Institutional Labor Economics, the New Personnel Economics and Internal Labor Markets: A Reconsideration

Paul Osterman

MIT Sloan School of Management

April, 2010

This paper originated as a presentation to a conference organized by Professor Isabel Calderon Gutierrez at the University Carlos III in Madrid and I am grateful to her and to the participants at that conference for their support. Helpful comments on later drafts were provided by Lucio Bacarro, Frank Dobbin, Robert Gibbons, Sandy Jacoby, Thomas Kochan, Mike Piore, and participants at the SASE annual meetings in Costa Rica where I presented an earlier version of the paper and, in particular, Peter Auer, Jerome Gautie, Harry Katz, David Mardsden, and Wolfgang Streeck. In addition, two referees and the editor offered very useful suggestions.
Abstract

The goal of this paper is to reconsider institutional labor economics and make a renewed case for it. Two recent developments motivate this effort: the rise of New Personnel Economics (NPE) as a significant subfield of labor economics and the very substantial shifts in work organization that have taken place in recent years. Understanding how and why firms have reorganized work opens the door for a renewed interest in institutional approaches.

The focus is upon how organizations structure their internal labor market (ILM) rules. The argument is that these rules emerge from the interplay of competing groups and competing views of legitimate objectives. As such the role of power and influence in establishing work rules is central although more conventional NPE considerations also remain important.

The paper illustrates the utility of institutional labor economics by reviewing evidence regarding features of ILMs and how they are changing. The empirical material includes the emergence of ILMs, patterns of diffusion of high performance work organization, variation across firms in obtaining effort and commitment from their workforce, the use of contingent labor, and the rise of equal employment opportunity and due process functions. The final section draws together the argument and its implications and discusses limitations and possible next steps.
A chestnut in any course on Industrial Relations is a review of old debates between the so-called institutionalists and more mainstream labor economists. Lester (1946) and Machlup (1946) debated the role of equilibrium and marginal analysis in labor economics while Lampman (representing what he termed the “empiricists”) and Rottenberg (representing mainstream theory) argued about realism and the standards against which theory should be judged (Lampman and Rottenberg, 1956). The local labor market studies of Reynolds (1951), Myers and Shultz (1951), and others critiqued the standard assumptions of mobility and perfect information and these scholars believed that “where frictions, imperfections, and the ‘human element’ in labor markets bulk larger…the competitive…model frequently gives answers that are incorrect, misleading, or unduly narrow and simplistic.” (Kaufman, 2008, p. 287)

By the mid-1960s these disputes seemed to have been settled with the rise of human capital theory and the elaboration of search and information models which accounted for many of the institutionalists’ observations regarding the operation of labor markets. The institutionalists were not eliminated but they retreated into schools of Industrial Relations and, less commonly, into business schools. This outcome notwithstanding, the goal of this paper is to reconsider institutional labor economics and make a renewed case for it.

One stronghold of institutional thinking has been researching and explaining the personnel practices of firms and, in particular, how they organize the rules of their internal labor markets (ILMs). Understanding ILMs is an important arena for institutional theory because, at least as a first approximation, the very existence of administered employment rules seems to point to a different understanding of labor
markets than that which emerges from invisible hand models of supply and demand. In addition, the current period is witnessing a remarkable number of changes in how firms organize work. Understanding what is driving these transformations is central to the intellectual agenda of the field and this in turn at least raises the question of what value an institutional perspective can provide.

Related to this point is that in the past several decades the standard models for organizing work and have lost their hold and we now observe substantial diversity in ILMs. This diversity is not difficult to understand if firms face different constraints but there is also diversity among firms that appear to be in the same economic circumstances and this pattern suggests a role for institutional considerations.

Despite the seeming appeal of institutional factors for understanding these patterns, in recent years in a replay of the old pattern, New Personnel Economics (NPE) has emerged as a significant subfield of labor economics and has a tendency to be imperialistic regarding its claims (Lazear, 2000). This line of research seeks to explain the personnel and internal labor market practices of organizations from an optimization perspective. For example, NPE has addressed the general question of why internal labor markets exist, has studied optimal compensation and promotion practices, retirement policies, and many other human resource topics. Because NPE is congruent with most of standard economic theory it has gone unchallenged yet the older tradition of institutional labor economics appropriately updated takes a different perspective, has a great deal to say on the same topics, and is worth a serious reconsideration. This is particularly true because somewhat below the radar screen more institutionally oriented scholars, many of whom are sociologists but some of whom are more direct descendants of the older
institutionalists, have enlarged and deepened their own analysis of organizations. Hence the substance of institutional economics is deeper and more vital than has been understood.

In the next section of the paper I describe the institutional approach that I will take to understanding developments in ILMs and I will outline the boundaries and limits of my argument. Subsequent sections of the paper illustrate the utility of institutional economics by reviewing evidence regarding features of Internal Labor Markets and how these ILMs are changing in the current environment. The final section draws together the argument and its implications and discusses its limitations and possible next steps.

**The Argument**

There are three logical steps in a research program aimed at understanding ILMs. The first is to ask whether in fact organizations have rules that at least nominally shape employment practices and outcomes. The second step is to determine whether these rules are binding in the sense that they lead to outcomes that are different than those predicted by standard supply and demand models. Assuming that the answers to the first two steps are positive the third step is to explain the origins of those rules and why they persist.

