HISTORICAL INSTITUTIONALISM IN COMPARATIVE POLITICS

Kathleen Thelen
Department of Political Science, Northwestern University, Evanston, Illinois 60208; e-mail: thelen@nwu.edu

KEY WORDS: path dependency, rational choice, institutional change, critical junctures, policy feedbacks

ABSTRACT

This article provides an overview of recent developments in historical institutionalism. First, it reviews some distinctions that are commonly drawn between the “historical” and the “rational choice” variants of institutionalism and shows that there are more points of tangency than typically assumed. However, differences remain in how scholars in the two traditions approach empirical problems. The contrast of rational choice’s emphasis on institutions as coordination mechanisms that generate or sustain equilibria versus historical institutionalism’s emphasis on how institutions emerge from and are embedded in concrete temporal processes serves as the foundation for the second half of the essay, which assesses our progress in understanding institutional formation and change. Drawing on insights from recent historical institutional work on “critical junctures” and on “policy feedbacks,” the article proposes a way of thinking about institutional evolution and path dependency that provides an alternative to equilibrium and other approaches that separate the analysis of institutional stability from that of institutional change.

INTRODUCTION

Institutional analysis has a distinguished pedigree in comparative politics, and the “new” institutionalist literature of the past two decades has both sustained this venerable tradition and deepened our understanding of the role of institutions in political life. At the same time, recent work has given rise to new debates. It is now conventional to distinguish three different varieties of institutionalism: rational choice institutionalism, historical institutionalism, and socio-
logical institutionalism. Each of these three schools in fact represents a sprawling literature characterized by tremendous internal diversity, and it is often also difficult to draw hard and fast lines between them. The differences that have been identified amount to tendencies that apply unevenly across particular authors within each school of thought (Hall & Taylor 1996). The walls dividing the three perspectives have also been eroded by “border crossers” who have resisted the tendencies toward cordonning these schools off from each other and who borrow liberally (and often fruitfully) where they can, in order to answer specific empirical questions. A few examples will suffice to illustrate this point.

First, a group of prominent theorists working out of a rational choice perspective have become proponents of a more eclectic approach that combines elements of deductive theory—the hallmark of rational choice—with an explicit attempt to contextualize the analysis in ways that historical institutionalists have long advocated (Bates et al 1998b). This strategy, which they call analytic narratives, represents an attempt to construct explanations of empirical events through analyses that “respect the specifics of time and place but within a framework that both disciplines the detail and appropriates it for purposes that transcend the particular story” (Levi 1999). The analyses offered thus incorporate elements of deduction and induction in ways that overcome traditional distinctions between historical institutionalism’s characteristic focus on specific contextual conditions and rational choice’s characteristic search for generalizable features of political behavior rooted in the incentive structures that individuals face.

A second illustration of border crossing is the equally important impact of rational choice on the work of many historical institutionalists. One development has been an enhanced appreciation of the need for explanations that rest on firm micro foundations. Although much macro-historical work was already implicitly sensitive to these issues, articulating the micro-foundational logic of the arguments offered was not always a top priority, and recent work reflects a heightened appreciation for such issues. Many historical institutionalists have also taken on board the notion that institutions that solve collective action problems are particularly important in understanding political outcomes (Rothstein 1996:159). This has long been a central concern of rational choice theory, and it has set an important agenda for historical institutionalists as well. Thus, an increasing number of historical studies focus precisely on explaining the emergence and persistence of institutions that do (or do not) facilitate coordination among employers and other groups (see e.g. Hall 1994, Thelen & Kume 1999).

Third, there has been some important borrowing and cross-fertilization between historical institutionalism and sociological institutionalism. The works that lie at this intersection often embrace a more expansive view of institutions, not just as strategic context but as a set of shared understandings that affect the way problems are perceived and solutions are sought. Katzenstein’s analysis of the evolution of Japanese security policy, for example, shows how collectively held norms define appropriate conduct, shape actor identities, and influence actor interests (1996:23), and in doing so, “inform how political actors define what they want to accomplish” (1996:ix). Katzenstein roots his cultural approach in a political analysis of how some norms (and not others) came to be institutionalized, and so his perspective resonates especially with those versions of institutional sociology that specifically incorporate considerations of power and/or legitimacy in explaining how institutions emerge and are reproduced (e.g. Fligstein 1991, DiMaggio 1988, Stinchcombe 1997). In sum, there have been some rather fruitful developments at the intersection of these various schools of institutionalism, and in my view historical institutionalism has been enriched by encounters with alternative perspectives.

This article provides an overview of recent developments within the historical institutional tradition. Given the vast scholarship in this area, this tour is necessarily selective. I attempt to capture the current state of the literature in two passes. First, I review some distinctions that are commonly drawn between the “historical” and the “rational” choice variants of institutionalism. This exercise reveals that there are more points of tangency than commonly assumed. However, differences remain, and here I focus on the difference between rational choice’s emphasis on the coordinating functions of institutions (generating or maintaining equilibria) versus historical institutionalism’s emphasis on how institutions emerge from and are embedded in concrete temporal processes. This discussion serves as the foundation for the second half of the essay, in which I revisit one of the key frontiers in historical institutionalism (identified in Thelen & Steinmo 1992) and assess the progress that has been made in our understanding of institutional formation and change. Taking up an important challenge by Orren & Skowronek (1994), and drawing on insights from recent historical institutional work on critical junctures and on policy feedbacks, I propose a way of thinking about institutional evolution and path dependency that provides an alternative to equilibrium and other approaches that separate the analysis of institutional stability from that of institutional change.

2Some of this literature also relates to ongoing work by rational choice theorists that emphasizes the role of culture in defining “focal points” that influence which of a number of possible equilibria is actually achieved (Ferejohn 1991, Greif 1994, Bates et al 1998a).
RED HERRINGS AND REAL ISSUES

Theoretical Versus Empirical Work

One of the lines that is frequently drawn between historical institutionalism and rational choice institutionalism is between “theoretical” and “empirical” work. In a well-known critique, Green & Shapiro (1994) charge that rational choice has produced elegant theories but has generated little to explain real observed events. From the other side, rational choice theorists have often argued that historical institutionalists are engaged in something less than theory building; they are stringing details together, “merely telling stories.” Even where the distinction is not drawn so starkly, the assertion is that the difference is fundamental. Thus, for example, Levi argues (in opposition presumably to historical institutionalism) that rational choice is “almost always willing to sacrifice nuance for generalizability [and] detail for logic” (1997a:21).

This dichotomy is often exaggerated, and in my view it is misplaced, for the best work in both perspectives is concerned with generating hypotheses that are then brought to bear on empirical phenomena. For example, Luebbert’s (1991) analysis of the origins of fascism, social democracy, and liberal democracy is based on a comparative analysis of class relations in the interwar period in Europe. Combining comparative method with close historical process tracing in individual cases, Luebbert singles out as the crucial explanatory variable the issue of how the landed peasantry was mobilized politically. Where working-class–based parties allied themselves with the landed peasantry, this produced the mass base necessary for the establishment of social democratic regimes. By contrast, where social democrats failed to forge this alliance, the landed peasantry turned against the working class and provided the mass base on which fascism grew. One can disagree with his findings, but Luebbert’s work is an exemplary model of the testing, through the comparative method, of a strong and clear hypothesis that is obviously capable of being falsified.

Another example is Collier & Collier’s (1991) impressive study of regime transformation in Latin America. Based on structured comparisons and historical process tracing for individual cases, they found that differences in patterns of labor incorporation were key to the type of regime that emerged. Like Luebbert’s work, this study was not meant to capture every detail; on the contrary, Collier & Collier’s account is precisely designed to probe the plausibility of alternative hypotheses. Much is thus lost in terms of a comprehensive history of each country, but the payoff is a set of general propositions about the way in which labor incorporation affected subsequent regime outcomes across a number of countries.

Many other works could be invoked that use the comparative historical method to sort out the causal mechanisms behind observed empirical patterns (see the discussion, below, of the critical junctures literature). All of them “go
beyond conventional history’s preoccupation with historical particularity and aim for theoretical generalization” (Rueschemeyer et al 1992:4). In this, they share much with rational choice analyses as characterized by Levi, sacrificing much detail in order to identify general causal patterns that hold across a number of countries.

Conversely, rational choice work in comparative politics arguably has become more empirical over time. For example, the authors of the “analytic narratives” project mentioned above specifically emphasize that the papers they present are “problem driven, not theory driven; they are motivated by a desire to account for particular events or outcomes. They are devoted to the exploration of cases, not to the elaboration of theory...” (Bates et al 1998b:11). Some of the analyses are comparative, others focus on single cases; however, like good single case studies by historical institutionalists, the latter use close analysis of critical cases to illuminate important general issues.

