

Levy J. (2005) Redeploying the State: Liberalization and Social Policy in France. In Streeck W. and Thelen K. (eds.), *Beyond Continuity: Institutional Change in Advanced Political Economies*. New York, NY: Oxford University Press, 103–126.

Mahoney J. and Thelen K. (2010) A Theory of Gradual Institutional Change. In Mahoney J. and Thelen K. (eds.), *Explaining Institutional Change: Ambiguity, Agency and Power*. New York, NY: Cambridge University Press, 1–37.

Marmor T. (2000) *The Politics of Medicare*, 2nd ed. New Brunswick, NJ: Transaction Publishers.

Mayhew D. (2002) *America's Congress: Actions in the Public Sphere, James Madison Through Newt Gingrich*. New Haven, CT: Yale University Press.

North D. (1990) *Institutions, Institutional Change and Economic Performance*. Cambridge, UK: Cambridge University Press.

Palier B. (2005) Ambiguous Agreement, Cumulative Change: French Social Policy in the 1990s. In Thelen K. and Streeck W. (eds.), *Beyond Continuity, Institutional Change in Advanced Political Economies*. Oxford: Oxford University Press, 127–144.

Patashnik E. (2008) *Reforms at Risk: What Happens After Major Policy Changes are Enacted*. Princeton, NJ: Princeton University Press.

Pierson P. (2004) *Politics in Time: History, Institutions, and Social Analysis*. Princeton, NJ: Princeton University Press.

—— (2005) Public Policies as Institutions. In Shapiro I., Skowronek S. and Galvin D. (eds.), *Rethinking Political Institutions: The Art of the State*. New York, NY: New York University Press, 114–134.

Pressman J. and Wildavsky A. (1973) *Implementation: How Great Expectations in Washington are Dashed in Oakland; OR, Why It's Amazing that Federal Programs Work at All*. Berkeley, CA: University of California Press.

Rothstein B. (1998) *Just Institutions Matter: The Moral and Political Logic of the Universal Welfare State*. New York, NY: Cambridge University Press.

Salines M., Glöckler G. and Truchlewski Z. (2012) Existential Crisis, Incremental Response: The Eurozone's Dual Institutional Evolution 2007–2011. *Journal of European Public Policy* 19(5): 665–681.

Selznick P. (1949) *TVA and the Grass Roots*. Berkeley, CA: University of California Press.

Schickler E. (2001) *Disjointed Pluralism: Institutional Innovation and the Development of the U.S. Congress*. Princeton, NJ: Princeton University Press.

Sheingate A. (2010) Rethinking Rules: Creativity and Constraint in the House of Representatives. In Mahoney J. and Thelen K. (eds.), *Explaining Institutional Change: Ambiguity, Agency and Power*. New York, NY: Cambridge University Press, 168–203.

Soss J., Fording R. and Schram S. (2011) *Disciplining the Poor: Neoliberal Paternalism and the Persistent Power of Race*. Chicago, IL: University of Chicago Press.

Streeck W. and Thelen K. (eds.) (2005) *Beyond Continuity: Institutional Change in Advanced Political Economies*. New York, NY: Oxford University Press.

Teles S. (2008) *The Rise of the Conservative Legal Movement: The Battle for Control of the Law*. Princeton, NJ: Princeton University Press.

Thelen K. (2004) *How Institutions Evolve: The Political Economy of Skills in Germany, Britain, the United States and Japan*. Cambridge, MA: Cambridge University Press.

Tilly C. (2004) Observations of Social Processes and their Formal Representations. *Sociological Theory* 22(4): 595–602.

Tsebelis G. (2002) *Veto Players: How Political Institutions Work*. Princeton, NJ: Princeton University Press.

Weaver K. (2000) *Ending Welfare as we Know it*. Washington, DC: The Brookings Institution.

Appendix

Table A1. A rough guide to mechanisms of gradual institutional change

Mechanism	Definition	Canonical Citations
Conversion	"The changed impact of existing rules due to their strategic redeployment"	Thelen (2004), Selznick (1949)
Drift	"The changed impact of existing rules due to changes in the environment"	Hacker (2005), Hacker and Pierson (2010a, 2010b)
Displacement	"The removal of new existing rules and the introduction of new ones"	Weaver (2000), Patashnik (2008)
Layering	"The introduction of new rules on top of or alongside existing ones"	Schickler (2001), Thelen (2004)

Source: Mahoney and Thelen (2010, 15–16).

Table A2. Types of gradual change

	Displacement	Layering	Drift	Conversion
Removal of old rules	Yes	No	No	No
Neglect of old rules	–	No	Yes	No
Changed impact/enactment of old rules	–	No	Yes	Yes
Introduction of new rules	Yes	Yes	No	No

Source: Mahoney and Thelen (2010, 16).

Table A3. Sources of institutional change

	Characteristics of Targeted Institution	
Characteristics of the Political Context	Low Level of Discretion in Interpretation/Enforcement	High Level of Discretion in Interpretation/Enforcement
Strong veto possibilities	Layering	Draft
Weak veto possibilities	Displacement	Conversion

Source: Mahoney and Thelen (2010, 19).