The simple existence of ILM rules is well accepted by scholars of all persuasions and has been so at least since the publication of Doeringer and Piore (1972). The second step in the chain of logic is more difficult because of the challenge of establishing the counterfactual (that is, what the results in a more standard labor market would look like)
and testing against it. Firm level data are required because national datasets lack detail on the specific personnel practices that shape outcomes for the individuals who are surveyed. As a result the research that addresses this question typically studies individual firms although in some instances the form in which the data comes is descriptive or ethnographic accounts and in other cases econometric analysis of personnel records. The sophisticated econometrics should not, however, obscure the fact that essentially each study deals with a sample of one firm. In this literature NPE researchers have played a very important role by providing more clarity about where observed patterns diverge from those predicted by traditional labor economics (see for example Baker, Gibbs, and Holmstrom, 1994, Gibbs and Hendricks, 2004, Gibbons and Waldman, 2006). In general a fair conclusion is that scholars of all camps accept the view that ILM rules do in fact shape outcomes in ways that differ from those predicted by models that ignore the administrative rules of the firm.

The debate arises regarding the third question, the origins and persistence of these ILM rules. New Personnel Economics (NPE), as summarized in a recent review by Edward Lazear and Katherine Shaw “is aimed at modeling firms’ use of optimal management practices. (Lazear and Shaw, 2007, p. 92).” The prototypical research strategy in the NPE is to identify a personnel practice that is empirically prevalent and develop a model in which the practice in question emerges as the most efficient solution to the problem posed in the model. In other words, practices are seen as flowing from the firm’s effort to maximize efficiency. As a second step, NPE adopts the standard economics equilibrium assumption that if organizations err and adopt sub-optimal
practices they will lose out in the competitive market. As a result natural selection will enforce the diffusion and dominance of practices that flow from the optimizing models.

In developing these models NPE, according to a recent review (Gibbons and Waldman, 1999), draws on four theoretical building blocks: human capital theory, optimal job assignment models, incentive contracting, and tournaments. The central ideas of each of these theoretical streams are well known. Human capital theory addresses the acquisition of skill and, in the context of a firm the determination and consequences of learning trajectories. Assignment models ask about how organizations can best match skills and job requirements. Incentive contracting concerns solutions to the well known agent-principal problem while tournament models focus on the use of promotions to provide incentives and to sort employees by quality.

I present the alternative institutional approach below. However, in making my argument it is also important to be clear that the claim is not that mainstream labor economics does not pay attention to institutions. Whatever the merit of this claim might have been in the past when the dominate research stream was human capital theory and earnings regressions it is certainly no longer true. There is a rich “standard” literature that examines the impact of minimum wages, unions, employment security legislation, temporary help agencies, and a variety of other labor market institutions. However, in this literature institutions such as those listed above are viewed as external constraints on the firm and the model of firm behavior, the model exemplified by NPE, remains one of optimizing a well defined objective function subject to external constraints be they supply and demand or external institutions.

The Institutional Perspective
The central theoretical idea developed below focuses on an alternate perspective on how firms make decisions regarding how to organize work. To address this question this paper joins other recent contributions that have sought to elaborate the meaning of institutional labor economics and to create space for this school of thought (Jacoby, 1990, Kaufman, 1994, Kaufman, 2004, Marsden, 1999).

The argument proceeds in two stages. The first is that organizations are characterized by groups with competing objectives and competing perspectives regarding the legitimate purposes of the organization. The second stage of the argument is that the ILM outcomes we observe are the result of an internal political process that is made necessary by the need to resolve these competing objectives. In this context the optimizing pressures emphasized by NPE are an important player in the political contest but only one of multiple players and the actual outcome we observe represents a particular resolution of the conflict. Over time, depending in part on the external economic and political environment, the relative power of different internal constituencies varies which in turn lead to differences in the relative importance of NPE versus other considerations. However, in all periods the group political process is important.

Groups and ILMs

The importance of organized work groups, in both union and non-union settings, is very well demonstrated in a very large and diverse ethnographic literature. Furthermore this literature establishes that these groups act collectively, in non-union as well as union settings, to pressure management to modify practices and policies. In the words of one review of this ethnographic literature (a review that contains numerous
“Workers’ goals can include oppositional elements that focus on resistance to management practices…. (Hodson, 1977, p. 426).” Another review of this literature summarized it as finding that “managers and workers jointly negotiate the trade-offs among management control, worker consent, and labor effort. (Choi, Leiter, and Tomaskovic-Devey, 2008, p. 423). It is important to understand that by “negotiate” the authors refer to informal interaction, pressure, and accommodation and not to formal union-management collective bargaining. The ethnographies include white as well as blue collar settings. It follows from the nature of these implicit negotiations and political struggles that the outcomes that we observe cannot be presumed to result from optimal decision making. It is not the case that the competing perspectives described above can simply be modeled as additional constraints to be considered by an optimizing leader.2

There are, of course, other models and ways of thinking that recognize that conflicts exist within organizations over goals and over behavior. That point is not new but what distinguishes the argument here is the emphasis on group rather than individual interests. The NPE approach to conflict focuses on principal-agent issues but the core of this approach is the view that the individual is the focal actor and that person is seen as an atomistic and selfish utility maximizer looking after his or her particular interests. On the other hand, social-psychologists emphasize a more expansive view of individual motivation (Pfeffer 2007; Baron and Kreps, forthcoming). However, as useful as is this socio-psychological viewpoint it is once again important to see that the discussion and examples used to support the argument center on such individual level behaviors as role of intrinsic motivation, the importance of reciprocity and gift-exchange, status inconsistency, the anchoring of expectations, and interpersonal comparisons of utility.
By contrast the perspective developed here emphasizes group norms, conflict, and negotiation.³

This focus on group conflict over goals and rules is consistent with the work of many scholars who have recognized that firms contain multiple groups with quite different perspectives on the legitimate objectives of the organization. Nearly a century ago John Commons wrote that “…economic conflicts are not merely conflicts between individuals…but are conflicts between classifications or even classes of individuals…”⁴ In his description of a French bureaucratic organization Crozier wrote of “opposing forms of rationality” (p. 298) and went on to observe a “complex equilibrium affecting the patterns of action, the power relationships, and the basic personality traits characteristic of the cultural and the institutional systems of a given society” (p. 294).

