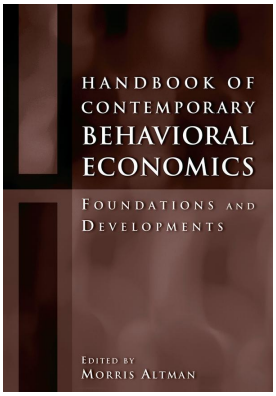


Handbook of Contemporary Behavioral Economics Foundations and Developments

Morris Altman

What a Difference an Assumption Makes

Publication details
https://test.routledgehandbooks.com/doi/10.4324/9781315703879.ch7
Morris Altman TMAN
Published online on: 15 Jul 2006

PART 2

CONTEXT AND MODELING

Page Intentionally Left Blank

CHAPTER 7

WHAT A DIFFERENCE AN ASSUMPTION MAKES

Effort Discretion, Economic Theory, and Public Policy

MORRIS ALTMAN

One of the most intriguing and important focal points of behavioral economics relates to research that falls under the analytical umbrella of efficiency wage and x-efficiency theories. Fundamental to these theories is the assumption of effort discretion—that individuals have some control over both the quality and quantity dimensions of the effort they put into the production process, and that this effort is maximized only under very special circumstances. Conventional economic theory assumes that effort is not a variable and that effort is maximized given the firm's production function and the economic agent's human capital endowment. For example, effort is assumed to be maximized in both its quantity and quality dimensions irrespective of market conditions or the firm's industrial relations system. In terms of modeling effort variability and its consequences, the assumption of effort variability dominates any of the other assumptions that are often part and parcel of the efficiency wage/x-efficiency narrative, such as irrationality, quasi-rationality, and imperfect product markets. Indeed, such assumptions, I argue, are subsidiary or secondary assumptions that are not necessary to modeling key economic scenarios in terms of effort variability.

This essay focuses upon the analytical importance and public policy consequences of introducing the behavioral assumption of effort variability and thereby effort discretion into the objective function of the individual and the firm. Special and critical attention is paid to the foundational contributions of Leibenstein and Akerlof and extensions to their modeling and analysis based on efficiency wage and x-efficiency theories. This review of the foundational efficiency wage/x-efficiency literature, as well as the revisions offered herein, builds to a large extent upon my own earlier contributions to the efficiency wage and x-efficiency literature. (See Berg's essay in this volume for an elaborate general discussion of this literature and Frantz 2004 for a comparison of the contributions of Akerlof and Leibenstein.) Implications of effort-variability-based theories for an understanding of involuntary unemployment, economic efficiency, economic convergence, globalization, economic efficiency, and altruism are briefly addressed, as is the overall sustainability of low- and high-wage firms and economies in the context of competitive markets. With regard to the latter, special attention is devoted to a critical appraisal of the reciprocity ultimatum game literature, which predicts the convergence toward relatively high-wage organizational forms.

This discourse on efficiency wage/x-efficiency theory is also placed in the context of the larger behavioral economics narrative. Related to this, I discuss a hypothesis that dominates much of

contemporary behavioral economics: that individuals are fundamentally irrational in choice behavior. I argue that behavioral economics has little to do with irrationality, and this is certainly true of efficiency wage and x-efficiency theory, although leading proponents of both efficiency wage and x-efficiency theory embed their models and analytical frameworks in the context of irrational or quasi-rational individuals or economic agents.

BEHAVIORAL ECONOMICS

Research on efficiency wages and x-efficiency, both of which are critically concerned with the productivity of labor as this relates to effort variability, is integral to behavioral economics in terms of its concern for the realism of the underlying behavioral and institutional assumptions that underlie the theory of the firm. A good example of a behavioral approach to the economics of the firm is presented in Cyert and March 1963. On a general level, one of the pioneers of behavioral economics, Herbert Simon (1959, 1987), distinguishes behavioral economics from contemporary economic theory largely in terms of the analytical significance of behavioral assumptions to the construction of economic theory. As opposed to neoclassical economics, which assumes that the realism of behavioral assumptions are of no analytical consequence (Altman 1999; Friedman 1953; Reder 1982), in behavioral economics the realism of behavioral assumptions can be of fundamental analytical importance. Behavioral economics is therefore distinguished by its efforts to test the empirical validity of a theory's behavioral assumptions and to determine the analytical and policy-related significance of introducing more realistic assumptions into standard theories.

> Thus, behavioural economics is best characterized not as a single specific theory but as a commitment to empirical testing of the neoclassical assumptions of human behaviour and to modifying economic theory on the basis of what is found in the testing process. And not all of the economists who hold a behavioural point of view also hold a common theory, or are all preoccupied with examining the same parts of the economic mechanism. (Simon 1987, 221)

With regard to the importance of the social and institutional parameters of economic analysis as opposed to the psychological, which is the mainstay of contemporary behavioral economics, Simon writes:

> The principal forerunner of a behavioral theory of the firm is the tradition usually called Institutionalism. It is not clear that all of the writings, European and American, usually lumped under this rubric have much in common, or that their authors would agree with each other's views. At best, they share a conviction that economic theory must be reformulated to take account of the social and legal structures amidst which market transactions are carried out. . . . The name of John R. Commons is prominent—perhaps the most prominent—among American Institutionalists. (Simon 1978, 499)

The realism of assumptions—and they can be of the psychological, social, or institutional variety—matter in general for behavioral economics, given that they impact on the accuracy of one's analytical predictions and one's causal analysis. This, in turn, can have powerful consequences for public policy and can impact on choices individuals make.

Leibenstein makes similar methodological points that are fundamental to his x-efficiency approach to the firm. Leibenstein discusses the relationship between theory and facts, the connection between *ex ante* predictions and the *ex post* confrontation of theory with empirical results.