Although the difference between the historical and rational choice variants of institutionalism cannot be summed up accurately in a strict dichotomy of “empirically” versus “theoretically” oriented work, there do appear to be some differences between the two traditions’ approaches to theory building. First, most historical institutionalists are working at the level of mid-range theory of the sort that Bendix and others have advocated. Often, though not always (e.g. Rueschemeyer et al 1992, Karl 1997), this involves focusing on a limited range of cases that are unified in space and/or time, for example, explaining transitions to democracy in Eastern Europe (Stark & Bruszt 1998) or the effects of new international pressures on the political economies of the advanced industrial democracies (Hall 1994). Rational choice theorists, by contrast, sometimes aspire to produce more general (even universal) theoretical claims3 or use historical examples not so much for their intrinsic importance (e.g. as “critical cases”) but to demonstrate how widely applicable are the theoretical claims (e.g. Knight 1992, Levi 1988, Tsebelis 1990). Still, the object of both is to test theoretical propositions against observed phenomena, in order not only to explain the cases at hand but also to refine the theory.

A second difference with respect to theory building lies in the ways that historical institutionalists and rational choice institutionalists approach hypothesis formation. Very frequently, historical institutionalists begin with empirical puzzles that emerge from observed events or comparisons: Why did the policies of the advanced industrial countries differ so much in response to the oil shock of 1973 (Katzenstein 1978)? Why have some industrial relations sys-

---

3This is not always the case, and in fact the tendency may be weaker among comparativists. For example, those who study the political economy of the advanced industrial countries (e.g. Scharpf and Soskice) tend to formulate their hypotheses and conclusions at the same, mid-range, level as most historical institutionalists.
tems proved more stable than others in the face of globalization pressures (Thelen 1993, 1994)? Why do some countries tax and spend more than others (Steinmo 1993)? The analyst then uses the comparisons to test hypotheses that can account for the observed differences. Rational choice theorists often proceed somewhat differently, deriving their puzzles from situations in which observed behavior appears to deviate from what the general theory predicts: Why, given free rider problems, do workers join unions (Wallerstein 1989)? Why would unions lead workers into hopeless battles (Golden 1997)? Why would citizens ever volunteer for war (Levi 1997b)? This difference in the analytic point of departure accounts for the extensive use of counterfactuals in rational choice research (e.g. Bates et al 1998a), which function much as cross-national comparisons often do in historical institutional research. The question is, what is “off the equilibrium path” (Levi 1997a:31)? The theory gives us prior expectations that we can then hold against the observed outcomes (see also Scharpf 1997:29).

The point is that it is not the case that one perspective’s analysis is guided by clear hypotheses and the other’s is not. The issue is where these hypotheses come from. Moreover, the difference in sources of hypotheses is not a hard and fast rule. As Scharpf points out, rational choice theories seek “to explicate what the authors of ‘good’ case studies always have in the back of their minds: a ‘framework’ that organizes our prior (scientific and prescientific) knowledge about what to expect in the province of the world that is of interest to us, that emphasizes the questions that are worthwhile asking, the factors that are likely to have high explanatory potential, and the type of data that would generally be useful in supporting or invalidating specific explanations” (1997:29–30).

Moreover, there is much overlap between the two perspectives when it comes to testing hypotheses against empirical cases, for this invariably involves contextualizing the theory (assumptions and propositions) and demonstrating that the hypothesized processes are actually at work (Pierson 1996:158). As Scharpf puts it, “We need to have hypotheses that specify a causal model showing why and how a given constellation of factors could bring about the effect in question,” but equally, “we need to have empirical evidence that the effect predicted by the hypothesis is in fact being produced” (1997:28). There is, in other words, no dichotomy between theoretical and empirical work because good analyses have to be both. The generation of the hypotheses is not the analysis, although it is the vital starting point for engaging the empirical material. The utility of a theory, after all, cannot be assessed apart from the empirical material it is meant to explain.

Preferences: Problematical or Not

In 1992, Steinmo and I argued that “one, perhaps the core, difference between rational choice institutionalism and historical institutionalism lies in the
question of preference formation, whether treated as exogenous (rational choice) or endogenous (historical institutionalism)” (Thelen & Steinmo 1992:9, emphasis in original). I now think this may be a less stark difference than before. Setting aside some important ambiguities on this issue in both the rational choice and the historical institutionalist literature, one of the core claims of historical institutionalism from the beginning was that institutions do more than channel policy and structure political conflict; rather, “the definition of interests and objectives is created in institutional contexts and is not separable from them” (Zysman 1994:244). In almost any version this is quite different from strong versions of rational choice theory, which begin with a (universal, not context-specific) rationality assumption.

Or is it? As Levi has suggested, in rational choice “the trick is in defining the preferences in general, ex ante to a particular application” (1997a:24). She notes that this is frequently difficult; for example, in the case of citizens, “given the range of interests [they] might have; there is nothing comparable to the economics dictum of getting the most for the least for one’s money in the marketplace” (1997a:24). In a candid discussion of the problem of imputing preferences, she notes the danger that assuming utility maximization “can...produce tautology: Whatever people do becomes a ‘revealed preference’” (1997a:24).6

Even the assumptions traditionally considered “safe” may be trickier than we thought, the most common perhaps being that politicians are seeking reelection. This cannot be valid for Mexico, where the rules of the game (read: the specific institutional configuration) rule out, by law, reelection of legisla-

---

4Within the rational choice literature on norms, for example, there is some ambiguity as to whether norms affect belief formation or preference formation. “Beliefs” simply refer to how you think others will behave, whereas “norms” are collectively shared convictions (see Ferejohn 1991; Bates et al 1998a; Levi 1997b, 1998). (I am indebted to Guillermo Trejo for pointing this out to me.) Among historical institutionalists, Immergut (1997) argues that the claim that institutions shape preferences has often conflated several analytically separate issues, and she prefers “to distinguish more sharply between preferences, interests and choices” and to focus on interests as “publicly expressed, organized demands” rather than preferences in the sense of the views of individuals. She argues that “most institutionalists focus on how institutions may foster the emergence of particular definitions of mutual interest, or advantage particular political choices, without necessarily re-socializing citizens in a fundamental way” (1997:339–40).

5In economics, theories of the firm have been more successful than theories of consumption for this very reason. The assumption that firms seek to maximize profits is much more tangible, concrete, and therefore useful, than the assumption that consumers seek to maximize “utility” (whatever that might be). The quality of the theory appears to be a direct function of the degree of specificity of the core assumptions (G Trejo, personal communication). See also the discussion in Satz & Ferejohn 1994, esp. p. 72.

6See also Kuran (1991), who distinguishes between the private and public preferences of individuals and shows that people may in fact falsify their preferences for pragmatic reasons, making it impossible to know whether their behavior reflects their true preferences or is simply strategic. See also Rothstein (1996:148).
tors. (The same is true for presidents across a much larger number of countries.) In such cases, an analysis that starts with the reelection-seeking assumption will be widely off the mark. How do rational choice analysts deal with this situation? By and large they try to figure out, within a given context, what would make sense for a politician to seek, which is to say that they contextualize the preferences based on the particular institutional incentives that these politicians face in different settings. This is not so different from how historical institutionalists often proceed. The point is that the move from general propositions about what political actors are seeking to maximize inevitably brings the theorist face to face with the question of what it means to, say, maximize power within a given context. Until this step is complete, the analysis cannot begin. This seems to be what Bates et al (1998a) mean when they note that the “cultural’ knowledge required to complete a rational choice explanation reveals the complementarity of the [‘rational’ and ‘interpretive’] approaches. Game theorists often fail to acknowledge that their approach requires a complete political anthropology…. Game-theoretic accounts require detailed and fine-grained knowledge of the precise features of the political and social environment within which individuals make choices and devise political strategies” (1998a:628).

In addition, norms and culture, which for a long time were of concern mainly to historical institutionalists and institutional sociologists, appear to be assuming an increasingly important role in rational choice analysis. Ferejohn’s recent work, for example, argues that “culturally shared understandings and meanings” are crucial to selecting among the many possible strategic equilibria (1991:285). He argues that

in social action, human agents make strategic or allocative choices while simultaneously enacting (ontologically) prior understandings about the nature of the strategic situation in which they find themselves, the characteristics or identities of the players (including themselves), and the common understandings or expectations as to how the game will be played. Thus, when it comes to explaining action, rational accounts, no less than interpretive ones, must appeal to principles external to the individual agents. (Ferejohn 1991: 285)

7 A revision of the reelection assumption might say that politicians are seeking to maximize their careers or trying to increase their power, but again, what that means depends on the context—among other things, where power is located (institutionally) in different systems. For example, given the historical weakness of Congress in Latin America, “maximizing power” for a legislator there may mean angling for an appointive job in the bureaucracy. This is why rational choice theory works best in highly structured settings in which the players and the rules are stable and well known (Satz & Ferejohn 1994:72,81; Bates 1997:704; Bates et al 1998:222), in other words, where the structure itself does some of the analytical work.