ILM Rules As The Result of a Political Process

The view that ILM outcomes are the result of a political process that aims to resolve group level conflict stands in contrast to what March and Simon called the “machine model” of behavior (March and Simon, 1958, p. 37). This is the view that organizations can been seen as an algorithm that take inputs, e.g. data on prices and production processes, and spit out the optimal response which they then execute. One concern is that this view assumes an organizational computational capacity doubts about which have long formed the basis of critiques (Nelson and Winter, 1982, Cyert and March, 1963). More fundamentally this machine model represents an inaccurate characterization of organizational decision making and is overly functional. As Jacoby
notes in his critique of standard theory, “There is a tendency to rationalize employment practices in functional or efficiency terms without regard to their historical-causal (how did the practice arise)…complexities (Jacoby, 1990, p. 317).”

In thinking about this political process it is important to understand that while it may be tautologically true that there is a decision maker in an organization who has to sign off on a policy in fact that person does not have unilateral power to impose an “optimal” decision but rather must take into account the views of constituencies within the organization and the power of those constituencies. Once this point is accepted then the door is opened for an institutional perspective on the determinants of ILM rules.

An extended illustration of the political process perspective is found in Gouldner’s classic book Patterns of Industrial Bureaucracy (1954). In a case study he analyzed the origins and implementation of bureaucratic rules regarding issues such as attendance, safety, bidding for jobs, and effort. In modern language he explored how the firm created the rules of its internal labor market. He described how each rule came into being and argued that they were not based on a universal functional rationality but rather the result of competing interests. For example, he showed how the world view and interests of several distinct groups of employees as well as both incumbent managers and newly hired outside managers all differed from each other and how these differences varied depending on the topic at hand. The point was not just that they had different self interests but also that they had different views about the underlying rationale of the organization and about what was appropriate and inappropriate behavior. As a result the firm’s personnel practices, its ILM reflected a political and social process that resolved these conflicts. Gouldner wrote that “Rather than assuming that bureaucracy
possesses a Gibraltar-like stability, the perspective directs attention into the tensions and problems evoked by bureaucratization” (p. 11).6

Dynamics Over Time

The institutional perspective offered here challenges the view that ILM rules emerge from an optimization calculation. Within organizations the urge towards optimization and efficiency competes with other objectives held by powerful constituencies. The form that ILMs take, the nature of the personnel rules, can only be understood as flowing from a complex interaction of these objectives and interests. This competition for policy includes an important role for group level factors such as norms and customs, social structure, competing interests, search for legitimization, and power.7

This said, the viewpoint presented here leaves ample room for NPE type considerations as represented by constituencies in the organization that place a high value on standard efficiency considerations. In recent years as competitive pressures on firms have ratcheted up many older ILM practices increasing looked costly and hence the pressure to change them have grown. In addition NPE scholars might also plausibly argue that underlying technologies have shifted in ways that affect optimal retention policies (for example in the direction of more general skills. See, for example Gould (2002)).

All of the above is certainly supportive of the idea that some of what we observe today as ILMs are transformed can be understood via standard models as shifting constraints that lead firms to implement new ideas about organizations, what Cappelli (1999) has characterized as “market-in,” such as new uses of on-site suppliers,
increased utilization of contractors and out-sourcers, and innovations in pay-for-performance. The empirical material presented in the next section begins with a discussion of the history of the emergence of ILMs and it will be apparent that the power of optimizing considerations has indeed been stronger in some periods than in others. However at all times it remained true that understanding the particular ways in which the economic pressures are refracted through institutional considerations is essential for a full understanding of the origins and operation of the practices that we observe. Groups within organizations seek to redirect pressures for change to their own liking and the actual ILMs rules which emerge reflect this institutional fact of life.

**Evidence**

The arguments developed above point to an understanding of the evolution of ILMs that combines pressures for economic optimization with the kind of institutional considerations that are often overlooked in the literature. In the sections that follow this perspective will be illustrated by five examples: the historical emergence of ILMs and present day political struggles over occupational boundaries, the slow diffusion of High Performance Work Systems, the strategies of firms to obtained increased levels of effort and commitment from their workforce, the use of contingent employees, and the spread of new promotion rules in response to equal employment opportunity pressures. In discussing each example the approach will be to identify the political contenders that seek to shape the nature of the ILM practice at hand. The kind of considerations emphasized by NPE models are important but must compete with other considerations pressed by various internal constituencies. As the earlier discussion argued, the final
resolution—the ILM practices that we observe—are the result of this political competition.

Job Ladders And ILMs

ILMs emerged gradually among American firms from roughly the beginning of the twentieth century until the beginning of the Depression and then accelerated in their diffusion and took on a somewhat different form with the arrival of the New Deal and the rapid growth of union power (Jacoby, 2004, Kaufman, 2008). Underlying the NPE story is the effort by firms to find more efficient ways to manage increasingly large organizations as well as the need, which varied over the cycle, to limit the costs of turnover. Turnover is costly because it entails a waste of human capital and because in a high turnover organization provision and investment in on-the-job training is problematic. In order to reduce turnover compensation is back loaded and this leads to a system in which people spend long periods of time with the same employer. This effect is intensified by the nature of human capital acquisition. The early investments in training need to be rewarded over time, from both the perspective of the employees and the firm, and again this implies long careers.