> But the *ex ante, ex post* distinction does not really get at the roots of the issue. If one is to test a theory by making a "prediction," then the result appears to be more convincing if the prediction is a consequence of an inference from the theory, and the result is not known when the prediction is made. This has to do with effect rather than scientific significance. There is room for much *ex post* analysis of data. Economic historians could hardly perform their tasks if they could not apply their theories to existing data. A major purpose of theory is to *account* for the facts, obtained either *ex ante* or *ex post.* However, a significant problem is to sufficiently constrain the use of "free" undetermined variables, that is, to be sensitive to the danger that a fit may be obtained by attributing values needed to one or more unmeasured, unspecified, or possibly hidden variables involved. (Leibenstein 1983, 840)

Related to this point, Leibenstein argues:

> On the F-twist methodological issue, I believe that counterfactual postulates are unlikely to lead to correct *coherent* explanations. If the postulates cannot be tested, then we are forced to consider only the implications. But we would believe there is something wrong if we had a theory whose postulates were known to be counter to fact but which lead to correct predictions. . . . I do not believe that the only purpose of theory is as an engine for prediction, nor do I see that we should look at any particular set of methodological views as imposing decisive constraints on our scientific procedures at this stage in our knowledge. (Leibenstein 1983, 840)

With regard to x-efficiency theory, what is critical to appreciate is that by counterfactually constraining effort to some unspecified maximum irrespective of behavioral or institutional context, one is omitting a key explanatory variable in productivity and related analyses and thereby misspecifying causality and often generating models with relatively poor predictive power. If effort is assumed to be maximized, one might assume that productivity is maximized when it is not, one might ignore an important source of productivity change (effort variability), or one might assume that the economic pie is fixed and unaffected by distributional issues when in fact that is not the case. Leibenstein argues that by building theory upon more realistic and relevant analytical assumptions, one will be better able to explain economic reality and predict the consequences of individual choice behavior and public policy.

With regard to contemporary efficiency wage theory and its connection with behavioral economics, George Akerlof writes: "My dream was to strengthen macroeconomic theory by incorporating assumptions honed to the observation of such behavior" (2002, 411). His rendition of macroeconomics has therefore a behavioralist foundation, since it

> incorporates realistic assumptions grounded in psychological and sociological observation, [and has] produced models that comfortably account for each of these macroeconomic phenomena . . . Instead of denying the very existence of involuntary unemployment, behavioral macroeconomists have provided coherent explanations. Efficiency wage theories, which first appeared in the 1970's and 1980's, make the concept of involuntary unemployment meaningful. These models posit that, for reasons such as morale, fairness, insider power, or

> asymmetric information, employers have strong motives to pay workers more than the minimum necessary to attract them. Such "efficiency wages" are above market clearing, so that jobs are rationed and some workers cannot obtain them. These workers are involuntarily unemployed. (Akerlof 2002, 413–14)

Contemporary efficiency wage theory's raison d'être is to explain the existence of persistent involuntary unemployment based on realistic microeconomic assumptions. This is to correct for the arbitrary and unrealistic assumptions of the classical and new classical economists whose incorrect assumptions (according to the efficiency wage theorists) predict that all persistent unemployment must be voluntary.

CHOICE RATIONALITY

To situate one's modeling in the context of more realistic assumptions in no way implies that such behavioralist modeling presumes that economic agents need be irrational, although this is the sense one gets from much of contemporary behavioral economics (e.g., Kahneman and Tversky 1979; Tversky and Kahneman 1981; Kahneman 2003; Thaler 1992; for alternative behavioralist perspectives, see Altman 2003b; Gigerenzer 2000; Gigerenzer and Selten 2001; Goldstein and Gigerenzer 2002; Rieskamp, Hertwig, and Todd, this volume; Smith 2003, 2005; Todd and Gigerenzer 2003). For example, Kahneman writes:

> Our research attempted to obtain a map of bounded rationality, by exploring the systematic biases that separate the beliefs that people have and the choices they make from the optimal beliefs and choices assumed in rational-agent models. The rational-agent model was our starting point and the main source of our null hypotheses. (Kahneman 2003, 1449)

Deviations from the conventional norm imply, according to this perspective, errors and biases—the absence of rationality and even intelligence—in choice behavior. Given that the notion of individuals as fundamentally irrational is thought to be incompatible with general economic modeling, is this assumption of irrationality a necessary condition for behavioral economic theories, especially those related to the x-efficiency and efficiency wage discourse? Or is it simply the case that the conventional definitions of rationality are wanting from both descriptive and normative perspectives and need to be modified? In this case, one need not and should not drop the assumption of rational-intelligent economic agents. Behaving rationally, such as with regard to effort choice, may simply not be what the conventional wisdom presumes it to be. Indeed, behaving as a neoclassical rational agent in terms of effort choice might be the height of irrationality.

Herbert Simon argues that

> the term "rational" has long had in economics a much more specific meaning than its general dictionary signification of "agreeable to reason; not absurd, preposterous, extravagant, foolish, fanciful, or the like: intelligent, sensible." As is well known, the rational man of economics is a maximizer, who will settle for nothing less than the best. . . . It is this concept of rationality that is economics' main export commodity in its trade with the other social sciences. It is no novelty in those sciences to propose that people behave rationally—if that term is taken in its broader dictionary sense. Assumptions of rationality are essential components of virtually all the sociological, psychological, political, and anthropological theories with which I am familiar. What economics has to export then is not rationality, but a

> very particular and special form of it—the rationality of the utility maximizer, and a pretty smart one at that. (Simon 1978, 2)

Simon also makes the case that

> almost all human behavior has a large rational component but only in terms of the broader everyday sense of rationality, not the economists' more specialized sense of maximization [; moreover,] economics itself has not by any means limited itself to the narrower definition of rationality. [In addition,] economics has largely been preoccupied with the *results* of rational choice rather than the *process* of choice. (Simon 1978, 2)

From this perspective, deviations from the maximization assumption cannot be taken as evidence of irrational choice behavior.

More generally, Marsh argues:

> Engineers of artificial intelligence have modified their perceptions of efficient problem solving procedures by studying the actual behavior of human problem solvers. Engineers of organizational decision making have modified their models of rationality on the basis of studies of actual organizational behavior. . . . Modern students of human choice behavior frequently assume, at least implicitly, that actual human choice behavior in some way or other is likely to make sense. It can be understood as being the behavior of an intelligent being or group of intelligent beings. (Marsh 1978, 589)

Therefore Marsh and Simon, amongst the founders of behavioral economics, argue that if one searches deep enough, one is likely to discover the rational underlying choice behavior even if this behavior is not maximizing in the traditional sense. More recently, Vernon Smith, the founding father of experimental economics, whose work ties into and overlaps with behavioral economics, has made a case for ecological rationality:

> It is shown that the investor who chooses to maximize expected profit (discounted total withdrawals) fails in finite time. Moreover, there exist a variety of non-profit-maximizing behaviors that have a positive probability of never failing. In fact it is shown that firms that maximize profits are the least likely to be the market survivors. My point is simple: when experimental results are contrary to standard concepts of rationality, assume not just that people are irrational, but that you may not have the right model of rational behavior. *Listen to what your subjects may be trying to tell you*. Think of it this way. If you could choose your ancestors, would you want them to be survivalists or to be expected wealth maximizers? (Smith 2005, 149)

(On this point also see Gigerenzer 2002; Gigerenzer and Selten 2001; Goldstein and Gigerenzer 2002; Todd and Gigerenzer 2003.) Survival implies that economic agents have behaved sensibly and intelligently—that is to say rationally, if survival is a rational goal. Behaving neoclassically and not surviving would not be ecologically rational, nor is it how individuals actually behave, nor would one want such ecologically erroneous behavior to be recommended. Moreover, the ecological rationality perspective, referred to by Gigerenzer and associates as "fast and frugal heuristics," suggests that it is often the case that neoclassical rationality yields suboptimal results. From the perspective of economic efficiency, a rational neoclassical agent can be expected to

perform less efficiently than an ecologically rational, "fast and frugal" economic agent. What are errors and biases for Kahneman and Tversky and fellow travelers are often optimal behavior from the benchmark of ecological rationality. If individuals behave neoclassically, they would be therefore behaving suboptimally with regard to economic efficiency, given their level of effort input.

