



Ackerman, Frank. "Overview Essay: New Directions in Labor Economics," Part II, Frontier Issues in Economic Thought, Volume 4: The Changing Nature of Work. Island Press: Washington, D.C. 1998. pp.15-25.

Social Science Library: Frontier Thinking in Sustainable Development and Human Well-being

“New Directions in Labor Economics” by Frank Ackerman

How, if at all, is labor different? Is it a unique entity that calls for a unique economic analysis, or just a case in which the well-established theory of market behavior explains yet another mass of data? Odd as it may seem to outsiders, the latter view has its supporters. As Robert Solow has said,

Among economists, it is not obvious at all that labor as a commodity is sufficiently different from artichokes and rental apartments to require a different mode of analysis.¹

Fortunately, there are also many economists (including Solow) who do believe that labor is different enough to require its own analysis. Indeed, the theory of labor economics is one of the most active and creative fields of economics; this essay and the summaries that follow it will identify some of the leading new themes that have emerged in recent years.

Labor economics is also an area with an unusually close connection between theory and empirical evidence. Systematic, quantitative evidence is available in the form of massive computer databases describing individual workers and jobs. More than in many other areas of economics, it is common for innovative theoretical articles -- including several that are summarized here -- to include empirical tests of their theories.

The need for new theories flows directly from the obvious facts about labor. Two related types of evidence make it clear that labor is not just another commodity, and cannot usefully be understood as if it were simply a particularly prickly form of artichoke. First, the process of wage determination and its relationship to employment is unique; labor markets are indeed different from other markets. The most striking fact is the persistence of unemployment: even a substantial level of involuntary unemployment leads to only gradual and modest changes in the wages of employed workers and the total supply of labor. (Artichokes do not experience unemployment; if there are too many on the market, the price and/or quantity supplied will fall relatively quickly to the market-clearing level.) Among workers who are employed, there is persistent inequality in wages and salaries, which is only partially related to visible differences in skill requirements or the inherent difficulty of the work being done.

A second type of evidence for the uniqueness of labor concerns the nature of work itself. Unlike other inputs into production processes, workers are conscious of what they are doing. Thus the subjective experience and involvement of the workers cannot be ignored. Systems of labor relations reflect the unique character of work; many employers as well as workers accept union contracts, seniority-based pay scales, and other long-term employment commitments that bear no clear or direct relationship to the marginal product of individual workers. (Needless to say, artichokes, machinery, and other inanimate resources do not demand or get similar treatment.)

These categories of facts, which are discussed throughout this volume, make it hard to accept the simple market paradigm, and thus provide much of the motivation for the creation of new approaches in labor economics.

THE HISTORY OF LABOR ECONOMIC

Before turning to contemporary developments, it will be helpful to take a look at the history of labor economics.² In the earliest stages of economic theory, it was taken for granted that labor was both unique and central to economic life. Classical economists from Adam Smith through Karl Marx relied on the labor theory of value, and examined at great length such questions as the division of labor, the level of employment, and the determination of wages.

The view of labor as just another commodity first became possible with the “marginalist revolution” of the 1870s. The labor theory of value was replaced by the idea that values and prices both depended on marginal utility; the formalization of general equilibrium reflected a fascination with the symmetrical treatment of all markets. The leading figure of late nineteenth-century economics, Alfred Marshall, played a contradictory role in this, as in other areas. On the one hand, Marshall developed the analysis of the marginal productivity of labor, which made it possible to extend the theoretical formalism of supply and demand to the labor market. On the other hand, he wrote in more practical, applied terms about the social, institutional, and political aspects of the labor questions of the day.

Later theorists generally chose one or the other of Marshall’s contradictory leanings. In the 1930s the macroeconomics of John Maynard Keynes acknowledged the significance of unemployment and examined the factors affecting wages, emphasizing labor’s uniqueness and institutional specificity. At the same time, the elaboration of microeconomic theory by John Hicks and Paul Samuelson completed the formal symmetry of the market model, later epitomized by Samuelson’s famous remark that, in theory, it makes no difference whether capital hires labor, or the other way around.