In some versions of the human capital story there is also a process by which firms need time to learn about the abilities of their employees and this implies firm based careers (e.g. Farber and Gibbons, 1996). The incentive interpretation is somewhat different in that it is less about skills and more about behavior. In this line of thought ILMs help resolve agent-principal problems by creating long-term employment settings.
In such settings opportunism can be minimized because there is always another round in the game (Williamson, 1975).

The institutional argument regarding the rise of ILMs takes two tacks. The first is to point to the role of unions in pressing firms to rationalize their employment practices in general as well as to adopt specific practices such as seniority-based promotion. The second thread of the institutional story is the role of personnel profession in acting in their own self-interest by encouraging the diffusion of formal employment practices which then required a personnel staff to administer.

The evidence is that both the NPE and the institutional stories were important over the course of the twentieth century. At the turn century American firms, even large ones, did not have well developed internal labor markets. Instead, they managed their workforce via the so-called “drive system” in which substantial discretionary (often arbitrary) power was put in the hands of the foreman (Jacoby, 2004). Turnover was high and was problematic during periods of growth and labor shortage and this led to efforts to stabilize the workforce via creation of job ladders and personnel systems. In his history of the rise of Human Resource Management Kaufman argues that while in the pre-New Deal period “HRM also moved up and down in response to the threat exerted by unions and government (Kaufman, 2008, p. 295)” he nonetheless places more explanatory weight on the problems posed by managing large organizations (Kaufman, 2008, p. 300). Jacoby writes that “employers now found personnel management an attractive alternative, one that promised to simultaneously relieve labor shortages, improve productivity, and promote labor peace,” (Jacoby, 2004, p. 102). Furthermore, Jacoby relates the ebb and flow of adoption of ILMs to shifting pressures of supply and demand.

With the onset of the Depression the relative importance of the different considerations shifts. In sectors that were unionized internal political pressures led firms to establish the formal job ladders and seniority systems that characterize industrial ILMs. Non-union firms then imitated these systems in order to avoid unionization. In addition, personnel specialists within firms were very active in pushing for the formalization of employment practices, a point of view that was clearly in their self-interest. Indeed, during the Depression and the World War II periods there was very little correlation between the standard optimizing correlates of ILMs—firms size and turnover rates—and the adoption of specific formal personnel practices. (Baron, Dobbin, and Jennings, 1986).

The core conclusion from this material is that when it comes to understanding how the prototypical ILM of the post-war period emerged the story is considerably more complex than that implied in NPE models and the complexity is consistent with the institutional perspective developed above. There was certainly considerable room for the play of market forces and efficiency maximizing considerations but to these were joined the actions of two important interest groups within firms, unions and personnel professionals, as well as the actions of an external institution, the government. Internal Labor Markets can only be understood as resulting from the interaction of these actors as they used their resources to pressure organizations in the directions that they preferred.

The role of internal political struggles over the shape of ILMs is an on-going story, not limited to a particular historical episode. A powerful illustration of this point
is the extensive literature on the determination of boundaries among occupations. The struggle over these boundaries is essentially a group level struggle over the definition of jobs and the steps in the ladder of an ILM.

At the core of this struggle, and central to the literature, is a rejection of the idea that there is some natural order regarding which tasks are assigned to which professions or occupations (Abbott, 1988). Rather, the creation of boundaries and steps in the job ladder is the result of political conflict, some of which is carried out in the public arena via professional societies and licensing authorities, but some of which takes place within the firm. A classic example is the conflict in the medical field between doctors, nurses, and medical technicians over who could do what for a patient (Freidson, 1970), a contest which is on-going and fueled by changes in technology and external shifts in funding. Even among doctors there are internal political tensions, for example between internists and specialists over who can treat which conditions. Abbot (1988) provides a wide range of examples of struggles of this kind in fields such as information technology, law, medicine, engineering, and construction. Furthermore, the literature on how self-conscious groups engage in political struggles within organizations to define the scope of their tasks and to protect themselves against competing groups is not limited to a few highly educated elite occupations. Barley (1986) and Nelson and Barley (1997) provide a wide range of evidence of similar processes in a range of “lower level” technical occupations.

The point for our purposes is the ample evidence of self-conscious groups engaged in a power struggle over the shape of job ladders within an organization. It would be hard to understand the resulting ILM as simply the result of an optimizing
process by managers who survey the technology and tasks and decide upon the most efficient way to structure work.

The Slow Diffusion of High Performance Work Systems

High Performance Work Systems, which involve the use of teams and various forms of quality programs, have been implemented in a range of firms (Osterman, 1994, Osterman, 2000; Handel and Gittleman, 2004) because there is considerable evidence that they lead to higher levels of quality and productivity (Pfeffer, 2007, MacDuffie, 1995, Ichniowski, Shaw, and Prennushi, 1997, Kochan and Osterman, 1994). Yet despite this evidence adoption has been slow which implies, as Pfeffer has observed, that money has been left on the table. Even when adopted the final form that the work organization takes is often not in line with the initial expectations of management.

It is also important to understand that the role of internal political struggles between groups of employees and management is apparent in accounts of adoption or the lack thereof. In a case study of non-union and white collar Hewlett-Packard Michael Beer and Mark Cannon examined why an effort to implement pay-for-performance compensation failed. They found that teams of employees resisted because of fears of unexpected changes in their compensation levels and that their resistance took the form of refusing new job assignments and refusing to admit new members to their teams. Beer and Cannon conclude that the process was one of “implicit negotiation” and that “high commitment [work systems] can only be created if employees…feel fairly treated and this in turn is a function of how much voice they have. (p. 14, 17).” Other non-union
examples include the observation that first line supervisors resist HPWO because they feel threatened by the role of self-managed teams in taking over some of their functions (Coyle-Shapiro, 1999).