Simon makes specific reference to theories of the firm in his discourse on rationality:

> The general features of bounded rationality, selective search, satisficing, and so on have been taken as the starting points for a number of attempts to build theories of the business firm incorporating behavioral assumptions [this includes Leibenstein's x-efficiency theory]. Characterized in this way, there seems to be little commonality among all of these theories and models, except that they depart in one way or another from the classical assumption of perfect rationality in firm decision making. A closer look, however, and a more abstract description of their assumptions, shows that they share several basic characteristics. Most of them depart from the assumption of profit maximization in the short run, and replace it with an assumption of goals defined in terms of targets—that is, they are to greater or lesser degree satisficing theories. If they do retain maximizing assumptions, they contain some kind of mechanism that prevents the maximum from being attained, at least in the short run. In the Cyert-March theory, and that of Leibenstein, this mechanism can be viewed as producing "organizational slack," the magnitude of which may itself be a function of motivational and environmental variables. (Simon 1978, 508–9)

The implications of deviating from the conventional maximization assumption can be of profound analytical significance, but many conventional economists find it disturbing. According to Simon:

> The presence of something like organizational slack in a model of the business firm introduces complexity in the firm's behavior in the short run. Since the firm may operate very far from any optimum, the slack serves as a buffer between the environment and the firm's decisions. Responses to environmental events can no longer be predicted simply by analyzing the "requirements of the situation," but depend on the specific decision processes that the firm employs. However well this characteristic of a business firm model corresponds to reality, it reduces the attractiveness of the model for many economists, who are reluctant to give up the process-independent predictions of classical theory, and who do not feel at home with the kind of empirical investigation that is required for disclosing actual real world decision processes. But there is another side to the matter. If, in the face of identical environmental conditions, different decision mechanisms can produce different firm behaviors, this sensitivity of outcomes to process can have important consequences for analysis at the level of markets and the economy. Political economy, whether descriptive or normative, cannot remain indifferent to this source of variability in response. (Simon 1978, 509)

This speaks to the notion of multiple equilibria, which is developed in some detail below in the discussion of a behavioral model of the firm.

In terms of a theory of the firm, here we have individuals who from the perspective of the firm deviate from maximizing behavior, which might take the form of not maximizing firm profit through organizational slack. Do such deviations from conventionally defined (in the economics literature) rational behavior imply that individuals are not behaving rationally or intelligently, and should we therefore define such deviant behavior as irrational or quasi-rational? With regard to efficiency wage and x-efficiency theory, does this then imply that if such deviancy (or biases

and errors with regard to the conventional norm) is corrected, the economy would perform more efficiently? Or is it simply that rational behavior can yield economic inefficiency under normal circumstances and therefore that rational inefficiency need not be an aberration? In other words, if non-neoclassical behavior is intelligent given the circumstances faced by individuals and the substance of individuals' cognitive abilities, we must address the question of whether neoclassical efficiencies are realizable, and if they are, the circumstances under which they can be achieved. Both the efficiency wage and x-efficiency literatures suggest that economic efficiency can be achieved under reasonable circumstances.

Akerlof and Leibenstein and the modeling approaches that their contributions represent are quite sympathetic to the quasi-rational approach to economic agency because individuals deviate from neoclassical norms. However, Leibenstein's economic agents are certainly intelligent, as are Akerlof's. Leibenstein's agents simply do not conform to neoclassical behavioral norms with regard to maximizing effort input into the production process irrespective of circumstances. It is for this reason that Leibenstein refers to individuals as being quasi-rational—but this unnecessarily confounds the question of rationality with arbitrary neoclassical behavioral norms. Leibenstein also refers to neoclassical efficiency as his theoretical benchmark for ideal efficiency. But, once again, x-efficiency theory is all about a discourse on the necessary conditions for intelligent agents to perform optimally and to move the economy toward more efficiency in production. Also, in the efficiency wage literature agents are largely intelligent and rational, but given effort variability, real wages tend to be above their market-clearing values. But such behavior need not be inconsistent with utility maximization. Utility is not maximized by maximizing effort input. Quasi-rationality enters the picture as the solution to the market-clearing problem in terms of money illusion. And given how money illusion is defined in efficiency wage theory, related to the cognitive costs of computing and reacting to small decreases in real wages, it is not clear how such illusions are irrational even from a conventional neoclassical perspective.

I argue, in the traditional of Simon, March, Smith, and Gigerenzer, that economic agents can be modeled as rational or intelligent in behavior even if they do deviate from neoclassical norms. Thus individuals can be rational even if they do not maximize effort inputs under all circumstances. Indeed, individuals can be viewed as broadly rational:

> Individuals maximize utility in a consistent way, and . . . they consider the effect of their actions on future as well as present utilities . . . individuals are still assumed to make forward-looking, maximizing, and consistent choices. (Becker 1996, 22–23)

But in this modeling framework the individual's objective function is not specified. As Becker points out, rational individuals' preferences are influenced by circumstances inclusive of past experiences and social interactions.

Following from such an understanding of rationality, each individual can have a different objective function, and there is no a priori guarantee that objective functions are consistent across individuals. Rational agents might have conflicting preferences. There is no reason to expect a unitary objective function, nor that the dominant objective function with regard to decision making will be consistent with firms producing efficiently (effort maximization in terms of x-efficiency theory) or with an economy minimizing unemployment, which is the analytical focus of efficiency wage theory. I argue that the assumption of rationality, broadly defined and inclusive of maximizing individuals, is consistent with the realization of both economic efficiency and inefficiency as well as of either significant or marginal amounts of voluntary unemployment. At least the important efficiency wage/x-efficiency vein of behavioral economics can be effectively

and efficiently mined and processed without reference to irrationality as a key causal variability. It is in fact quite telling that significant inefficiencies can be generated by the choices of rational individuals. In the behavioral model discussed below, persistent and significant rational inefficiencies can be obtained even in a competitive environment. Survival does not imply efficiency. Being fit enough to survive as a firm can take on various organizational forms and economic results, only a subset of which need be consistent with economic efficiency. In this sense ecological rationality need not imply economic efficiency, where such efficiency is realizable only in particular environmental (institutional) settings.