8 Actually, this interest goes back to Schelling (1960), and Elster (1989) too has long been concerned with norms and culture. More recently they have been joined by a host of others, including Ferejohn (1991), Bates et al (1998a), Levi (1998, 1999), North (1990), and Greif (1994). See also APSA-CP (1997), especially the contributions by Bates, Johnson, Laitin, Rogowski.
Many rational choice theorists follow North in his argument that norms constitute informal institutions—another set of rules that create incentives or constraints on behavior. However, theorists differ in explaining how norms and culture fit into the analysis of political outcomes. Some authors see cultural symbols and norms as resources invoked in strategic interactions (e.g. Johnson 1997); others view them as signaling devices in games of incomplete information (e.g. Bates et al 1998a); still others as focal points that affect which of a number of possible equilibria prevails (Rogowski 1997, Greif 1994). Here it is sufficient to note that in order to operate in any of these ways, norms must exert some independent power over individual behavior, and in this sense most of these works appear to go well beyond traditional “instrumental rationality” assumptions. In light of such developments, the issue of preferences (exogenous/endogenous) no longer seems to provide a clear line of demarcation between the different approaches.

Micro-Foundational Versus Macro-Historical Research

The issue of micro foundations is a third one that is typically cited as distinguishing historical institutional research from rational choice. The idea is that aggregate outcomes need to be understood in terms of the actions and behavior of individuals behaving strategically. This is frequently contrasted to broad macro-historical research that either sees interests as structurally generated in one way or another or stays at the level of aggregations (such as class) without regard for the way in which the strategic actions of individuals figure into the aggregation process.

But this too is a false dichotomy. The issue of the embeddedness of interests has been discussed above; however, there is the additional question of how to think about aggregate behavior. “Micro-foundational” obviously does not preclude dealing with collectivities, since most rational choice work deals with collective actors.9 The analysis remains, nonetheless, “actor-centered” in the sense that the players are defined as “any individual or composite actor that is assumed to be capable of making purposeful choices among alternative courses of action” (Scharpf 1997:7). This may be unproblematical, as for example in Scharpf’s case, since unions (or at least the union leadership) and not individual workers are the relevant players who actually engage in the strategic bargaining that generates the policy outcomes of interest. In other cases, however, it is worthwhile questioning whether the collectivities to which stra-

---

9Scharpf’s (1997) work on economic policy outcomes, for example, examines the interaction of unions, employers, and the government. Golden (1997) explains apparently suicidal struggles on the part of unions by analyzing the behavior of union leaders, employers, and the rank and file. Weingast (1998) examines the incentives faced by Southern states in their interaction with the North. In all these cases, the relevant players are aggregations of individuals, that is to say, collectivities.
strategic action is attributed in fact constitute players in the sense that Scharpf identifies.

Take the example of Rogowski’s (1987) analysis of coalitional alignments and realignments in the late nineteenth century. Rogowski’s analysis hinges on the interaction of land (agricultural interests), labor, and capital, aggregations that are viewed as acting purposefully and strategically in the face of changing conditions in international trade. To be persuasive, we need to do more than impute (actor-centered) motives and strategies to these aggregations; we have to demonstrate that these actors were in fact players in Scharpf’s sense, i.e. that these aggregations were cohesive and strategic. But we know from historical work—and indeed from the logic of Rogowski’s own analysis—that these aggregations were not coherent players, each being riven by internal tensions that derive from precisely the changes in international trade that are the focus of Rogowski’s analysis. Rogowski’s argument itself suggests that business interests would have been divided, depending on whether firms were oriented toward the domestic market or the international market. And we know from historical work that this was in fact the case. Similarly, among labor, workers’ interests frequently followed the divisions within business, opening the door for “cross-class alliances” that were crucial to the outcomes in which Rogowski is interested (Gourevitch 1986, Swenson 1989). Swenson’s (1989) deeply inductive, historical approach is clearly more micro-foundational than Rogowski’s rational choice approach in the sense that Swenson gets much closer to the actual players whose strategic interactions produced the outcomes at issue.

There is no dichotomy because taking micro foundations seriously means that we cannot be content to impute coherence to actors identified by the analyst; we must do the empirical work to make sure that the actors to whom we attribute certain strategic behaviors are in fact “players” in the first place.10 In other words, the fact that a particular analysis employs a rational choice perspective does not necessarily mean that it stands on strong micro foundations. Conversely, historical institutional research is not necessarily not micro-foundational—quite the contrary, as we have seen.

**Functional Versus Historical View of Institutions**

It is frequently noted that, unlike most historical institutionalism, a good deal of rational choice theory embraces a functional view of institutions (e.g. Hall & Taylor 1996:943–44, Pierson 1996). This distinction probably originated in the two schools’ differing approaches to the issue of institutions. Zysman notes

---

10A major theme in historical institutionalism is the way actors and their interests are constituted historically; see Thelen & Steinmo 1992, as well as the discussion, below, of the critical junctures and policy feedback literatures.
succinctly that “rational choice institutionalists start with individuals and ask
where institutions came from, whereas historical institutionalists start with in-
stitutions and ask how they affect individuals’ behaviors” (1992, comments at
Conference of Europeanists, Chicago). Indeed, much of the early and
pathbreaking work in rational choice did pose the question in the way Zysman
suggests, and the question of why institutions emerge and are sustained has
been answered in terms of the functions that they perform. One example is the
literature on the US Congress and the way its rules eliminate “cycling”; an-
other is the rational choice literature on institutions in international relations,
which are viewed as mechanisms by which states can reduce transaction costs
and achieve joint gains in an anarchic world (e.g. Shepsle & Weingast 1981,

However, there has been much work within rational choice that, like his-
torical institutionalism, embraces a non-functionalist, more historical view of
institutions (Pierson 1996:131). North’s later work (e.g. 1990), for example, is
concerned with tracing, historically, the emergence of different kinds of institu-
tional arrangements that either promote or distort development. Knight too
has criticized a good deal of rational choice literature for embracing either an
evolutionary or a spontaneous-creation view of institutions (1992:ch. 1).
Against more functionalist accounts, Knight sets his own model of institu-
tional formation and change, which places issues of distributional conflict at
the center of the analysis.11

These works share much with the more historical view of institutions
embraced by historical institutionalists. Zysman writes of political-economic
institutions, “The institutional approach begins with the observation that mar-
kets, embedded in political and social institutions, are the creation of govern-
ments and politics” (1994:244). This is close to Knight’s view, and for that
matter March & Olsen’s (1989); none of these works makes any assumptions
about the social efficiency of institutional arrangements, and all of them allow
for suboptimality and inefficiency. Differences remain, of course. Here I sim-
ply wish to point out that these differences do not boil down to the commonly
cited divide between rational choice’s “functionalist” and historical institu-
tionalist’s more “historical” approaches to institutions.

Synthesis or Creative Borrowing?

Should we conclude from the above that there has been a blending of these dif-
ferent approaches? In my view, no. What we see is a partial convergence of the

11Moreover, in some cases a functionalist view of institutions seems perfectly warranted
because it is consistent with the historical record. Weingast’s study of the “balance rule” in
antebellum America, for example, provides evidence that this institution was precisely designed to
issues at stake as, for example, historical institutionalists have come to a deeper appreciation of micro foundations and problems of collective action, and rational choice theorists have come to treat preferences, norms, and beliefs as a more central (also more complicated) issue than heretofore. But, as pointed out above, differences remain—in how theorists working out of these different traditions approach these issues, in how they generate the hypotheses that guide their work, and in the level at which they attempt to build theory, for example.

Rather than a full-fledged synthesis, we might instead strive for creative combinations that recognize and attempt to harness the strengths of each approach. This seems to be the strategy advocated by prominent proponents within each school, Zysman from a historical institutionalist perspective and Scharpf from a rational choice perspective. Thus, for example, Zysman (1994: 277) argues that institutions and broad processes of social change certainly have micro-foundations. The “naked” institution emerging from a state of nature by rational choice and the “socially embedded” institution are one and the same, but they represent two different narratives whose perspectives highlight different processes within a common story. That is, the arguments built around institutions and historical dynamics should be consistent with notions of the “rational” dynamics of individual behavior. Inconsistencies are instructive to both those who would build micro-foundations and macro-theories.

He likens the difference between the two perspectives to that between “high-level computer languages (historical narrative) and the bit-level machine language of the computer (microeconomic narrative)” and maintains that “inherently they must work together, they must be consistent” (Zysman 1994:277) and that “issues must be segmented to make appropriate use of the perspectives, not to reject the insight of one or the other as part of an ideological quarrel” (1994:278; see also Ostrom 1995, esp. pp. 177–78).