Such microeconomic theories were not shared by the first few generations of American labor economists. In the early twentieth century the institutionalist school, including John Commons, Richard Ely, and their successors, produced many important studies of labor problems. They typically focused on questions of equity for labor, and were supportive of unions and of government intervention in the market on behalf of workers; their work had little connection to the simultaneous development of neoclassical theory. In the 1940s and 1950s a later generation

of economists, including John Dunlop and Clark Kerr, used neoclassical tools to analyze institutional issues concerning labor, addressing the problems of imperfect competition, bargaining structures, wage contours, and internal labor markets.

In short, at midcentury there was a diverse, pluralist understanding of the economics of labor. While the complex mathematics of the simple market model was of growing importance in microeconomics, a Keynesian view of unemployment and wage determination held sway in macroeconomics, and a variety of institutional insights into bargaining and wage structures remained central in labor economics itself. The labor movement was at its peak of membership, representing one-third of the paid workforce in the U.S. in the 1950s; thus a theoretical focus on the bargaining process seemed only natural.

THE MARKET MODEL AND THE NEW DISSENT

As in other areas of economics, the neoclassical microeconomic model eventually came to dominate over alternative theoretical approaches to labor. Yet only in the 1960s, with the addition of two innovations, human capital and search theory, did neoclassical theory begin to offer a plausible account of employment, unemployment, and wage determination.

Discussion of the role of skills in explaining income differentials dates back at least to Adam Smith. In the 1960s Gary Becker formalized this idea in the notion of human capital: education and training could be seen as investments in workers' skills, which, like capital goods, are then available for use in production. At about the same time, George Stigler suggested that time and effort spent in searching for a new job could also be viewed as a productive investment; a longer search leads to a greater probability of a higher return.³ As a result, it appeared that many types of inequality, and many types of unemployment, could be explained in conventional market terms without reference to institutional or historical factors. If not quite artichokes, perhaps workers are like machine tools, whose varying degrees of sophistication, age, and obsolescence account for their differential rewards.

These theories have led to valuable insights -- but they also threaten to exaggerate the explanatory power of the simple market model. It is one thing to say that observable investment in education and training explains *some* observable patterns of employment and wages; it is another thing to assume that there must be unobservable differences in human capital that explain *all* differential labor market outcomes. Likewise, productive investment of time and effort in searching for a new job explains selected aspects of unemployment, but it defies common sense to interpret all unemployment in this manner. However, defying common sense often seems fashionable in economic theory. This was never more true than in the heyday of "rational expectations" and "microfoundations" models in the 1980s, when the principle of well-informed individual maximization was axiomatically assumed to explain virtually all behavior, and involuntary unemployment was defined out of existence.

Fortunately, a new wave of dissent, and a variety of theoretical innovations, have appeared in labor economics in recent years. Some of the new approaches preserve most of the neoclassical mathematical apparatus of individual maximization, showing that even modest changes in the

assumptions can lead to different, more realistic outcomes. Others question the adequacy of the assumptions of individualism and maximization, raising more fundamental, if messier, questions about the direction of economic theory. A sampling of these new approaches are represented in the summaries following this essay, to which we now turn.

WORK INCENTIVES, EFFICIENCY, AND UNEMPLOYMENT

The strength of the new approaches to labor economics rests on their ability to explain a range of facts which fit awkwardly, if at all, into the standard market model. Thus it is appropriate to begin our look at the theories with a massive empirical study, the “wage curve” of David Blanchflower and Andrew Oswald. The wage curve addresses a deceptively simple question: is higher unemployment associated with higher or lower wages? The standard model of supply and demand suggests that if wages are raised by unions or minimum wage laws, employers’ demand for labor should drop (since the demand curve is downward-sloping), and unemployment should rise; in other words, higher wages should be associated with higher unemployment.