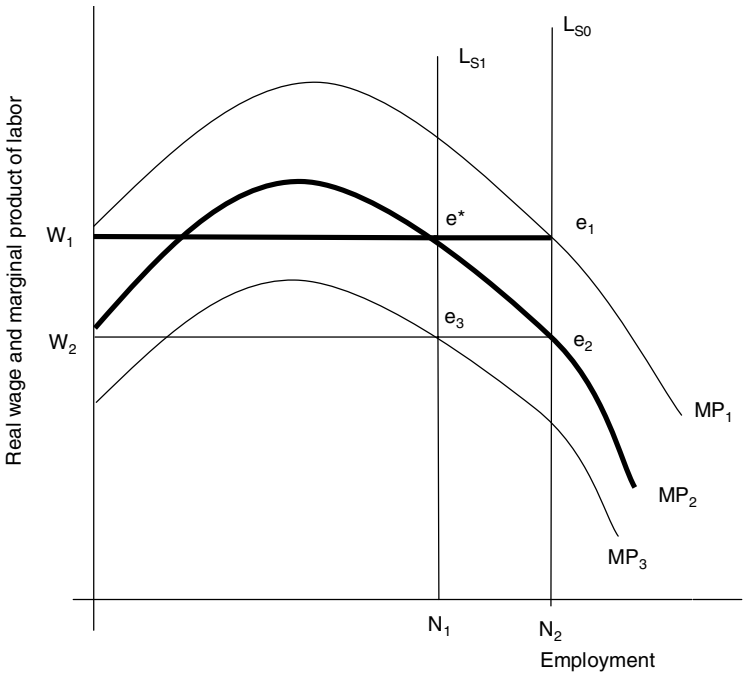
EFFICIENCY WAGE THEORY AND EMPLOYMENT

Leibenstein's Foundational Modeling: Introducing Effort Variability

What is today dubbed efficiency wage theory was pioneered by Harvey Leibenstein (1957a, 1957b) to help resolve the apparent paradox of there being long-run surplus labor in less developed economics at positive real wages. Surplus labor can take the form of either hidden or actual unemployment. All agents in his modeling framework are rational and maximizing with regard to profit (marginal cost equals marginal benefit) and utility. Leibenstein breaks with the conventional wisdom by introducing the empirically based assumption that effort input varies positively with real wages since real wages affect the physiological ability of workers to work more intensively. Leibenstein assumes that the changing capacity of workers to work more or less hard will be translated into actual changes in effort input on the job. In the conventional modeling effort input remains constant in the face of change in real wages. Given the assumption of effort variability, cutting real wages in response to downward pressures on the labor market generated by surplus labor need not yield a profit-maximizing solution to the firm, as it would in the conventional model. Thus, profit-maximizing rational firms will not cut real wages as they would in the conventional model, yielding surplus labor. In other words, long-term unemployment is a product of the downward stickiness of real wages, as it is in the pre-Keynesian classical world and in the more recent new classical rendition of the classical model. If only real wages could be cut, surplus labor would be employed. But Leibenstein maintains that rational firm owners attempt to form coalitions, often unstable, to prevent market forces from driving wages downward, since this would cut into profits, or else simply hire more workers than would be warranted by profit-maximizing conditions—they hire a subset of workers whose marginal product exceeds their wage—to relieve downward market pressure on wages. This is, of course, a form of featherbedding.

By introducing effort variability into his modeling framework, Leibenstein posits that each real wage can be modeled as being related to a unique marginal product of labor curve and that where the wage equals the marginal product of labor there is a unique net profit accruing to the firm. Any cut in the wage rate shifts the marginal product of labor curve inward as compared to the conventional model, wherein a change in real wages simply results in movement along the marginal product curve. Therefore, with effort variability, cutting real wages could result in either the same or a lesser level of employment, instead of resulting in more employment. I would illustrate this basic argument in terms of Figure 7.1, where full employment is given by N_2, but current employment stands at only N_1 given the wage rate W_1, where the wage equals the marginal product of labor at e*. The wage W_1 yields marginal product curve MP_2. In the conventional model, market forces would force the real wage down, yielding full employment, N_2. But given effort variability, a cut in the real wages yields a lower marginal product curve, such as MP_3, where equilibrium is given at e_3 and employment is no higher than it was at the higher

Figure 7.1 **Labor Demand and Marginal Product**

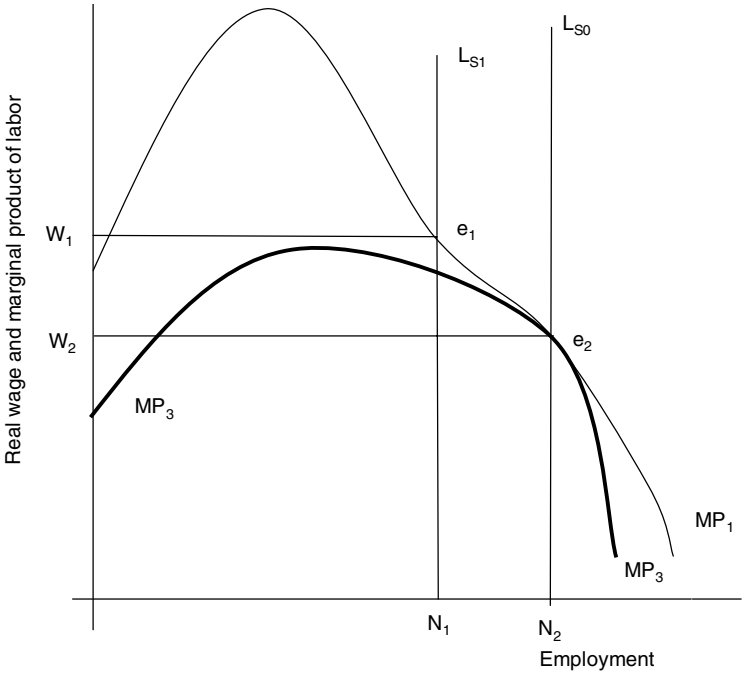


wage. Of course, employment could have fallen had the marginal product curve shifted inward any further as a result of the diminution in real wages.