Scharpf (1997) elaborates in some detail how this might be achieved. He advocates the use of rational choice and game theory as a way of generating hypotheses that can discipline empirical analysis, but he acknowledges that “even when we can rely on models with high predictive power, they are likely to be of limited scope and will only represent certain subsets of the complex, multiarena and multilevel interactions that are characteristic of real-world processes” (1997:31). This being the case, he argues that “it is usually necessary
to combine several such modules into a more complete explanation” (1997:31, emphasis in original). The composite explanation of particular processes is likely to be unique for each country but...the modules employed in constructing it may reappear more frequently in other cases as well and thus are more likely to achieve the status of empirically tested theoretical statements. Even then, however, the linkages between these modules remain problematical.... Thus we will often depend on narrative, rather than analytical, connections between partial theories that have analytical as well as empirical support—which also means that the composite explanation itself remains vulnerable to charges of being ad hoc. (Scharpf 1997:32)

**Equilibrium Order Versus Historical Process**

One area does seem to distinguish these two analytic approaches, at least in emphasis. This distinction can be characterized in terms of the relative centrality of “equilibrium order” versus “historical process” in the analysis of political phenomena. One of the defining features of rational choice institutionalism is its assumption of equilibria and its view of institutions as coordinating mechanisms sustaining these equilibria (Levi 1997a:27, Scharpf 1997:10, Shepsle 1989:145). As Levi (1997a) has emphasized, this does not imply efficient equilibria; indeed, a major concern of rational choice has been to sort out the reasons why individual actors, behaving rationally, often produce suboptimal or inefficient collective outcomes. Nor does the equilibrium assumption imply that there exists a single, unique equilibrium outcome; another central problem in rational choice is to understand the process through which one equilibrium rather than another is reached. And finally, this assumption does not mean that rational choice theorists are uninterested in political change—merely that they tend to treat it as involving a transition between equilibrium orders. As Orren & Skowronek put it, “institutional politics appears as ‘normal,’ as politics as usual, explicitly or implicitly opposed to an extraordinary politics, in which equilibria are upset, norms break down, and new institutions are generated” (1994:316). Bates et al (1998a) appear to concur, also with the implicit critique: “The greatest achievement of rational choice theory has been to provide tools for studying political outcomes in stable institutional settings…. Political transitions seem to defy rational forms of analysis” (1998a:604–5). (Bates et al address this problem by incorporating elements of an “interpretivist” approach into the analysis).

---

13This section is strongly influenced by Orren & Skowronek (1994). I perhaps draw a less sharp distinction than they do between equilibrium analysis and process analysis, since the equilibria described by rational choice theories are not static but dynamic, and conversely, the “feedback” literature in historical institutionalism (see below) can be seen as a kind of equilibrium analysis as well.
Whereas rational choice theorists tend to view institutions in terms of their coordinating functions, historical institutionalists see institutions as the legacy of concrete historical processes. In embracing this view, historical institutionalism “brings questions of timing and temporality in politics [rather than equilibrium order] to the center of the analysis of how institutions matter” (Orren & Skowronek 1994:312). This does not mean that historical institutionalists are uninterested in regularities and continuities in politics; it just means that the emphasis tends to be on political development as a (structured) process and on the way institutions emerge from particular historical conflicts and constellations (e.g. Steinmo 1993, 1994). As Pierson puts it, historical institutionalism “stresses that many of the contemporary implications of…temporal processes are embedded in institutions—whether these be formal rules, policy structures, or norms” (1996:126; see also Skocpol 1992:58–59).

Orren & Skowronek emphasize two features of political life that have been central to historical institutional analyses. The first is that “institutions, both individually and collectively, juxtapose different logics of political order, each with their own temporal underpinnings” (1994:320). That is, the various institutional arrangements that make up a polity emerge at different times and out of different historical configurations. For this reason, the various “pieces” do not necessarily fit together into a coherent, self-reinforcing, let alone functional, whole. For example, some analysts treat the German political economy as a well-integrated “system” in which various institutional subsystems—vocational education and training, collective bargaining institutions, financial institutions, bank-industry links—form a mutually reinforcing whole. Historical institutionalists, by contrast, are likely to be concerned with the origins rather than the functions of the various pieces, and indeed, historically oriented research has demonstrated that the evolution of the German model was highly dysynchronous and full of unintended consequences (Streeck 1992; P Ma-now, unpublished manuscript; Thelen & Kume 1999). This has important implications for our view of the operation of this system. Streeck (1997), for example, has drawn attention to the ways in which industrial-relations institutions actively generate problems and pressures in other parts of the system, especially vocational education and social welfare institutions. As in the more functionalist view, the interdependencies among the various parts of the system are central to the analysis; however, in Streeck’s work the frictions as well as the functional interdependencies come to the fore. The result is very much in the spirit of Orren & Skowronek’s characterization: “The single presumption abandoned is that institutions are synchronized in their operations or synthetic

14As the large literature on comparative statics demonstrates, they are indeed interested (see Thelen & Steinmo 1992); see also the discussion of the “feedback effects” literature below.
in their effects; the more basic idea, that institutions structure change in time, is retained” (1994:321).

The second claim, related to the first, is that one important source of change comes from the interactions of different institutional orders within a society, as “change along one time line affects order along the others” (Orren & Skowronek 1994:321), that is, as interactions and encounters among processes in different institutional realms open up possibilities for political change. Stark & Bruszt’s (1998) work on the transition to democracy and market economy in Eastern Europe provides one example of what this looks like in practice. In language that resonates with Orren & Skowronek’s, they argue that in Eastern Europe, “we see social change not as a transition from one order to another but as transformation—rearrangements, reconfigurations, and recombinations that yield new interweavings of the multiple social logics...” (1998:7). Like Orren & Skowronek, they stress the incongruities among the multiple processes as they unfold: “...within any given country, we find...many [transitions] occurring in different domains—political, economic, and social—and the temporality of these processes is often asynchronous and their articulation seldom harmonious” (Stark & Bruszt 1998:81). Change in one arena affects other ongoing processes, which is what drives institutional evolution.16

Pierson’s (1996) analysis of the evolution of social policy in the European Union provides further examples. In one instance (the case of EU policy on gender equality), he shows how provisions adopted by the EU member states in one period—largely symbolic and without much effect—were later picked up by emergent women’s groups, who were able to use these provisions to achieve gains at the EU level that had eluded them at the domestic level. Changes in the political and socioeconomic context brought new actors into the game (in Pierson’s case, women’s groups) who were able to use existing but previously latent institutions (in Pierson’s case, Article 119 of the Treaty of Rome), whose new salience had important implications for political outcomes (Thelen & Steinmo 1992:16).

These examples point to the importance of examining politics as a dynamic process that frequently produces unintended consequences as different, ongo-

---

15Weir (1992) draws attention to this idea in her analysis of “collisions” between different policy streams; it is also related to Pierson’s (1996) notion of “gaps” and “lags” in policy processes that produce openings for institutions to evolve in ways their designers did not anticipate (see below).

16The concept of institutional bricolage, which Stark and other students of Eastern Europe employ, describes “an innovative process whereby new institutions differ from but resemble old ones” (Campbell 1997:22). This echoes Orren & Skowronek’s comment that “more often than not, we expect to find that continuities along one dimension of order and time will be folded into, and formative of, the extraordinary changes we are observing along another” (1994:322).
ing processes interact. Perspectives that conceive of change as the breakdown of one equilibrium and its replacement with another do not capture this well. Nor, however, do alternative conceptions, for example, some early versions of the new institutionalism in sociology, in which the definition of institutions as “shared cultural scripts” obscures political struggles among competing scripts and/or conceives of change as the displacement of one script by another.¹⁷

In sum, although historical institutionalists are just as interested as “other” institutionalists in the regularities of politics over time, they tend to emphasize historical process over equilibrium order. Whereas alternative conceptions view institutions in terms of their coordinating functions, historical institutionalists see them as the product of concrete temporal processes. Thus, rather than conceiving of institutions as “holding together” a particular pattern of politics, historical institutionalists are more likely to reverse the causal arrows and argue that institutions emerge from and are sustained by features of the broader political and social context. In this approach to institutions, path dependency involves elements of both continuity and (structured) change; institutions are conceived in relational terms (Immergut 1992, Katznelson 1997:104); and institutional arrangements cannot be understood in isolation from the political and social setting in which they are embedded.

**PATH DEPENDENCY**

Two ways of thinking about path dependency— one from the literature on economics and technology, the other from the work of “new” institutional sociologists— have gained prominence, and a brief discussion of each provides a baseline for a discussion of the historical institutional approach to path dependency. I argue that both contain insights into the mechanisms that sustain particular patterns of politics, but some of the most prominent formulations tend to obscure the distributional consequences of political institutions and blend out important sources of dynamism in political life.

*Technological Models of Path Dependency from Economics*

The most widely invoked model of path dependency is the one that comes out of the work of economists seeking to understand technological trajectories. Most closely associated with the “QWERTY keyboard,” the argument developed by David (1985) and elaborated by Arthur (1989) holds that certain technologies, for idiosyncratic and unpredictable reasons, can achieve an initial

advantage over alternative technologies and prevail even if in the long run the alternatives would have been more efficient. (Krasner 1988, Kato 1996a, and Pierson 1997 all review these arguments in detail.) What political scientists have taken from this is the intuitively attractive idea that technology, like politics, involves some elements of chance (agency, choice), but once a path is taken, then it can become “locked in,” as all the relevant actors adjust their strategies to accommodate the prevailing pattern.

Some features of politics are undoubtedly subject to the kinds of “positive feedback” effects to which the David/Arthur model of technological change draws attention, and the notion of increasing returns certainly has important applications to politics. But as a general guide to understanding political development, the QWERTY model is both too contingent and too deterministic. It is too contingent in that the initial choice (call it a “critical juncture”) is seen as rather open and capable of being “tipped” by small events or chance circumstances, whereas in politics this kind of blank slate is a rarity, to say the least. The openness implied in this model is belied by the vast literature on critical junctures (discussed below) that traces divergent trajectories back to systematic differences either in antecedent conditions or in the timing, sequencing, and interaction of specific political-economic processes, suggesting that not all options are equally viable at any given point in time.