Case studies in unionized settings reach the same conclusion. The poster-child is the automobile industry in which tensions between local unions and management often delayed or reshaped the implementation of high performance work systems (Kochan, Lansbury, and MacDuffie, Landsbury, 1997). More broadly the literature reports similar issues. Unions can be oppositional when their adversary function is threatened by the more cooperative nature of HPWO systems (Walton, McKersie, and Cutcher-Gershenfeld, 1994) and management resists giving up its power to groups of employees (a useful review of failed HPWO partnerships in unionized settings is provided in Chapter 2 of Kochan, Eaton, McKersie, and Adler, 2009).

The case study research is supported by survey methods although the surveys cannot by their very nature provide as textured a picture. For example, in their survey of the adoption of HPWO practices across a sample of automobile firms Pil and MacDuffie (1996) find that variables capturing the presence or absence of internal organizational barriers and resistance perform better than simple measures of economic performance. And as the Hewlett-Packard case makes clear, this pattern extends well beyond the unionized sector of the job market. Whether and how a firm adopts these practices is due to much more than simply an optimizing decision by senior managers.

**Effort and Commitment**
Obtaining higher levels of employee commitment has emerged as important in markets where quality and customer satisfaction are key to competitive success. As a consequence another recent development is new strategies by firms aimed at addressing this need.

The New Personnel Economics addresses employee effort and commitment from the perspective of efficiency wages and incentive compensation models that are grounded in principal-agent theory. These perspectives are somewhat different in that efficiency wage/shirking models are about level of effort whereas in principal-agent models the worry is that the effort may be in the wrong direction, i.e. for the self-interest of the agent rather than the principal. However, what both flavors have in common is the view that effort is largely decided at the level of the individual and that the key to any solution is to get that individual’s incentives correct.

An empirical illustration of this perspective, and evidence that it has power, is Lazear’s research on Safelite glass (Lazear, 2000). This paper, which is prominently cited in the literature, discusses the consequence of introducing piecework pay systems for auto-glass installers. The research showed that the installers responded to these incentives and that there was a selection process in which the less able employees quit the organization. The point then is that people do indeed respond as predicted to financial incentives but it is also important to remember that the circumstances of Safelite were special. Output was easily measured and monitored and employees worked individually, not in teams or groups. Nonetheless, one can take the point that incentives are important.
Although it is clear that incentives as conventionally understood play a role in eliciting effort group processes of the sort emphasized by the institutional perspective are also central. The level of effort delivered by the workforce is at least partially determined at the group level and by considerations other than incentive schemes. Consider Toyota or Southwest Airlines, both of whom are probably taught in Human Resource courses even more frequently than Safelite. These firms are tremendously successful in their industries and both are widely known for obtaining high levels of effort from their workforce. They are also successful in obtaining organizational commitment, an idea that is basically foreign to New Personnel Economics thinking (Lincoln and Kalleberg, 1990). But, in contrast to NPE expectations, both are quite conventional in their incentive systems (and this is as true of Toyota of America as it is of Toyota of Japan) and in both organizations seniority (frequently seen in the NPE literature as the enemy of meritocracy) is important. Why is it that these organizations are able to obtain exceptional effort and organizational commitment while other firms with essentially the same compensation arrangements are not? The answer lies in factors that institutionalists emphasize: group norms and culture.

In thinking about this it is important to recognize that norms and culture are group phenomena and not examples of the kind of individual level social-psychological exchange discussed earlier. It is also important to contrast this perspective with the individualist viewpoint typical in the NPE literature. Consider, for example, the comment made by John Roberts in his widely cited book on the internal economics of organizations: “While motivation problems can be manifested at the group level, we focus our discussion on the problems of motivating individuals because these are
logically prior and because most of the arguments are more directly made in this context” (Roberts, 2004, p. 177). This perspective misses the fundamental point that group behavior and group objectives are not simply the aggregation of the behavior and objectives of individual actors.

There is a long tradition in institutional research of recognizing that group culture and norms are important. Seashore (1954) showed that norms operate at the group level, Dunlop (1957, p. 130) emphasized the importance of custom in both the short and long-run process of wage determination, and Roy (1959) described how group processes lead to withholding of effort. In its more modern form organizational culture is a rapidly growing research field (Barley, Meyer, and Gash, 1988; Martin, 1992).

In both Southwest Airlines and Toyota the employees have developed strong group commitments to the success of the organization (Womack, Jones, and Roos, 1997; Pfeffer, 1998; Gittell, 2003; Rubinstein and Kochan, 2001). It is important to avoid romanticism in that Southwest has had labor trouble on occasion with its pilots and Toyota has worked very hard to keep out the United Automobile Workers. Nonetheless it is clear to observers (see the citations above) that at the level of the workgroup each organization has succeeded in obtaining effort via broadly shared norms and not simply via individual incentives or individual gift-exchange.8

There is also a very large ethnographic literature that generalizes the point of the Southwest and Toyota examples and suggests that the reality is that the workforce essentially engages in a bargaining process with management over effort, a negotiation that can be informal and is not linked to formal collective bargaining or even the presence
of a union. Culture and norms represent the “opening” position of the workforce and what ensues is implicit negotiation between employees and management over effort.