Leibenstein takes this argument further, making specific assumptions about the shape and positioning of the marginal product of labor curve, illustrated in Figure 7.2, where the two marginal product curves are drawn in such a fashion that there is one point, e_2, where the curves are tangent. This would be at the point of full employment, N_2. In this scenario, the marginal product curve shifts inward from MP_1 to MP_3 with a cut in the real wage from W_1 to W_2. Moreover, given the shift in the marginal product curve as wages fall and the assumed shape of the marginal product curves, the area above the wage rate line W_1 for MP_1 is greater than the area above wage rate line W_2 for MP_2. Thus net profit is greater at the higher wage given effort variability. Given these assumptions, the firm maximizes profit at wage rate W_1, yielding a less than full employment level of employment, N_1. Full employment could be obtained at a lower wage, but this would be inconsistent with the firm maximizing its profit. And for this reason firms resist market pressure to pay workers a lower wage rate.

Another point made by Leibenstein and taken up by contemporary efficiency wage theorists is that there is one wage that yields optimal net profit. This argument is embedded in the implicit assumptions made by Leibenstein with regard to the elasticity of effort input with respect to changes in the real wage and the elasticity of productivity with respect to changes in effort input. Leibenstein assumes diminishing returns in the relationship between wage changes and changes in effort, and changes in effort and changes in productivity. For example, for any wage below the one that yields maximum profit, the elasticity of effort to wage changes is greater than 1, and above this wage the elasticity is less than 1. The unique wage that maximizes net profit is referred in the contemporary literature as the efficiency wage. This point can be illustrated in Figure 7.3,

Figure 7.2 **Labor Demand, Marginal Product, and Profits**

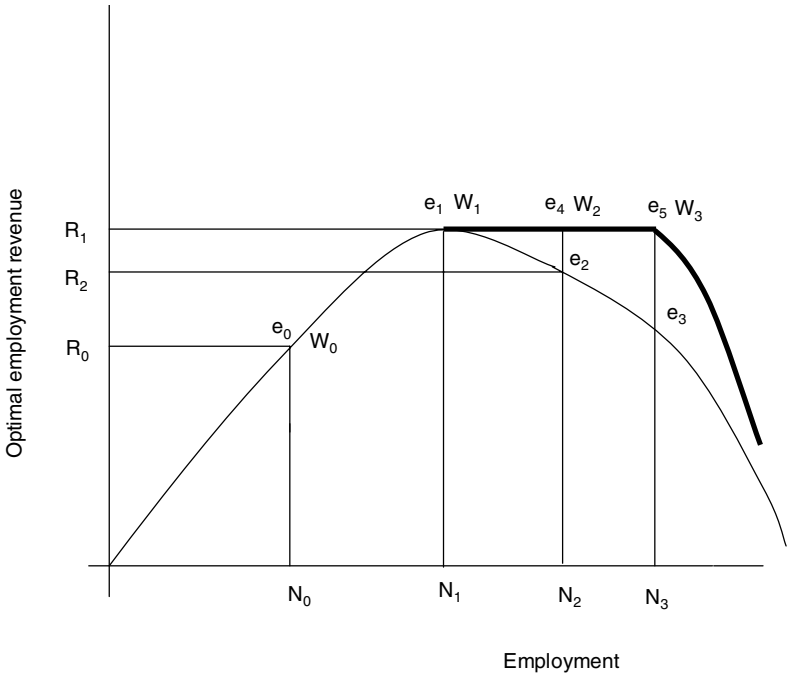


where optimal employment revenue is mapped along the vertical axis and employment along the horizontal axis. The concept of optimal employment revenue is derived from Leibenstein and refers to the maximum net revenue that is generated for any given real wage. Only one real wage yielding one level of employment maximizes net revenue, W_1 and N_1, respectively. This level of employment is less than full employment, where the latter can be achieved only at a lower real wage and at a lower net revenue or profit. There is an unambiguous trade-off between profit and higher levels of employment in this modeling framework. Thus, more employment is obtained here only if firms hire more workers than is warranted by maximum profits, such as occurs with featherbedding. Leibenstein argues that firms, independently and jointly, search for ways to hire more workers than can be justified by the criterion of profit maximization so as to relieve downward pressure on the labor market. And this makes sense as long as net profits remain greater than they would be at the market-clearing wage. However, each firm remains under severe pressure to pay lower real wages given the existence of surplus labor, thereby resulting in the efficiency wage being unstable. Alternatively, if workers could be made to accept lower wages without reducing effort inputs, more employment could be realized. But this could not come to pass if effort is strictly tied to real wage levels in a world where real income is so low that variations in real wages affect workers' capacity to generate more or less effort input.

A Behavioral Interlude to the Efficiency Wage Narrative

Prior to discussing contemporary efficiency wage theory it would be helpful to further elaborate upon the implications of efficiency wage theory on the cost side. Based on Altman 1992

Figure 7.3 **Labor Demand and Marginal Revenue**



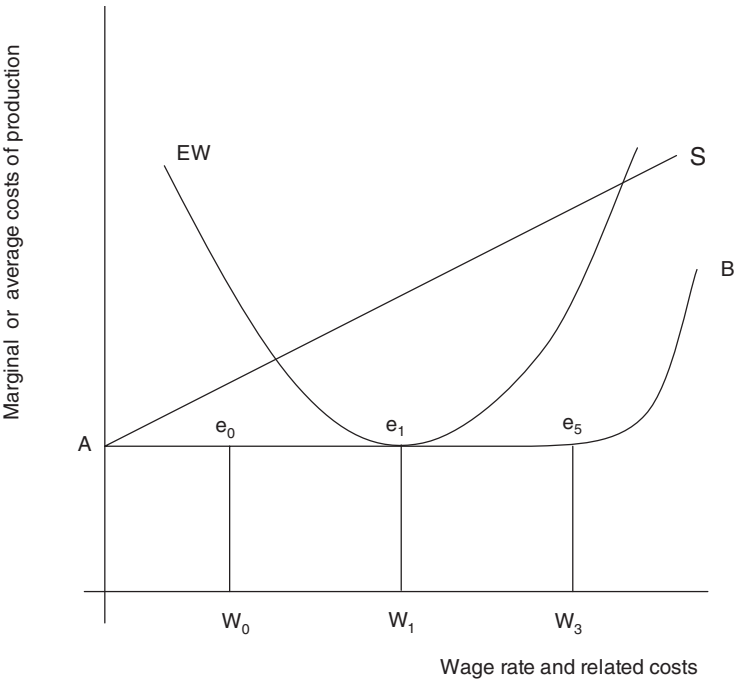
and 1996, we assume for simplicity that labor is the only factor input. Given this assumption, the cost dimension of efficiency wage narrative can be given by equations 1a and 1b and Figure 7.4.

$$MC = \frac{w}{\left(\frac{dQ}{dL}\right)} \tag{1a}$$

$$AC = \frac{w}{\left(\frac{Q}{L}\right)} \tag{1b}$$

where *MC* is marginal cost, *w* is the wage rate, (*dQ/dL*) is the marginal product of labor, *AC* is average cost, and (*Q/L*) is the average product of labor. In the efficiency wage narrative there is only one wage that minimizes marginal cost and average cost. Any downward deviation of the wage from the efficiency wage increases unit cost for example, just as it will reduce net profits. This point is further illustrated in Figure 7.4, where W_1 is the efficiency wage and cost is minimized at e_1. The efficiency wage is unique—cost increases if the wage rises or fall. This is given, it is critically important to note, by the assumed smooth convexity of the cost curve.