The QWERTY model is also too deterministic in that once the initial choice is made, then the argument becomes mechanical. There is one fork in the road, and after that, the path only narrows. In this model, actors adapt to prevailing institutions by investing in them in ways that reinforce the institutions (e.g. people learn to type in a particular way, firms make products that fit with the standard). In other words, in the world of firms and users of technology, adapting to the standard means adopting it; those who do not adapt lose, and—importantly—the losers disappear (for example, as firms go out of business).

Politics is characterized by disagreement over goals and disparities in power, and in fact institutions often reinforce power disparities (Hall 1986, Knight 1992, Riker 1980:444–45). However, the losers do not necessarily disappear, and their adaptation can mean something very different from embracing and reproducing the institution, as in the technology model. For those who are disadvantaged by prevailing institutions, adapting may mean biding their time until conditions shift, or it may mean working within the existing framework in pursuit of goals different from—even subversive to—those of the in-

---

18 The best treatment of these issues is by Pierson (1997), and the policy feedback literature discussed below provides numerous examples of increasing returns arguments in politics.

19 David (1985), especially, emphasizes “chance elements” (p. 332) and “essentially random” factors (p. 335) in determining among an apparently very wide range of possible outcomes; Arthur (1989) is overall more circumspect and nuanced.
stitution’s designers. Such considerations provide insights into the reasons why, in politics, increasing returns do not necessarily result in an irrevocably locked-in equilibrium; further choice points exist.

**Path Dependency in Institutional Sociology**

Another strong argument about path dependency emerges from the work of the new institutionalists in sociology. Whereas economic models start with individuals or firms in the market, sociological perspectives begin with society. Institutions, in this view, are collective outcomes, but not in the sense of being the product or even the sum of individual interests. Rather, institutions are socially constructed in the sense that they embody shared cultural understandings (“shared cognitions,” “interpretive frames”) of the way the world works (Meyer & Rowen 1991, Scott 1995:33, Zucker 1983:5). Specific organizations come and go, but emergent institutional forms will be “isomorphic” with (i.e. compatible with, resembling, and similar in logic to) existing ones because political actors extract causal designations from the world around them and these cause-and-effect understandings inform their approaches to new problems (DiMaggio & Powell 1991:11, Dobbin 1994). This means that even when policy makers set out to redesign institutions, they are constrained in what they can conceive of by these embedded, cultural constraints.

The strong emphasis on cognition in the new institutionalism in sociology gives us powerful insights into the persistence of particular patterns of politics over time, but as DiMaggio & Powell point out (1991, esp. pp. 1, 11–12), the early formulations (e.g. Meyer and colleagues) were less helpful in understanding change. Some versions of new institutionalism in sociology make it hard to see any forks in the road at all; for example, Zucker (1991:85) argues that each actor fundamentally perceives and describes social reality by enacting it, and in this way transmitting it to the other actors in the social system....The young are enculturated by the previous generation, while they in turn enculturate the next generation. The grandparents do not have to be present to ensure adequate transmission of this general cultural meaning. Each generation simply believes it is describing objective reality.

20 An example of the latter is the job classification system in American industrial relations. This system was originally imposed on unions by employers as a way of controlling labor. Unable to change the system, emergent unions adapted their strategies to it but sought to attach rules to these job classifications, and in doing so, they eventually turned it into a system of union control. In this case, “adapting” to the institution had the effect of transforming it altogether, so much so that now it is employers who attack the system, unions who defend it.

21 The New Institutionalism in Organizational Analysis (Powell & DiMaggio 1991) is partly a response to these weaknesses. See also Powell & Jones (1999), which explicitly addresses institutional change.
The notion of institutions as shared scripts sometimes obscures conflicts among groups (because the scripts are by definition shared), and the notion of isomorphism emphasizes continuity across time and space (because new problems are solved using the same cultural template). Yet we know that dominant cultural norms emerge out of concrete political conflicts, in which different groups fight over which norms will prevail (Katzenstein 1996); we know that dominant policy paradigms can and do shift at times (Hall 1993); we know that organizational fields are often imposed by powerful actors (Fligstein 1991, DiMaggio 1991), and that legitimacy, not automaticity, explains why people follow scripts in the first place (Stinchcombe 1997). This is why recent formulations argue that the cognitive dimensions (though important) should not eclipse the strategic and political elements of action, and frequently find that in questions of institutionalization and institutional change, the political part of the story (and not the cognitive) is more important (DiMaggio & Powell 1991:27,31; see also Katzenstein 1996). At a minimum, then, much work remains to be done to sort out the relationship between the political (decision/power) and the cognitive (script) aspects of institutional stability and change.

Both the economic-technological and the sociological-institutional perspectives provide strong tools for understanding continuity, but by stipulating and privileging particular mechanisms of reproduction (coordination effects for the former, isomorphism for the latter) they have a hard time incorporating notions of conflict and power, and they are not particularly helpful in talking about change. Dynamism in both models has to come from some exogenous shock, or, as Orren & Skowronek (1994) argue for equilibrium models generally, these perspectives strongly imply that political change is not amenable to the same type of analysis we use to understand the operation of the institutions themselves.

PATH DEPENDENCY IN HISTORICAL INSTITUTIONALISM

Ikenberry captures the essence of a historical institutional approach to path dependency in his characterization of political development as involving “critical junctures and developmental pathways” (1994:16ff). As the phrase implies, this approach includes two related but analytically distinct claims. The first involves arguments about crucial founding moments of institutional formation that send countries along broadly different developmental paths; the second suggests that institutions continue to evolve in response to changing environmental conditions and ongoing political maneuvering but in ways that are constrained by past trajectories. These two lines of argument tend to be reflected in a bifurcation of the literature in this area. A number of important analyses of critical junctures explore the origins of diversity across nations
(e.g. Skocpol 1979, Luebbert 1991, Ertman 1997); other studies focus on the logic and self-reinforcing properties of particular national trajectories over time, drawing out comparisons to other countries where relevant (e.g. Weir 1992, Skocpol 1992).22

Although obviously related, these two literatures have characteristic strengths and weaknesses, and each would be enriched by a more sustained engagement of the other. The great strength of the critical junctures literature lies in the way in which scholars have incorporated issues of sequencing and timing into the analysis, looking specifically at the different patterns of interaction between ongoing political processes and at the effect of these interactions on institutional and other outcomes. Where this literature has generally been weaker is in specifying the mechanisms that translate critical junctures into lasting political legacies. Here the policy feedback literature, which has provided many insights into the mechanisms that account for continuity over time, is useful. However, in this second literature, strong tools for understanding continuity are not matched by equally sophisticated tools for understanding political and institutional change. In the next three sections of this essay, I argue that greater insight into the different types of reproduction mechanisms behind different institutional arrangements holds the key to understanding what particular kinds of external events and processes are likely to produce political openings that drive (path-dependent) institutional evolution and change.

**Historical Institutional Analyses of Critical Junctures**

As Katznelson (1997) suggests, the macro-historical analysis of critical junctures that set countries along different developmental paths has long been the bread and butter of historical institutionalism. Rejecting a functionalist view of institutions, historical institutionalists see institutions as enduring legacies of political struggles. The classics in this genre (ably reviewed by others, e.g. Ikenberry 1994) include Moore (1966), Gerschenkron (1962), Lipset & Rokkan (1967), and Shefter (1977). All of these works emphasize sequencing and timing and, related to these issues, different patterns of interaction between ongoing political and economic processes in the formation and evolution of institutional arrangements. These studies are “configurative,” as Katznelson (1997) puts it, in the sense that they do not view political processes in isolation but rather focus specifically on how the temporal ordering of, and interactions among, processes influence outcomes—in these cases, institutional outcomes.23

---

22There are, however, studies that treat both the cross-national differences and the over-time continuities within countries. Steinmo (1993) and Collier & Collier (1991) are good examples.

23However, I disagree with Katznelson’s distinction between “configurative” and “variable centered” analyses. I understand one of the great strengths of these works to be the framing of key variables in a way that captures the interactive nature of these processes.
This venerable tradition is alive and well in historical institutional research, as a review of a few select works demonstrates. Collier & Collier’s landmark study, *Shaping the Political Arena*, links differences in patterns of labor incorporation to variation in party and regime outcomes across a wide range of Latin American countries. As in the earlier works cited above, Collier & Collier emphasize the central importance of sequencing and timing; their study “confront[s] the interaction between a longitudinal and a cross-sectional perspective: between the unfolding over time within each country of phases of political change, and a sequence of international developments that influenced all the countries roughly in the same chronological time, but often at a different point in relation to these internal political phases” (Collier & Collier 1991:19–20). In fact, one of the book’s central themes is the way in which “common” international events or trends translate into different challenges in different countries as a result of their intersections and interactions with ongoing domestic processes (see also Locke & Thelen 1995, Collier 1993).