Blanchflower and Oswald show, however, that the opposite is true, in a study encompassing data on literally millions of workers in a dozen countries. They find that, when controlling for a vast range of individual and industry characteristics, workers in regions of a country with higher unemployment have lower wages than comparable workers in regions with lower unemployment. This finding, consistent with the results of other investigators, shows that something “non-neoclassical” is going on in labor markets. As Blanchflower and Oswald explain, their results are compatible with models of bargaining strength for countries with high rates of unionization, or with “efficiency wage” models for less unionized countries such as the U.S.

The efficiency wage model is the best-known recent innovation in labor economics, and has been widely studied and debated. The essential idea is that if payment of higher wages has a positive effect on a worker’s productivity, then employers face contradictory incentives; their usual desire to cut costs will be tempered by the recognition that lower wages mean lower productivity. This idea can be traced back to Harvey Leibenstein’s 1963 suggestion that in the poorest of developing countries, higher wages lead to better nutrition and hence to increased productivity. As a model of labor markets in developed countries, the efficiency wage theory first gained prominence in the 1980s.⁴

The most common version of the theory assumes that it is costly or difficult for employers to observe employee effort on the job, and that employees in general want to shirk, or avoid effort, at work. A worker weighing the costs and benefits of working hard will compare the joys of laziness with the probability of being caught shirking and the costs of being fired (the assumed result of being caught). A higher wage increases the cost of losing the current job, and hence makes it more attractive to work hard. An alternate, but compatible, explanation assumes that payment of above-market wages will induce worker loyalty to the firm, leading to greater effort. In either case, firms find it profitable to pay some workers more than their full-effort marginal product, i.e. more than they add to production, even when working as hard as possible. This

means that the demand for labor is below the market-clearing level, implying that there is involuntary unemployment.⁵

In the next article summarized here, Sam Bowles presents a model formalizing aspects of Marxian theory, using the mathematical techniques of neoclassical economics. The efficiency wage mechanism is central to Bowles' model of the firm, since profit-maximizing employers face a tradeoff between increased surveillance of workers and increased wage payments, as alternate means of extracting additional work effort. Bowles extends the theory to its implications for the choice of technology, and explains the functional role of unemployment, from the employers' point of view, in terms consistent with Blanchflower and Oswald's findings. Discrimination, too, is functional for employers, since it reduces labor unity and bargaining strength. Although relentlessly pursuing the mathematics of individual maximization in developing his model, Bowles also stresses the unique and social character of the work process. He is perhaps least persuasive in differentiating his version of efficiency wage theory from the very similar approaches that have appeared in recent non-Marxian theories of the firm.

Does the individualism of the efficiency wage model misrepresent the essential nature of the labor process? This question is raised in the article by Robert Buchele and Jens Christiansen. They observe that many workers, especially the best-paid, are protected by union agreements, seniority systems or other mechanisms that limit the threat of individual firing based on individual performance. Moreover, individual effort is not always the decisive factor in determining productivity. The effective organization of individual effort into a productive, collective process is at least as important.

For Buchele and Christiansen, a cooperative system of industrial relations can elicit worker participation, a source of productivity that is not available in an antagonistic system. They present a comparative analysis of data on the leading industrial countries, suggesting both that cooperation boosts productivity, and that, in a cooperative regime, lower unemployment is good for productivity, while the opposite is true in an antagonistic setting. In more recent empirical work Buchele and Christiansen have used this framework to study the tradeoff between employment growth in the U.S. and productivity growth in leading European economies, suggesting that deregulating and "Americanizing" European labor markets could lead to greater insecurity and inequality, rather than to economic growth.⁶

Like Bowles, Buchele and Christiansen (and a number of other authors represented in this volume) are members of the school of "radical political economy" that emerged as part of the general political and cultural radicalism of the late 1960s and 1970s. In the next summary, James Rebitzer reviews the contributions of this school of thought to labor economics. Several of his themes echo the Bowles article, including the theory of the firm, the efficiency wage model, and the potential for adoption of technologies that enhance control over workers, even at the expense of efficiency. The important question of employee ownership, raised by Rebitzer and many of the authors he reviews, is discussed in much greater detail in Part 6 of this volume. While Rebitzer is generally sympathetic to the radical political economy school, he also makes a

number of important criticisms of its work, including a tendency to oversimplify the role of management, the actual process of bargaining, and the nature of work incentives.⁷