Figure 7.4 **Labor Demand and Production Costs I**



Contemporary Efficiency Wage Theory

Leibenstein's foundational efficiency wage narrative was finessed and enriched by a new line of scholars, exemplified by George Akerlof (Akerlof 1980, 1982, 1984, 2002; Akerlof and Yellen 1986, 1988, 1990; Akerlof, Dickens, and Perry 2000; Bewley 1999; Shapiro and Stiglitz 1984; Stiglitz 1987), whose objective was to provide a rigorous micro foundation for persistent involuntary unemployment. Indeed, Akerlof and Stiglitz were awarded the Nobel Prize in Economics in part for these efforts. Contemporary efficiency wage theory is not tied to the notion that effort variation is a function of the physiological capacity of workers to supply more or less effort input. Rather, it is argued that effort varies for sociological reasons such as fairness, which is tied to relative real wages. Workers retaliate against employers when they feel hard done by. The latter theme is developed in the ultimatum game and related reciprocity literature (see, for example, Fehr and Schmidt 1999; Fehr and Gachter 2000a, 2000b; Güth 1995; see also Güth and Ortmann, this volume).

This modeling of the labor market attempts to identify a rationale for the downward stickiness of real wages, which is out of the choice domain of workers. Firms refuse to cut real wages over the business cycle because they are quite aware of the fact that such wage cuts would only serve to reduce effort input on the part of their employees, thereby reducing profits and increasing unit cost. It is assumed that there is one wage, the efficiency wage, that serves to maximize profits and minimize unit cost. Thus workers are paid above the market-clearing wage, just as they are in Leibenstein's narrative. But unlike in Leibenstein's discourse, the efficiency wage is stable—individual firms have no incentive to pay a lower wage given the downward market pressures on the labor market. Surplus labor as manifested in involuntary unemployment is a stable equilib-

rium. Here we have the classical and new classical narrative revisited and refined such that workers bear no responsibility for the supply-side cause of persistent unemployment, as they must in the classical model of the labor market. Needless to say, the public policy metaphor is the same in both narratives—real wages must be cut if one wishes to secure more employment. The demand side might play a role in determining the level of employment, but supply conditions establish the micro priors for determining the extent to which more workers can be profitably employed. However, unlike Leibenstein, contemporary efficiency wage theorists, especially Akerlof, argue that is it quite possible for real wages to be cut over the course of the business cycle so as to secure more employment without unleashing the efficiency wage effect, which would make such a wage cut untenable from the perspective of profit-maximizing firms.

Critical to cutting real wages without workers retaliating by reducing effort input is the assumption that workers are subject to money illusion. If government adopts a mild inflationary policy, real wages fall, yielding more employment, without workers responding by reducing effort input. Thus real wages cannot be reduced by individual firms, since workers will reduce effort input in retaliation, but macro policy can certainly do the trick. Therefore, if workers are quasi-rational (some might say irrational), higher levels of employment can be secured. But such irrationality on the part of workers is strongly denied by the new classical economics, an assumption that Akerlof argues is fundamentally wrong. Akerlof argues with regard to the significance of efficiency wage theory for understanding the persistence of involuntary unemployment:

> A major contribution of behavioral macroeconomics is to demonstrate that, under sensible behavioral assumptions, monetary policy *does* affect real outcomes just as Keynesian economics long asserted. Cognitive psychology pictures decision makers as "intuitive scientists" who summarize information and make choices based on simplified mental frames. Reliance on rules of thumb that omit factors whose consideration have only a small effect on profit or utility is an implication of such cognitive parsimony. In the wage-price context, simple rules cause inertia in the response of aggregate wages (and prices) to shocks—the exact "sticky wage/price" behavior that New Classical economists had so scornfully derided. In the New Classical critique, the inertial wage behavior hypothesized in the "neoclassical synthesis" is irrational, costly for workers and firms, hence implausible. Behavioral economists have responded by demonstrating that rules of thumb involving "money illusion" are not only commonplace but also sensible—neither foolhardy nor implausible: the losses from reliance on such rules are extremely small. (Akerlof 2002, 416)

Akerlof elaborates on this point, relating his discourse to that of Kahneman-Tversky type behavioral economics:

> Keynes' assumption that workers resist nominal wage cuts was consistent with his intuitive understanding of psychology. The assumption also coincides with psychological theory and evidence. Prospect theory posits that individuals evaluate changes in their circumstances according to the gains or losses they entail relative to some reference point. The evidence suggests that individuals place much greater weight on avoiding losses than on incurring gains. Daniel Kahneman and Amos Tversky (1979) have demonstrated that many experimental results that are inconsistent with expected utility maximization can be rationalized by prospect theory. Downward wage rigidity is a natural implication of prospect theory if the current money wage is taken as a reference point by workers in measuring gains and losses. In support of this view, [Shafir, Diamond, and Tversky 1997] found in a question-

> naire study that individuals' mental frames are defined not just in the real terms hypothesized by classical economists but also exhibit some money illusion. Numerous empirical studies document that money wages are, in fact, downward sticky. (Akerlof 2002, 420)

Thus, for Akerlof, the downward stickiness of wages applies specifically to money, not necessarily to real wages. Workers' cognitive limitations and the related costs of processing information, inclusive of the effects of low rates of inflation on real wages, open the door to the capacity of workers to be deluded into accepting lower real wages when this is a product of inflationary monetary policy.