We have already seen an example of this in the Gouldner study (Gouldner, 1954, pp. 161-162) cited earlier in which he showed that employee views about the legitimacy of management specific management practices varied and that if the practices (i.e. ILM rules) were seen as illegitimate then effort was withheld. This pattern was found to be important in a very broad range of published ethnographic workplace studies surveyed by Hodson (1999). The ethnographies represented in the survey were undertaken in varied settings both unionized and non-unionized factories as well as white collar settings such as banks and restaurants. Hodson systematically compiled and coded the reports and examined the relationship between management adherence to work group norms and the degree of effort provided by the labor force. By norms he meant group level expectations about fair treatment as well as expectations regarding the organization of work and of production. By effort he meant both what he termed “resistance behaviors” such as acting dumb and sabotage, as well as “citizenship behaviors” such as offering ideas, training peers, and providing extra effort. The coded research demonstrated a strong and significant relationship between management adherence to work group norms and citizenship behaviors. This represents strong evidence of implicit bargaining between groups of employees and management over a core element of the ILM.

Strategies For Using Contingent Employees
The use of temporary workers has surged in recent years and is an important element in the fraying of closed ILMs (Autor, 2003, Hausman and Osawa, 2003). One explanation that is consistent with the NPE approach is that firms often use contingent work as a strategy for working around compensation rigidities that result from the persistence of fixed wage differentials based on work-group ideas of fairness (Abraham and Taylor, 1996). Related to this is the view that contingent employment enables firms to avoid the constraints that some courts have imposed via restrictions on the employment at will doctrine (Autor, 2003). What ties these views together is that contingent work is seen as an effort to avoid ILM rules and to return to a more classical labor market. Again, as in our earlier discussion, there is certainly good reason to believe that these considerations are important. However, as before, in practice they are in a political contest with more institutional considerations.

An institutional perspective would begin by observing that the motivations for contingent work outlined above have their origins in the strength of institutional constraints on the firms’ employment practices and hence the spread of contingent employment is in a sense an indirect acknowledgment of the power of institutional considerations. However, institutional considerations are more than a backdrop against which contingent work is played out, they shape how precisely is contingent work used, i.e. the rules that govern its implementation.

One example is found in the ethnographic research of Smith and Neuwirth (2008) who studied how a temporary help firm interacted with both the HR staff and the line managers of a client firm. How the temps were utilized was the result of a complicated
three way political process within the organization with each interest group having a distinct perspective. The HR group struggled to maintain its role and status, line management sought flexibility, and the temp agency not only searched for business but also negotiated the relationship between the two internal factions. Smith and Neuwirth write, “What appear to be purely market-mediated employment relationships are constructed in the negotiations and occasional struggles between HR staff, line management, and personnel from labor market intermediaries. (Smith and Neuwirth, 2008, p. 143).”

Similar struggles occur between managers and regular employees. When firms use temps they have to pay attention to the reactions of the regular workforce and the specifics of the ILM rules around temps are often shaped by the need to maintain status differentials. In her study of a Boston area call center and a manufacturing firm Lautsch (2002) observed that the balance of relations between the two employee groups was a central concern of managers as they decided how to implement contingent work. She notes that “regular workers favored the maintenance of status differences between themselves and temporary workers” (Lautch, p. 30). Management responded by “asserting that management would limit the scope of contingent work and its impact on regular jobs.” (Lautch p. 32). She adds that “management learned…that regular workers would rebel if teams were set up so that temporary workers had majority membership, or so that temporary workers were in supervisory roles over regular staff. Gradually these practices were eliminated.” (Lautch p. 34) In short, efficiency considerations were important and managers pressed forward with the use of temps as a
source of cost savings but the manner in which they were used and rules governing their work were significantly shaped by institutional factors.

**Due Process and Equal Opportunity**

Although little remarked upon by NPE scholars, one of the most widespread shifts in ILMs—a shift certainly as important and as widespread as the diffusion of performance based pay—has been the implementation of a broad range of personnel practices associated with equal employment opportunity and with non-union due process and grievance procedures. These practices obviously have important implications for career pathways and promotions within firms.

A large literature has emerged seeking to explain the origins of these practices within firms. Much of this literature is rooted in the “institutional” stream of thought in Organizational Sociology, a line of thinking that focuses on how organizations adopt practices in order to legitimize themselves in the eyes of other powerful actors or because the practices enhance the power of interest-groups within the organization (Selznick, 1949; DiMaggio and Powell, 1983). This latter perspective is clearly congruent with the argument developed in this paper. The challenge in this literature is to distinguish the institutional explanations from more efficiency based arguments that the adopted practices make economic sense for the organization.

Although it is not always possible to definitively prove that efficiency considerations are unimportant, the body of research taken as a whole does convincingly demonstrate that institutional considerations played a central role in the diffusion of the
new practices and in determining the specific forms that they have taken. In an early
foray into this field Edelman (1990) studied the adoption of grievance procedures in non-
union firms, specifically in a sample of 52 firms in the Bay Area. Her broad argument
was that organizations adopted these practices not because of any specific legal mandate
(there was none) nor because of unions (the firms were non-union and a variable
measuring whether they faced a union threat was insignificant). Rather, the process
appeared to be that the general normative environment coming out of the Civil Rights
movement put heavy moral or societal weight on firms to build “fairness” into their
procedures and that firms with personnel departments were more likely to respond to this
pressure.

In later research Edelman, Abraham, and Erlanger (1992) put even more emphasis
on the role of personnel departments in pushing for practices that were self-aggrandizing.
In this instance they studied the adoption of practices related to wrongful-dismissal
charges. They show that the personnel profession as a whole systematically exaggerated
the risks that firms faced in these law suits, in some cases using estimates that were
“absurd” (p. 65) and in other cases simply misleading (by, for example, reporting data
from California as if it applied to all states). As result, senior managers were convinced
to implement personnel practices that enhanced the power and role of the personnel
professionals.