Another example of critical junctures work in historical institutional research is Ertman’s *Birth of the Leviathan*, which traces the origins of state institutions across a broad range of European countries from the twelfth to the eighteenth century. Employing a logic that parallels Gerschenkron’s, Ertman argues that “differences in the timing of the onset of sustained geopolitical competition go a long way toward explaining the character of state infrastructures found across the continent at the end of the 18th century” (1997:26). Like Collier & Collier, Ertman attends to variation in the ways in which international forces intersect with ongoing domestic political developments. Where state-builders faced geopolitical competition early, they were forced into greater concessions to the financiers, merchants, and administrators who financed and staffed the bureaucracy, resulting in patrimonial systems. Where rulers confronted geopolitical pressures later, “they found themselves in a quite different world,” where developments in education and finance made these side payments unnecessary, resulting in greater bureaucratic autonomy (Ertman 1997:28).

P Manow’s analysis of union formation in Germany and Japan (unpublished manuscript) is also centrally concerned with how the intersection and interactions among different processes affect institutional outcomes. Against conventional analyses that attribute the different institutional forms adopted by the two labor movements to the triumph of social democratic ideology (Germany) or to employer strategies (Japan), Manow reveals that, in both

---

cases, the evolution of state social policy intersected with union formation and affected it by systematically favoring certain organizational forms over others. A complementary analysis by Thelen & Kume (1999) reveals that the organizational forms ultimately embraced by the two labor movements were also powerfully supported by differences in the systems of vocational education and training as these emerged in the early industrial period. In both analyses, attention to the timing and sequencing of union development, in relation to and in interaction with other ongoing political processes (the institutionalization of social policy and of vocational training), helps to explain the organizational forms that the labor movement ultimately adopted.

What all of these comparative historical studies share is a perspective that examines political and economic development in historical context and in terms of processes unfolding over time and in relation to each other, within a broader context in which developments in one realm impinge on and shape developments in others. Each of these works demonstrates that “[c]ausal analysis is inherently sequence analysis” (Rueschemeyer et al 1992:4). All of them engage in close examination of temporal sequences and processes as they unfold, and perhaps even more importantly, as different processes at the domestic level or at the international and domestic levels unfold in relation to one another. They all focus on variables that capture important aspects of the interactive features of ongoing political processes, and in ways that explain important differences in regime and institutional outcomes across a range of cases.

However, many of these works tend not to emphasize or even sufficiently problematize how the outcomes of critical junctures are translated into lasting legacies. In other words, they neglect the mechanisms for the “reproduction” (Collier & Collier 1991) of the legacy over time within a particular country. (Collier & Collier 1991 and Skocpol 1992 are exceptions; see below.) A good deal of this literature, old and new, invokes a similar language, arguing that the events described are important because they had the effect of “filling the political space” in ways that were difficult to reverse or alter. Pierson articulates what appears to be the logic behind such arguments when he suggests that feedback effects are likely to be “most consequential in issue-areas (or in countries...) where interest group activity is not yet well established…. Factors that give one set of organizations an initial advantage—even a small one—are likely to become self-reinforcing” (1993:602–3). However, it is not clear that this reasoning stands up to historical scrutiny. The history of organized labor—to consider one important interest group—is actually littered with organizational forms that, despite some early-comer advantages, did not manage to

---

25The metaphor of political space comes up, explicitly or implicitly, time and again in “legacies” arguments. See Lipset & Rokkan (1967:51ff), Valenzuela (1979:ch. 6), Luebbert (1991, e.g. pp. 9–11), Skocpol (1992, esp. pp. 52 and 530), Thelen & Kume (1999), and Herrigel (1993).
survive, let alone dominate the available political space (e.g. materials-based unions in Sweden and skill-based organizations in Germany). Instead, they withered for lack of mechanisms to reproduce themselves.\textsuperscript{26}

Political space arguments (and related arguments about the “freezing” or “crystallization” of particular institutional configurations) obscure more than they reveal unless they are explicitly linked to complementary arguments that identify the mechanisms of reproduction at work. Without these, they are at best incomplete, for they cannot explain why these patterns persisted and how they continue to dominate the political space. Indeed, the language of freezing and crystallization can be deeply misleading because it suggests that things stand still, when in fact we know intuitively that organizations such as political parties or unions with roots in the nineteenth century must adapt to myriad changes in the environment in order to survive into the twentieth century. The reproduction of a legacy, in short, is a dynamic process, and this is not well captured in some of the dominant formulations.

With this perhaps in mind, several authors invoke Stinchcombe’s (1968) arguments about “sunk costs” and “vested interests” that make embarking on alternative paths costly and uncertain. But such references, though a promising starting point for the analysis, cannot themselves replace the analysis; these concepts need to be applied, not just invoked. Among other things, we need to know exactly who is invested in particular institutional arrangements, exactly how that investment is sustained over time, and perhaps how those who are not invested in the institutions are kept out. Attending to these issues is likely to generate insights into differences in the mechanisms of reproduction that sustain different kinds of institutional arrangements, or even the same kinds of institutions in different contexts.

Within the critical junctures literature, Collier & Collier stand out in explicitly drawing attention to the issue of the reproduction of critical junctures legacies, as well as the related matter of a legacy’s duration (1991, esp. pp. 31–34). For example, the authors describe a pattern of labor incorporation for Mexico that is based on the ability of the labor-mobilizing party, Partido Revolucionario Institucional (PRI), to use the resources of the state to maintain the link to labor at all levels. The legacy of this type of labor incorporation, in other words, is reproduced through patronage, and indeed a form of patronage that reaches deep into society, so that ordinary workers (who may owe their jobs to the governing PRI party) are also materially invested in this system. This pattern of reproduction is quite different from, say, Brazil, where (as Collier & Collier show) labor incorporation involved harassment, repression, and coercion and left a legacy of labor alienation that was sustained by ongoing labor exclusion.

\textsuperscript{26}To argue that the groups institutionalized first were the ones that “stuck” is to beg the question of why some were institutionalized and others not.
The general point I wish to make is that the different forms of labor incorporation identified by Collier & Collier are sustained by different mechanisms of reproduction—at the extreme, broad-based patronage in Mexico versus repression and control in Brazil. These considerations are crucial to understanding differences in the duration of the various legacies [some critical junctures produce very stable regimes, whereas others seem to contain the seeds of their own destruction (Collier & Collier 1991:34)], but also—maybe more importantly—they are crucial to understanding what kinds of events or processes have the capacity to undermine the legacy in different countries. In the Mexican case, for example, the reproduction of the legacy was (predictably) especially vulnerable to developments that make it hard for the PRI to continue to deliver patronage benefits. And so we find that ongoing economic crisis in recent years has shaken the hegemonic position of the PRI by interfering with the party’s use of state resources to shore up political support (Collier & Collier 1991:759).

Summing up, the critical junctures literature has taught us a great deal about the politics of institutional formation and the importance of the timing, sequencing, and interaction of ongoing political processes in accounting for cross-national variation. Where many of these analyses have been somewhat less explicit, however, is in explaining what sustains the institutional arrangements that emerge from these critical junctures. The issue of continuity over time and feedback mechanisms that sustain institutions dynamically are at the center of a related body of work, to which I now turn.

Feedback Effects

The literature on policy feedback in historical institutionalism has been elegantly and thoroughly summarized (see especially Pierson 1993 and Ikenberry 1994). This literature follows on Krasner’s observation that “path dependent patterns are characterized by self-reinforcing positive feedback” (1988:83). The literature in this area points to two broad types of feedback mechanisms (Ikenberry 1994:20), though, as the examples below indicate, many analyses combine elements of both. One set of mechanisms, which Ikenberry refers to as functional, is perhaps better described as incentive structure or coordination effects (see also North 1990). What this means is that once a set of institutions is in place, actors adapt their strategies in ways that reflect but also reinforce the “logic” of the system. Zysman captures the essence of these arguments when

\[\text{392 THELEN}\]

As Krasner (1988:84) and Pierson (1997) point out, many of these arguments could be put in the language of “increasing returns,” here understood simply as a situation in which once a particular path is chosen, actors adapt to the existing institutions in ways that push them further along that trajectory (and in so doing, also render the path not chosen increasingly remote) (see also Levi 1997:28ff).
he states that “the institutional structure induces particular kinds of...behavior by constraining and by laying out a logic to the market and policy-making process” (1994:243). A few examples from recent work can illustrate this idea.

Streeck’s (1992) work on the political economy of Germany has shown that the existence of particular institutional arrangements (e.g. a national system for vocational education and training and centralized collective bargaining) affects firm strategies in ways that not only reflect but also actively reinforce these institutions. These arrangements, as he puts it, “force and facilitate” the pursuit of strategies based on high-skill, high-value-added production. As business adapts its strategies to institutional incentives and constraints, its adaptation encourages further movement along this trajectory, as firms come to depend on the existence of these institutions for their continued success in international markets. PA Hall & D Soskice (unpublished manuscript) take this argument a step further by suggesting that the presence of certain institutions (e.g. strong works councils) can raise the returns to the presence of other, complementary institutions (e.g. strong bank-industry links). The authors use this argument to explain why certain kinds of labor market arrangements tend to be associated in many advanced industrial economies with certain kinds of financial arrangements (especially “patient capital”). Esping-Anderson’s (1990) “conservative-corporatist” welfare state provides another example of this feedback mechanism. Since the conservative-corporatist welfare state is premised on the notion of a single breadwinner, family structures adapt to the incentives and disincentives it embodies, which is one reason why female labor market participation in such economies is low by international standards.