Akerlof summarizes some of the empirics related to his argument that, contrary to the anti-Keynesian economists, monetary policy can have a permanent effect on employment:

> At low inflation there is a long-run trade-off between output and inflation if there is aversion to nominal pay cuts. Unlike the Friedman-Phelps model, in which such a trade-off is transitory, long-term increases in inflation (if it is close to zero) result in significantly less employment and more output. The logic goes as follows. In both good times and bad, some firms and industries do better than others. Wages need to adjust to accommodate these differences in economic fortunes. In times of moderate inflation and productivity growth, relative wages can easily adjust. Unlucky firms can raise the wages they pay by less than the average, while the lucky firms can give above-average increases. However, if productivity growth is low (as it was from the early 1970's through the mid-1990's in the United States) and there is no inflation, firms that need to cut their real wages can do so only by cutting the money wages of their employees. Under realistic assumptions about the variability and serial correlation of demand shocks across firms, the needed frequency of nominal cuts rises rapidly as inflation declines. An aversion on the part of firms to impose nominal wage cuts results in higher permanent rates of unemployment. Because the real wages at which labor is supplied are higher at every level of employment when inflation is low, the unemployment rate consistent with stable inflation rises as inflation falls to low levels. Spillovers produce an aggregate employment impact that exceeds the employment changes in those firms that are constrained by their inability to cut wages. Thus, a benefit of a little inflation is that it "greases the wheels of the labor market." Simulations of a model with intersectoral shocks and aversion on the part of firms to nominal wage cuts suggests that, with realistically chosen parameters, the trade-off between inflation and unemployment is severe at very low rates of inflation, when productivity growth is low. For example, a permanent reduction in inflation from 2 percent per year to zero results in a permanent increase in unemployment of approximately 2 percentage points. (Akerlof 2002, 421)

Contemporary efficiency wage theorists, best exemplified in the work of Akerlof, not only recast Leibenstein's theoretical frame, such that it relates to developed economies where workers are paid well above physiological subsistence and efficiency wages are stable given their profit-maximizing function. They also make the case that it is possible to subvert the workers' tendency to reduce effort input when real wages are reduced. Thus Akerlof could argue that a solid theoretical basis has been established to demonstrate not only the reasonableness of the assumption of the downward inflexibility of wages but also why real wages can be cut in face of this inflexibility through sound monetary policy (for further details see Akerlof, Dickens, and Perry 1996, 2000).

It is important to note, however, that Keynes completely rejects the notion of money illusion, although he assumes that reductions in real wages are a necessary supply-side condition for in-

creasing employment given the conventional assumption of diminishing returns to labor and therefore a downward-sloping marginal product of labor curve. Keynes argues:

> [Workers] do not resist reductions of real wages, which are associated with increases in aggregate employment and leave relative money-wages unchanged, unless the reduction proceeds so far as to threaten a reduction of the real wage below the marginal disutility of the existing volume of employment. Every trade union will put up some resistance to a cut in money-wages, however, small. But since no trade union would dream of striking on every occasion of a rise in the cost of living, they do not raise the obstacle to any increase in aggregate employment which is attributed to them by the classical school. (Keynes 1936, 14–15)

For Keynes, workers are not characterized by quasi-irrationality. They accept cuts in real wages up to some point, given their expectation of what this implies for increasing employment. This being said, the assumption that real wages must be cut for employment to be increased as an unequivocal derivative of efficiency wage theory requires some critical perspective given the policy implications of this assumption.

A BEHAVIORAL MODEL OF FIRM AND EFFICIENCY WAGE THEORY

One alternative narrative to conventional efficiency wage modeling stems from a behavioral model of the firm, which I've developed in some detail elsewhere (for example, Altman 1992, 1996, 1998, 1999, 2001b, 2002, 2003a, 2005). This modeling of the firm and economic agency also serves as a segue to my critical discussion of x-efficiency theory. Although effort is introduced into the short-run production function as a variable, unlike what is assumed in the traditional efficiency wage literature, in the behavioral model there exists some linearity with respect to the relationship between effort inputs and the wage rate, which, for simplicity, is assumed to embody the entire system of industrial relations within the firm. In this scenario, there is an array of wage rates consistent with a unique marginal, average cost, or net profit when productivity changes, brought about through effort input changes, just suffice to offset the cost impact that changes to the real wage might otherwise have. In the conventional efficiency wage view, there exists one unique wage that minimizes costs or maximizes profit.

These basic points can be illustrated in equations 1a and 1b, above, where labor is the only factor input. In this case, cost does not increase in the face of an increasing wage rate, nor does it fall when the wage rate falls if there exist corresponding and proportional changes in the marginal and average product, respectively. There is no one efficiency wage. It is important to note that productivity need not increase proportionally to increases in labor costs when labor is only one among many inputs into the production process, which is typically the case. When labor is not the only compensated factor input, labor productivity need increase less than proportionally to increases in labor costs so as to neutralize increases labor cost. This point is illustrated in equation 2:

$$\frac{dAC}{AC} = \left(\frac{dw}{w}\right) * \left(\frac{w * L}{w * L + NLC}\right) \tag{2}$$

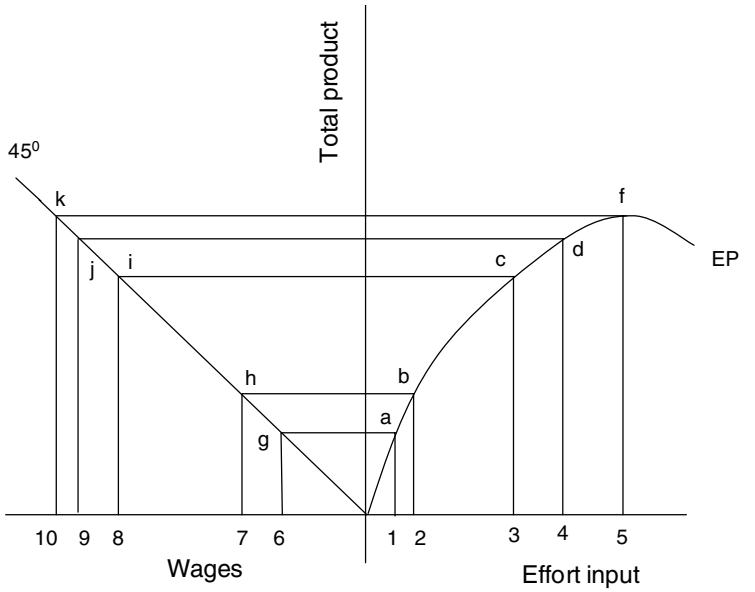
where *dAC* is the change in average cost, *dw* is the change in the wage rate, and *NLC* is nonlabor costs. If, for example, the wage rate is increased by 10 percent (*dw/w*) and wage costs (*w*L*)

represent 100 percent of total costs, labor productivity must increase by 10 percent to compensate for what would otherwise be a 10 percent increase in average costs. However, if wage costs represent only 50 percent of total costs, labor productivity must increase by only 5 percent to compensate for what would otherwise be a 5 percent increase in average costs.