LABOR MARKET SEGMENTATION

The second major area addressed by Rebitzer, theories of labor market segmentation, is also the subject of the next three summaries. Segmentation theories typically assume that the labor market is divided into two sectors: a primary sector with relatively high wages, long average job tenure, rewards for skills and experience, and chances for advancement on the job; and a secondary sector with none of these characteristics. As with efficiency wage theory, the boundary between radical and mainstream segmentation theories has blurred in recent years (to the benefit of both, as Rebitzer notes).

“Dual labor market” theories attracted attention in the 1970s, as a possible explanation of inequality. The theories typically suggested that most members of racial minorities, and many women, were stuck in the secondary labor market and denied access to good jobs. An influential early account of dual labor markets from an institutionalist perspective, by Peter Doeringer and Michael Piore,⁸ was extended by a number of radical economists exploring issues of discrimination. Yet the explanation of the basis for segmentation has been problematical; as Rebitzer explains, the most popular explanation, efficiency wage theory, does not fit some of the crucial facts.

A critique of labor market segmentation has suggested that the early theories could not explain the number of segments or the assignment of workers to those segments. Indeed, many economists have concluded that human capital theory provides a better explanation of inequality of labor market outcomes. The next article, by William Dickens and Kevin Lang, responds to this critique and insists that a model of market segmentation is superior to human capital theory in accounting for inequality. The defining characteristic of the theory, for Dickens and Lang, is not the inequality between better and worse segments of the market, but the nonmarket rationing of access to good jobs. This need not imply lifelong assignment of individuals to one or the other market segment; indeed, some patterns of racial inequality are consistent with the assumption that minorities spend more time queuing for access to primary-sector openings.

Using detailed individual data on male workers, Dickens and Lang estimate earnings functions for the primary and secondary sectors, based on education, race, age, location, and other factors. The distinctive feature of their model is that, while estimating earnings functions, it simultaneously assigns each worker to the segment in which he fits best. As predicted, the two-sector model fits the data much better than any one-sector model; returns to education and experience are substantial in the primary sector, while minimal in the secondary sector; and the definition of the sectors, while subtler than in most theories, follows the expected demographic patterns.

Dickens and Lang argue effectively against explanations of inequality based solely on human capital. But they, like Rebitzer, suggest a diverse range of possible explanations for segmentation. The next two articles explore two of those explanations in greater detail. Robert Drago and Richard Perlman present a “competing incentives” theory, in which employers must choose between two incompatible routes to worker motivation. One alternative is based on a high level of trust, requiring high wages and minimal supervision; the other involves little trust, low wages, and intense supervision. Drago and Perlman propose that the heterogeneity of technology determines which regime is more profitable for which firms. The choice facing employers is a discontinuous one; small moves toward the opposite system of motivation will simply reduce the efficiency of production. Thus there are two labor market segments because there are two incompatible systems of motivation; one or the other is more effective for each production technology. Even if workers are identical, many will queue for opportunities to enter the better jobs.

George Akerlof and Janet Yellen, two economists who have played an important part in the development of efficiency wage theories, examine the role of norms of fairness in the next article. In psychology, sociology, and elsewhere outside of economic theory, it is taken for

granted that people have strong feelings about fairness, and that norms of fairness are influenced by the observed treatment of others. Akerlof and Yellen assume that workers have a notion of fair wages, and reduce their work effort when wages fall below that level. If the standard of fairness depends in part on what one's co-workers receive, then, in an enterprise with varied skill levels, the low-skilled workers will set their standard of fairness above the market-clearing wage for their labor. As Akerlof and Yellen explain, this leads to higher than market wages, and lower demand, for low-skilled labor in major enterprises⁹; the result is either involuntary unemployment of lower-skilled groups, and/or the creation of low-wage firms that employ only low-skilled workers.