Schneider’s (1997–1998) analysis of the developmental state in Latin America provides another example. Schneider shows that the structures and policies of developmental states often have the effect of fragmenting business interests. Where states distribute economic benefits on a discretionary basis, firms orient their strategies toward direct, individual appeals to the government. State activity of this sort generates very weak incentives for firms to engage in collective action, resulting in anemic business associations. Levy (1999) comes to similar conclusions for France, showing how the traditionally dominant role of the state in the French political economy actively discouraged the emergence of strong intermediate (political-economic) associations. In what he calls “Tocqueville’s revenge,” recent attempts by the state to withdraw from its traditionally pivotal role in economic life have failed for lack of viable associations that can step into the regulatory void created by the state’s retreat. In other words, historically, the more the French state compensated for France’s weak associations, the less able it was to do anything else.

Vogel’s (1996) analysis of the politics of deregulation in the political economies of the advanced industrial countries also tracks the path-dependent evolution of institutions. His study looks at “how political-economic institu-
tions shape policy choices and also...how these choices in turn reshape the institutions” (Vogel 1996:9), and he identifies feedback mechanisms at both the ideational and the structural levels. Against contemporary theories predicting cross-national convergence in the face of globalization, Vogel finds that the regulatory reforms that individual governments have actually undertaken reflect and reinforce distinctive national trajectories based on different underlying ideas about the appropriate role of the state in the market and on structural features of the political-economic context. The picture that emerges from his analysis is one of evolution and change, but countries move along (nationally specific) well-worn paths, because the search for solutions to new international pressures is structured by prevailing domestic institutions.

The second feedback mechanism identified by Ikenberry (1994) has to do with the distributional effects of institutions. The idea is that institutions are not neutral coordinating mechanisms but in fact reflect, and also reproduce and magnify, particular patterns of power distribution in politics (see especially Pierson 1997). This body of work emphasizes that political arrangements and policy feedbacks actively facilitate the organization and empowerment of certain groups while actively disarticulating and marginalizing others. The distributional biases in particular institutions or policies “feed back” so that “over time, some avenues of policy become increasingly blocked, if not entirely cut off” as “decisions at one point in time can restrict future possibilities by sending policy off onto particular tracks” (Weir 1992:18,19).

Some of the best work in this area has been done by Skocpol and her collaborators. Protecting Soldiers and Mothers (Skocpol 1992) is an important contribution that reiterates but also elaborates some central themes of Skocpol’s earlier work. Skocpol explicitly problematizes the issue of interest formation, arguing that institutional arrangements “affect the capabilities of various groups to achieve self-consciousness, organize, and make alliances” (1992:47). (This is a major theme in Rueschemeyer et al 1992 and Weir 1992 as well.) For example, drawing on work by Shefter, Skocpol shows how the fragmentation of the state, as well as the organization of party competition along patronage lines, actively mediated against the development of a unified working class that could then spearhead the movement for comprehensive social policies in the United States. At the same time, the policies the government did devise powerfully shaped future possibilities for more comprehensive schemes. Specifically, the policy of granting social benefits to Civil War veterans contributed to the emergence of a self-conscious interest group (veterans and their widows) and endowed it with material and ideational resources that threw up barriers to other groups who might appeal to the state for protection on other grounds.

To take another example, Esping-Anderson (1990) draws attention to the “decommodifying” effects of universal welfare states, demonstrating that
these arrangements actively shore up the power of the political and economic organizations of the working class. The resulting pattern of politics contrasts sharply with patterns in the United States, where, as Skocpol and Katznelson emphasize, political institutions and government policy have if anything operated to disarticulate working class organizations and to disempower working class interests. Rather than taking the interests of political actors as given, all these authors step back to ask how groups originally got constituted in the particular ways they did, then to consider how this affects the groups’ understanding and pursuit of their interests. As Hall (1993:51) puts it, “The social construction of identities in other words is necessarily prior to more obvious conceptions of interest: a ‘we’ needs to be established before its interests can be articulated.” Weir has brought these insights to bear on the issue of coalition formation. “By channeling the way groups interact in politics and policy making,...institutions greatly affect the possibilities for diverse groups to recognize common interests and construct political alliances” (Weir 1992:24). The result is that, in some institutional settings, groups with the same “objective material interests” cannot find common cause.

My final example of the feedback literature emphasizing distributional effects is Karl’s (1997) study of “petro-states,” which paints an especially vivid portrait of path dependency that emphasizes the power-distributional biases of institutions. In countries as diverse in regime type, social structure, and culture as Venezuela, Iran, Nigeria, Algeria, and Indonesia, Karl finds that the adaptation of political-economic institutions to the oil economy produces pathologies that actively reinforce the dependence of these economies on oil, despite explicit efforts by many governments to use oil revenues to fuel more balanced economic development. Her view of path dependency stresses political-distributional feedback effects, arguing that the incentives embedded in political-economic institutions are “above all else...the reflection and product of power relations” (1997:xvi). In her cases, both societal and state institutions are irresistibly drawn to organize themselves around the oil economy (e.g. the domestic bourgeoisie shifts its activities to those linked to oil—where the money is—and the state becomes the center of rent-seeking behavior, so that state jurisdiction expands massively with the expansion of oil-related activities even as state autonomy and authority atrophy). For one case, Venezuela, Karl examines specific choice points in detail and shows how the various decisions “demonstrate that there was never an equal probability that other choices would be made in their place; that each decision was related and grew from the previous one; and that, except during uncertain moments of regime change, the range of choice narrowed from one decision to another as Venezuela moved further into its oil-led trajectory” (1997:226). In the end, the perverse and paradoxical effect is that oil revenues, rather than leading the way to development, perpetuate a dependence on oil and a failure to develop.
These works have taught us a great deal about the dynamic processes that help explain how stable patterns of politics persist and indeed reproduce themselves over time (Ikenberry’s “developmental pathways”). However, by emphasizing the mechanisms through which previous patterns are reproduced, many of these works downplay the factors that might tell us how they can be changed. The language of “lock-in” frequently obscures the fact that, because institutions are embedded in a context that is constantly changing, stability—far from being automatic—may have to be sustained politically. [Weingast’s (1998) analysis of antebellum America is a good case in point, for his characterization suggests that the balance rule was not at all automatic but had to be actively nurtured in light of changing external conditions.] Where the context is changing, those who are invested in particular institutions reevaluate their investment in light of these changes. Moreover, changes in one institutional arena can reverberate, provoking changes in other, complementary institutions (Skocpol 1992:59; PA Hall & D Soskice, unpublished manuscript). These considerations lead us to a discussion of institutional evolution and political change.

Institutional Evolution and Political Change

If positive feedback and increasing returns were the whole story, then prediction would be easy, since we could simply read the outcomes off the institutional configuration. But institutions evolve and change over time, and this is where Orren & Skowronek’s (1994) arguments about temporality and the unfolding of different processes over time become important. Above I argued that two alternative views of path dependency are overly deterministic; this weakness is at least partly attributable to the fact that both of them stipulate, at a very high level of abstraction, particular reproduction mechanisms that obscure conflict and make it difficult to talk about change. Orren & Skowronek, by contrast, present a more dynamic alternative, which—very importantly—focuses on the incongruities and intersections between different processes and institutional logics as they unfold over time. Illustrations of this approach in practice can be found in recent work by Weir and Pierson, among others. Weir’s (1992) study of US welfare policy demonstrates that the unexpected “collision” of two (previously) unconnected policy streams in the 1960s—the “War on Poverty” and the civil rights movement—had a profound impact on the evolution of employment policy in the United States, turning it in a direction that policy makers did not originally intend. Similarly, in his study of the evolution of the European Union, Pierson (1996) shows how “gaps” (between different levels of action—domestic and European-level) and “lags” (produced by disjunctures between short- and long-term events and considerations) created openings that allowed non-state actors (in his cases, women’s groups and EU bureaucrats) to influence institutional development in ways that the EU member states did not anticipate and could not control.
In some ways, however, Orren & Skowronek’s characterization is *too* fluid. The intersection and interaction of different processes unfolding in time is certainly an important feature of political life; the “collisions,” “gaps,” and “lags” to which Weir and Pierson have directed our attention are also clearly pervasive in politics. But what we need to know is which particular interactions and collisions are likely to be politically consequential— which of these, in other words, have the potential to disrupt the feedback mechanisms that reproduce stable patterns over time, producing political openings for institutional evolution and change. I suggest that the kinds of openings that particular institutional configurations offer depends on the particular mechanisms of reproduction that sustain them.