In another set of studies Dobbin, Jennings, Scott, and Meyer (1993) and Dobbin
and Sutton (1998) study the adoption of promotion policies in response to Equal
Employment Opportunity mandates from the Federal Government. There is certainly an
underlying efficiency explanation in that firms faced the risks of lawsuits and contract
loss if they did not respond in some way. However, in understanding what specific practices emerged there are important subtleties. First, Dobbin, Jennings, Scott, and Meyer (1993) show that the particular practices adopted “converged on a set of personnel practices that were isomorphic with the procedurally oriented, quasi-judicial administrative configuration of the federal government-formal, merit-based, employment and promotion conventions complete with an internal system of grievance adjudication.” In other words, the particular practices that were adopted were based upon imitation of those of the more powerful external organization, not upon an internal optimization logic. Of course, an NPE response to this might point to the search and information savings that such an approach might bring. However, Dobbin and Sutton (1998) demonstrate that even after Federal pressures sharply diminished in the Reagan years, and hence the costs of non-compliance sharply fell, personnel professionals continued to press for new practices and that they shifted their rhetoric from that of legal risk to the alleged internal benefits of diversity. Put differently, the personnel staff constructed (without any evidence) a new argument once the legal risk argument was no longer valid.

A fair reading of this literature is not that efficiency explanations were unimportant or irrelevant but that they are strikingly incomplete and perhaps even naive with respect to the forces that drove adoption of the wide range of personnel practices associated with equal employment opportunity and fairness in the workplace. The self-interest of a specific interest group within the firm, as it competed for resources and power, was a big part of the story.
CONCLUSION

This paper has aimed to describe an institutional perspective on ILMs that recognizes the role of standard efficiency and market considerations yet brings to bear a deeper understanding of how organizations make decisions about organizing work. One central characteristic of the argument is that it does not develop a universal model parallel to the universalistic tendencies of neo-classical economic theory. Rather this paper is pitched at the level of the firm or organization. Nor does it address the psychology of individuals (i.e. whether or not they are utility maximizers) or the operation of external markets. The creation of employment rules within organizations is the terrain on which I think that institutional ideas have the most purchase.

It is important in this context to understand that even at the level of the organization the approach proposed here does not reject market forces, this would be foolish indeed and particularly so in a period when established forms of work organization are under a great deal of pressure from competition. Rather the goal is to develop a more sophisticated understanding of the dynamics of Internal Labor Markets. The central proposition, one which as noted has a long tradition in the literature, is that these systems are the result of a political process in which competing objectives and norms held by groups play out a contest. Market or efficiency considerations are certainly one player in this contest and in the examples I presented I made an effort to give them their due. But they are not the only player. It is also important to recognize that the resources available to different positions in this contest will vary over
time. It would not be unreasonable, for example, to argue that in the Depression and World War II period the external environment favored internal actors who opposed NPE style thinking whereas today cost minimization advocates hold an advantage. But this ebb and flow does not alter the fundamental argument.

At the same time that market forces are given their due it is also important to understand that the institutional considerations emphasized here are persistent and are not simply transitory aberrations that will be quickly eliminated by the force of competition. First, the market has proved to be very slow in extinguishing long-standing practices as the history of the automobile industry makes clear. In addition important streams of research and theory in economic models of technology adoption, the sociology of organizations, and political science models of institutional change all argue that the forces of path dependence are strong and that practices become locked in.¹⁰

At the same time the emphasis on group processes and politics point to an important limitation of this paper: the question, which is not addressed, of how groups are formed and re-reformed. This, of course, goes to the issue of how people and collectivities come to identify themselves as having commonalities and how they frame what they regard as legitimate interests and goals. This is a topic of growing interest (Barley, 1989; Scully and Segal, 2002; Piore and Safford, 2006;) but it would take this paper too far afield to enter into it.

The goal of good social science should be to understand the determinants of the empirical realities that we observe, both at a point in time and as they change. Elegant theory is not an end in itself. This is important to understand because the arguments
developed in this paper are certainly less clean and elegant than are NPE models. Furthermore they run the risk of devolving into what has been a weakness of institutional theory, namely a tendency to simply provide a long list of factors that should be taken into account without any strategy for assigning relative importance or for testing the arguments. However, to say that institutional theory runs this risk is not to say that it should be abandoned. Rather, advocates of institutional models have to be willing to be clear about their arguments and the strategies for testing them. The gains from such an effort should be a richer understanding of how employment systems are evolving and what drives this evolution.
REFERENCES


Baron, James, M. Diane Burton and Michael Hannan, “Engineering Bureaucracy: The


Dunlop, John Wage Determination Under Trade Unions (New York: Augustus Kelley), 1950


Gouldner Alvin, Patterns of Industrial Bureaucracy (Glencoe: Free Press), 1954


Hannan and Freeman, “Structural Inertia and Organizational Change,” American Sociological Review, 1984

Hodson, Randy, “Group Relations At Work,” *Work and Occupations*, vol. 24, no. 4 (November, 1997), pp. 426-452


Jacoby, Sanford, “The New Institutionalism: What Can It Learn From The Old?”