THE PERSISTENCE OF INSTITUTIONS

Many of the recent innovations in labor economics preserve the methodology, though not all the underlying assumptions, of neoclassical theory. That is, an austere, minimalist set of assumptions are made about individual options and objectives, allowing a rigorous mathematical exploration of the implications of individual maximization subject to constraints. It is impressive that, with slight modifications of the assumptions about information flows and the nature of the firm and the labor market, this methodology produces much more realistic outcomes.

Yet at the same time there is something frustrating about this approach to theory. When the conclusions are more obvious than the proofs, how much value has been added by the theory? Several pages of calculus can "prove", for example, that rational employers do not always cut their employees' wages when sales decline, for fear of damaging morale and productivity -- or that more generous unemployment benefits allow laid-off workers to take longer looking for new jobs. Many non-economists would readily accept these and other theoretical results, while dismissing the proofs as hopelessly opaque.

Moreover, it is worth noting that the realism of the new models is less than complete. In some cases, the new models look like a success; in other cases, the narrowness of the assumptions still seems to lead to artificially narrow conclusions, suggesting the need for a broader theoretical framework.

The last three articles summarized in this section address a range of institutional questions, which remain crucial for an understanding of the economics of labor. In the best-known article included here, Richard Freeman asks a question that older, descriptive institutionalist theories had no trouble answering: what do unions do? Newer, quantitative theories often reduce the role of unions to their effects on wages, and hence indirectly on output and employment levels. Freeman insists that there is more to the impact of unions than the wages they win.

In terms of Albert Hirschman's dichotomy between exit and voice as means of expressing discontent, employees always have the option of exit from a job; unions can also give workers a voice for their concerns. This should reduce quit rates and lengthen average job tenure for unionized workers -- and Freeman demonstrates that this is exactly what happens. Much of the article is devoted to empirical analysis, disentangling the effect of union voice from the effect of other confounding influences, such as the effect of union wages (which would also be expected to reduce quit rates and lengthen tenure). Freeman also observes that grievance and arbitration procedures, seniority systems, and what he calls "industrial jurisprudence" may have a public-

goods character within the workplace; the only plausible means to pursue these objectives is through collective action.

Though they are clearly in a minority today, self-described institutionalist economists have not vanished. The article by David Marsden, an institutionalist labor economist, examines one of the ways in which conventional economic theory fails to describe the reality of labor markets. The common picture of markets matching jobs with workers who have transferable skills, says Marsden, ignores the problem of the origins of skills and the financing of the training process. The more transferable the skills, the more they resemble public goods; therefore, private firms will not provide the efficient level of training in such skills. Public education partially fills the gap, but many skills, even transferable ones, must be learned on the job.

In such circumstances, labor markets are fragile. Firms that provide training have a clear incentive to discourage inter-firm mobility; this is a reason for steep seniority-based pay increases. To the extent that firms succeed in discouraging mobility, there may be inefficient allocation of skilled labor between firms, and significantly different rates of pay for the same skills in the same region. Market outcomes will not be optimal by the usual criteria, and questions will arise of fairness and comparability among workers at different firms.

The final entry in this section takes the broadest view of the nature of work. It is taken from a new book by Chris Tilly and Charles Tilly (an economist and a historian, who are also son and father), which offers a comparative analysis of social change in the work process, focusing on case studies of the English and American textile, coal mining, and health care industries. Their most theoretical chapter, summarized here, draws on institutionalist, Marxist, and even some neoclassical insights into the labor process. They begin, not with individuals or firms, but with transactions and the contracts that embody them; these are grouped into social roles, particularly jobs, which are embedded in networks (such as markets and hierarchies) and organizations. Incentives that can be used to motivate the completion of transactions and contracts include compensation, commitment, and coercion; each of these three incentives can be separately present or absent, giving rise to eight different patterns of work (all three are present in a military command hierarchy, for example, while all are absent in scavenging). For Tilly and Tilly, the social context within which work transactions are embedded give rise to a range of possible outcomes and meanings, challenging us to think more expansively about the theory of labor economics.