Institutions rest on a set of ideational and material foundations that, if shaken, open possibilities for change. But different institutions rest on different foundations, and so the processes that are likely to disrupt them will also be different, though predictable. Take the case of the welfare state. Esping-Anderson’s (1990) three models—social-democratic, conservative-corporatist, and liberal welfare systems—not only rest on different levels of support (from broad and diffuse to narrow and weak) but also rely on different mechanisms of reproduction, and therefore they are differently affected by specific other “external” trends. For instance, changes in gender relations and family structures are likely to reinforce elements of the universalistic and liberal welfare states (which both, though in different ways, support a high level of labor-force participation by women), but these changes create new frictions and contradictions for conservative welfare states, which are premised on the single-breadwinner model of the family. In other words, we might well expect a (politically consequential) collision between changing gender roles and welfare state development, but only in the conservative welfare states.28

Universal welfare states, on the other hand, may be especially susceptible to other kinds of pressures. Rothstein’s (1998) analysis of the universal welfare state, for example, suggests that middle-class support is crucial (because this is the pivotal electoral group) but that— unlike the working class—the middle classes are neither clear material beneficiaries nor clear losers in the universal welfare state. Whereas working-class support for the system includes a very strong material component, Rothstein argues that the support of the middle class is premised more on their belief that the system is fair, in the sense that they are notshouldering an undue burden, and that all citizens are contributing their fair shares (this is what Rothstein, following Levi 1998, calls contingent

---

28See also Pierson’s (1994) comparative analysis of welfare state retrenchment in Britain and the United States. But in contrast to Pierson, I wish to emphasize that it is not just a question of whether policies are more locked in or less, but rather of the different ways in which the policies are reproduced, which makes them vulnerable to different kinds of pressures.
This type of analysis raises the issues flagged above, concerning who has vested interests in particular institutions and what sustains these investments over time. Rothstein’s study reveals that the foundations of working-class support for the welfare state are fundamentally different from those of the middle class; both are invested, but in different ways. The coalition behind the universal welfare state, far from being self-reinforcing (as in increasing returns arguments), may in fact have to be politically re-invented from time to time, as the environment changes and as these various groups reevaluate their investment in light of these changes. As Rothstein (1998) points out, one development that could complicate the reproduction of the system is the emergence of individualized private-sector social services, which open up opportunities for those with sufficient resources to opt out of the standardized univeralistic programs. The growth of such alternatives could then upset the contingent consent of the middle classes, who would resent paying privately for their own more individualized services while also shouldering the burden of the standardized system on which they no longer draw. In short, a rise in individualism associated with demands for more choice and less standardization is potentially subversive to the universal welfare state, given its particular material and moral foundations.

Understanding the different mechanisms of reproduction that sustain different institutions is also the key to understanding why common international trends frequently have such different domestic consequences, disrupting previously stable patterns in some countries while washing over others seemingly without effect (see Locke & Thelen 1995). As many authors have noted, the political-economic institutions of the advanced industrial democracies have proven surprisingly resilient in the face of globalization pressures, and this certainly speaks for the strong feedback mechanisms at work. However, some countries have in fact seen important changes. Prominent examples are wage bargaining institutions in Denmark and Sweden, both of which have experienced substantial reconfiguration in the last 15 years. How do we make sense of the special vulnerability of the Swedish and Danish systems of wage bargaining to these putatively common international trends? The crucial starting point for any such analysis is to examine what ideological and material foundations sustained these institutions prior to the onset of these new pressures (Swenson 1989). Both the Danish and the Swedish models of the 1960s and 1970s were characterized by a high degree of egalitarianism, which rested on (and also reproduced) a particular coalition of interests and a set of ideational claims supporting egalitarianism. These arrangements were highly resilient in the face of several important changes—both domestic and international—through the 1970s, but they began to unravel in the face of market changes in the 1980s that systematically enhanced the bargaining power of skilled over unskilled workers (Pontusson & Swenson 1996). Specifically, the trend toward “diversified
quality production” was deeply subversive to systems like the Danish and the Swedish ones because it encouraged the development of shop-floor structures that were completely at odds with the overarching national wage bargaining institutions. By contrast, the trend toward diversified quality production did not subvert, and in some ways reinforced, key institutions in the German political economy because these were premised on a very different foundation and different mechanisms of reproduction, which accommodated (indeed nurtured) certain kinds of inequalities between skilled and unskilled workers.

A final example is the stability of party systems. Lipset & Rokkan (1967) may be right in arguing that party systems become “frozen” at particular junctures, but we know from Shefter’s work (1977) that the mechanisms that sustain and stabilize different party systems vary. Shefter’s distinction between patronage-based and programmatic parties clearly makes a difference to the kinds of events that are likely to disrupt stable patterns of politics. For instance, throughout the postwar period, both Italy and Sweden had very stable party systems that revolved around the dominance of a single hegemonic party, the (patronage-oriented) Christian Democrats in Italy and the (more programmatic) Social Democrats in Sweden. The Swedish party system could absorb a defeat of the Social Democrats at the polls (as in 1976, for example) without a full-scale breakdown. In Italy, however, the defeat of the Christian Democrats (not coincidentally, precipitated by corruption scandals) prevented the party from continuing to use state resources to shore up its political support, and so the crisis of the Italian Christian Democrats created a massive opening that quickly brought about a major reconfiguration of the political landscape.

In all of these cases, understanding moments in which fundamental political change is possible requires an analysis of the particular mechanisms through which the previous patterns were sustained and reproduced. In contrast to equilibrium and other models that separate the question of stability from the question of change and propose that they require different analytic tools, the examples above suggest that an understanding of political change is inseparable from—and indeed rests on—an analysis of the foundations of political stability (Skowronek 1995:96, Orren & Skowronek 1994:329–30).

CONCLUSION

Gourevitch (1986) is responsible for the memorable aphorism that, among comparativists, happiness is a crisis that hits a lot of countries—for in moments of crisis, the elements that previously held a system together come into full relief. This is essentially what I have argued here. It is possible to do better than to separate questions of institutional reproduction from those of institutional change, resigning ourselves to the idea that each requires an entirely different toolkit. Instead, drawing together insights from the critical junctures
literature and the literature on path dependency and policy feedbacks, I have argued that the key to understanding institutional evolution and change lies in specifying more precisely the reproduction and feedback mechanisms on which particular institutions rest. I take from Orren & Skowronek the important insight that much of what moves politics is the intersection and interaction of different ongoing processes, although I qualify this by arguing that only some of the resulting collisions are likely to be politically consequential, specifically those that interfere with the reproduction mechanisms at work in particular cases.

Attention to these matters will provide insights into some of the provocative issues raised but not necessarily fully answered by recent historical institutional work. These include the important issue of differences in the duration of critical junctures legacies, as well as the related question of why some institutional legacies seem to contain the seeds of their own destruction. Understanding these issues will require work that is “genuinely historical” (Skocpol 1992:59) in the sense that it tracks the unfolding of processes, individually and in relation to one another, over time. The link between the critical junctures literature (on institutional formation) and the feedback literature (on institutional reproduction) is thus clear: Knowing how institutions were constructed provides insights into how they might come apart.

Functionalist perspectives will not take us far, since they skirt the issue of the origins of institutions and the all-important matter of the material and ideological coalitions on which institutions are founded. This does not mean that borrowing from other perspectives is impossible; on the contrary, it may be quite fruitful. One can imagine conceiving and analyzing consequential policy collisions as “nested games,” for example, employing some of the tools of rational choice to sort out the logic of the situation and the responses of the actors. This could certainly form one of the “modules,” as Scharpf (1997) puts it, in a more comprehensive analysis. It will not, however, substitute for the process-oriented analysis that is characteristic of historical institutionalism, which is often the only way to understand how some games came to be nested within others in the first place.

Many of the insights from the recent feedback literature will certainly play a role as well; this work has provided invaluable tools for exploring the key issues of who, exactly, is invested in particular institutions, and what sustains these institutions dynamically over time. Institutional research’s traditional focus on continuity and stability is thus maintained, but in some cases this should be combined with greater attention to what specific mechanisms sustain that stability, for it is there that we will find clues as to the particular external processes that can produce political opening and change. Attention to the different mechanisms of reproduction will also lend insight into the distinctive ways that different countries are affected by putatively common international forces
and trends. In short, a more precise specification of the reproduction mechanisms behind particular institutions is the key to understanding important elements of both stability and change in political life.

ACKNOWLEDGMENTS

For their comments on this paper, I thank David Collier, Frank Dobbin, Peter Hall, Peter Katzenstein, Margaret Levi, Paul Pierson, Nelson Polsby, Walter Powell, Bo Rothstein, Fritz Scharpf, Ben Schneider, Stephen Skowronek, Sven Steinmo, Guillermo Trejo, and the participants in the faculty seminar “The New Institutionalism in Comparative Politics” at the Centro de Investigación y Docencia Económicas in Mexico City.


Literature Cited

APSA-CP. 1997. Newsletter of the APSA Organized Section in Comparative Politics. Culture and Rational Choice. 8:2 (Summer)


Skocpol T. 1979. States and Social Revolutions. Cambridge, UK: Cambridge Univ. Press