*Industrial Relations*, vol. 29, No. 2 (Spring, 1990), pp. 316-340


Mahoney, James, “Path Dependence In Historical Sociology,” *Theory and Society*, 2000, v. 29, pp. 507-548


Myers, Charles and George Shultz, *The dynamics of a labor market: a study of the impact of employment changes on labor mobility, job satisfactions, and company and union policies*. New York, Prentice-Hall 1951


Pfeffer, Jeffrey, *The Human Equation: Building Profits By Putting People First* 


Prendergast, Candice, “The Provision of Incentives In Firms,” *Journal of Economic Literature*, vol. 37, 1999, pp. 7-63


Seashore, Stanley, *Group Cohesiveness In The Industrial Workgroup*, (Ann Arbor: University of Michigan Survey Research Center), 1954


There is considerable empirical work that demonstrates diversity. At the anecdotal level are examples such as the contrast in employment practices between Wal-Mart and Costco (Greenhouse, 2005; Cascio, 2006), the divergence of Southwest Airlines from the practices of other large carriers, or the HRM practices of the software firm SAS compared to its competitors (Pfeffer, 1998). More systematic work within industries, such as automobiles and telecommunications, also points to the same conclusion. For example, in their study of telecommunications and automobiles Katz and Darbishire (1999, p. 4) report that “within both union and non-union sectors the extent of variation in wages, work practices, and other employment conditions has increased.” In a review of the literature on call centers Batt and Moynihan (2002) identify alternative approaches towards organizing work (which they label “mass production,” “professional service,” and “mass customization”). Drawing on a survey of 354 American call centers they find that these models are all represented and the diversity of practices remained after the sample was divided into those centers serving residential and those serving business customers. Additional evidence of diversity of practices within customer segments is provided in Batt (2002).

---

1 There is considerable empirical work that demonstrates diversity. At the anecdotal level are examples such as the contrast in employment practices between Wal-Mart and Costco (Greenhouse, 2005; Cascio, 2006), the divergence of Southwest Airlines from the practices of other large carriers, or the HRM practices of the software firm SAS compared to its competitors (Pfeffer, 1998). More systematic work within industries, such as automobiles and telecommunications, also points to the same conclusion. For example, in their study of telecommunications and automobiles Katz and Darbishire (1999, p. 4) report that “within both union and non-union sectors the extent of variation in wages, work practices, and other employment conditions has increased.” In a review of the literature on call centers Batt and Moynihan (2002) identify alternative approaches towards organizing work (which they label “mass production,” “professional service,” and “mass customization”). Drawing on a survey of 354 American call centers they find that these models are all represented and the diversity of practices remained after the sample was divided into those centers serving residential and those serving business customers. Additional evidence of diversity of practices within customer segments is provided in Batt (2002).

2 The idea that organizational behavior/decision making is not of the optimizing “machine model” mode has certainly been put forward before in the literature by both economists and sociologists. Cyret and March (1963), in the Behavioral Theory Of The
Firms note that organizations are coalitions and that “basic to the idea of a coalition is the expectation that the individual participants in the organization may have substantially different preference ordering (i.e. individual goals)....with the obvious potential for group conflict” (p. 31). However, the book itself emphasizes different goals and coalitions among functional units as they make business decisions (production, accounting, sales, etc.) and devotes only two pages (186-187) to employment issues.

3 Other efforts to introduce “social” considerations into standard models are also similarly limited. For example, Robert Solow (1990) argues that the labor market is a “social institution” but his emphasis is basically upon variants of efficiency wage models of wage determination and the macro-economic consequence of the fact that individuals respond to fairness considerations and salary norms. There is no discussion of groups, group conflict or competing visions of the objectives of the firm (Solow, 1990).

4 quoted in Parsons, 1963

5 A similar comment, directed to economic modeling more generally, was made by Granovetter:

“...the main thrust ... is to deflect the analysis of institutions from sociological, historical, and legal argumentation and show instead that they arise as the efficient solution to economic problems. This mission and pervasive functionalism it implies discourages detailed analysis of social structure that I argue here is the key to understanding how existing institutions arrived at their present state” (Granovetter, 1985 p. 505)

6 In another important contribution Zald and Berger (1978) applied social movement theory to organizations and identified three types of political conflict within
organizations, coup d’états, bureaucratic insurgency, and mass movements. The first type represents a struggle for power among the elite and is not relevant for the arguments developed here and the third is too episodic to be relevant. However, bureaucratic insurgencies can indeed represent on-going struggles over the purposes and procedures of the organization that are driven by competing views of appropriate objectives held by different groups.

There is an NPE literature on organizational politics but it remains at the level of individuals and how they maximize their self interest rather than at the level of group political contests over rules. For example, Lazear (1989) examines whether wage compression is good or bad for efficiency given that it may enhance cooperation but perhaps also reduce incentives for the best employees. However, his analysis asks how individuals react to different wage distributions with respect their private self-interest and he maintains the view that there is a decision maker within the firm who can set the optimal wage structure. In another effort along these lines Prendergast and Topel (1996) model the behavior of supervisors who exhibit favoritism towards particular employees. But, again, the modeling is purely at the individual level. Supervisors simply “like” some workers and not others, there is no sense of group action or interest. Finally, Gibbons and Waldman discuss politics but again it is purely individual. They write, “In our usage political models are slightly different from agency models. Both describe interactions between individuals at different levels of a hierarchy...[but] in political
models the supervisor’s actions cannot be governed by a contract…” (Gibbons and Waldman, 1999, p. 2409).

8 The understanding that culture and norms are important in eliciting effort may be foreign to NPE but is increasingly shared in the broader economics literature. However, there remains a tendency to cling to the notion of a decision maker within the firm who, with this more sophisticated understanding, optimizes the firms organizational structure and culture to obtain effort (Akerlof and Kranton, 2005).

9 I am grateful to Sandy Jacoby for clarifying this point for me.

10 See, for example, Nelson and Winter (1982), Arthur (1989), Hannan and Feeman (1984), Stinchcomb (1965), and Thelen (2004).