CONCLUSION

How well have the new theories done in analyzing the unique characteristics of work processes and labor markets? The most widely discussed recent innovation, efficiency wage theory, offers a plausible mechanism that is compatible with the important “wage curve” findings, and can help explain the simultaneous existence of relatively well-paid “good jobs” and involuntary unemployment. Yet there are other realities, such as the social nature of production and the employment security of many good jobs, that call for a different conceptual framework. Not all the statistical evidence is compatible with an efficiency wage framework.

The hypothesis of labor market segmentation provides a straightforward explanation of many inequalities; relatively subtle empirical analysis appears to support the hypothesis, but more

work is needed here. Like efficiency wage theory, labor market segmentation could arise from any of several institutional and behavioral foundations. That is, market segmentation may be the mechanism leading to unequal outcomes, but that conclusion alone does not identify the underlying causes of segmentation, or of inequality.

Along with these relatively successful formal models, there is a lively debate over institutional questions which have, thus far, resisted useful formalization. The mathematics of individual maximization explains more than the simplest neoclassical theories would suggest, but less than we need to know in order to understand the economics of labor. The broader realities of work, as described throughout this volume, still present an ample unfinished agenda for the formulation of new theories of labor economics.

Notes

1. Robert M. Solow, *The Labor Market as a Social Institution* (Cambridge MA: Basil Blackwell, 1990), p.4.
2. This account draws heavily on Bruce E. Kaufman, "The Evolution of Thought on the Competitive Nature of Labor Markets," in Clark Kerr and Paul D. Staudohar, editors, *Labor Economics and Industrial Relations: Markets and Institutions* (Cambridge MA: Harvard University Press, 1994). See also the essays in that volume by George H. Hildebrand, Jack Barbash, and Clark Kerr.
3. For an introduction to these topics, see Jacob Mincer, "Human Capital: A Review," in Kerr and Staudohar, *op. cit.*, and David Sapsford, "The Labour Market: Unemployment and Search Theory," in David Greenaway, Michael Bleaney, and Ian Stewart, editors, *Companion to Contemporary Economic Thought* (New York: Routledge, 1991).
4. The articles by Leibenstein and other early contributors to the efficiency wage discussion, are reprinted in George A. Akerlof and Janet L. Yellen, editors, *Efficiency Wage Models of the Labor Market* (New York: Cambridge University Press, 1986).
5. There appears to be a conflict between efficiency wage theory and the wage curve, but in fact the two ideas are compatible. The efficiency wage theory examines the overall effects of wages on employment: the need to pay wages above the marginal product of labor implies a low demand for labor, "explaining" the existence of unemployment in general. However, it is also true that when unemployment is higher (for reasons unrelated to wage levels), then wages are lower. If unemployment is higher, the cost of job loss is greater, so workers are motivated to work hard even for relatively modest wages -- and profit-maximizing employers can therefore pay less. This argument provides an efficiency wage explanation of the wage curve.
6. Robert Buchele and Jens Christiansen, "Employment and Productivity Growth in Europe and North America: The Impact of Labor Market Institutions," forthcoming in *International Review of Applied Economics*.
7. For another sympathetic but critical view of the treatment of individual incentives and behavior in radical theories, see Nancy Folbre, "Hearts and Spades: Paradigms of Household Economics," *World Development* 14 (1986), 245-255.
8. Peter B. Doeringer and Michael J. Piore, *Internal Labor Markets and Manpower Analysis* (Lexington MA: Heath, 1971).
9. That is, the wage structure will be compressed upward in enterprises that have to hire high-wage workers; there is a premium on being an "insider" in a high-wage firm, even for those in low-wage occupations. Other "insider-outsider" models identify other mechanisms that lead to the same result: the firm's need for cooperation from incumbent workers, and its desire to avoid strikes when its operations are profitable, can lead to persistent payment of above-market wages for insiders. See, for example, Assar Lindbeck and Dennis J. Snower, "Cooperation, Harassment, and Involuntary Unemployment: An Insider-Outsider Approach," *American Economic Review* 78 (March 1988), 167-188.