The implications of introducing some linearity to the wage productivity relationship are further illustrated in Figure 7.4, above. Line segments S, EW, and B refer to the assumed wage-cost relationship in the conventional, efficiency wage, and behavioral models, respectively. In the standard model, where effort is fixed, increasing wages increases costs, ceteris paribus. In the efficiency wage model there is one wage that minimizes cost. But in the behavioral model, there is an array of wage rates consistent with one unique unit cost, at least to W_3, after which output cannot be increased sufficiently to offset increasing wage costs—here we enter into the realm of diminishing returns. Introducing a degree of linearity into the wage-productivity relationship, with effort input as an intermediary variable, yields significantly different analytical predictions with regard to the impact wage changes have upon unit cost as compared to what is generated in either the standard or efficiency wage models. This point is further illustrated in Figure 7.3 above, where the relationship between wages, net profit, and employment is modeled. In this scenario, there is an array of wage rates and levels of employment consistent with one unique maximum net profit (optimal employment revenue). As the wage rate increases above W_3, net profit falls as increasing labor productivity related to increased effort input no longer suffices to compensate for the increases in the wage rate. This compares to the efficiency wage narrative, wherein only one wage rate corresponds to maximum net profit, and this at a less than full employment level of employment.

The behavioral model suggests that there exists a considerable degree of freedom with regard to possible real wages and employment combinations sustainable for a competitive profit-maximizing firm if effort discretion exists and there is no assumed mechanical (nonlinear) relationship posited between the level of real wages, effort input, and productivity. It is important to note that one can have some linearity in the wage-productivity relationship even if one assumes there exist eventually diminishing returns in the wage-effort input relationship. This point is illustrated in Figure 7.5, where effort input is mapped against total product in the right quadrant, yielding an effort-product curve (EP). This curve is constructed to reflect diminishing returns. The wage rate is mapped against total product in the left quadrant such that the wage rate equals total output in a one-factor input model. In this scenario, each percentage increase in total output equals the percentage increase in the wage rate, yielding constant unit cost. This diagram therefore addresses the question of the extent to which effort must increase if unit cost is to remain constant in face of wage rate increases, and how this is contextualized given diminishing returns to effort. Clearly diminishing returns to effort do not preclude linearity between wage rates and unit cost. Diminishing returns simply imply that effort input must increase at an increasing rate for every given change in the wage rate. Initially effort might have to increase much less than the proportional increase in the wage rate. But as the wage rate keeps increasing, effort will eventually have to increase much more than the proportional increase in the wage rate if unit cost is to remain constant. Eventually, effort input is maximized and increasing wages are matched by increasing unit cost, at point f along EP. Unit cost might also increase if economic agents will not increase effort input sufficiently to generate the productivity increase required to offset the impact of increasing wages. Thus unit cost can increase prior to point f. But there is no a priori reason to expect, in theory, that there will be one wage that will minimize unit cost when effort input is a variable in the production function. Much depends on how economic agents respond to changes in wages and this is an empirical question. The critical point here is that the assumption in efficiency

Figure 7.5 **Labor Demand and Production Costs II**

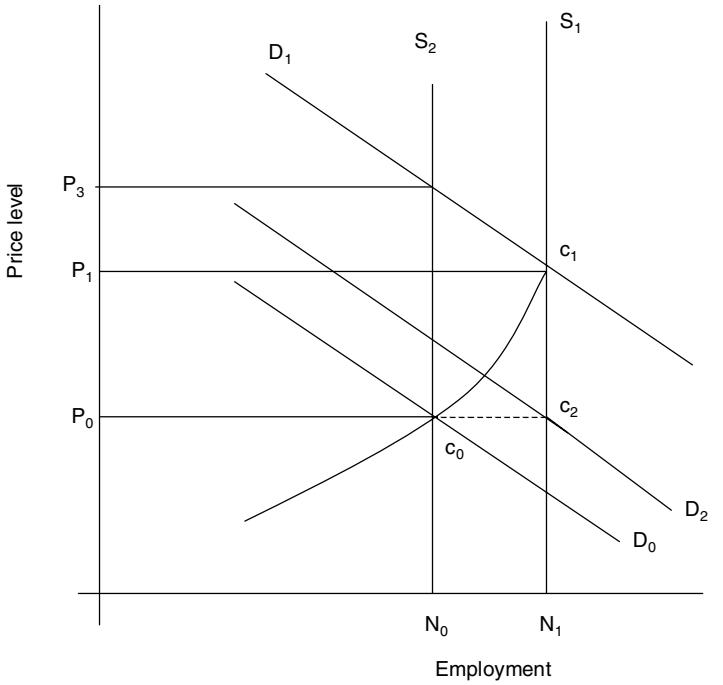


wage theory that relatively high wages yields excess unemployment due to their generating higher unit costs and lower net profits is simply an assumption, and it need not be the ideal one given the facts at hand.

Introducing some linearity into the wage-cost relationship is consistent with real wages being inflexible downward over the business cycle, as in the traditional efficiency wage narrative. Firm owners and managers have no incentive to reduce real wages if there is no perceived advantage in doing so and if there are short-term costs involved, such as worsening the labor-management relationship. This is especially true if it is perceived that real wages will be restored during the upswing in the business cycle. But the behavioral model suggests that relatively high real wages need not impede employment growth, given adequate aggregate demand, if such real wages are compensated for in terms of higher labor productivity. Moreover, in this case, increasing prices to reduce real wage, given that workers suffer from money illusion, is no longer a necessary condition to generate more employment (Altman 2006a).

In Figure 7.6, the increase in employment is illustrated using vertical supply curves wherein changes in employment are not affected by price changes. Holding the price level constant at P_0, employment increases from N_0 to N_1 as demand rises as a consequence of an outward shift of the marginal product of labor curve, which compensates for any increases in real wages required to induce the requisite increases in labor supply underlying increasing employment. This is reflected in the movement of the employment supply curve from S_2 to S_1. The employment supply curve is thus perfectly inelastic with respect to price, and employment becomes "full employment" once employment is at a maximum given the labor supply and production functions. In the short run, increasing demand from D_0 would simply increase price, given employment curve S_2. In the longer run, as the supply curve shifts outward to S_1, employment increases to N_1. An employment supply curve can be given by $P_0c_2c_1$. In this scenario, a price increase can accompany the increase in employment if the demand increases sufficiently, for example, to D_1. But unlike the Akerlof narrative, it is not the increase in the price level that actually causes the increase in

Figure 7.6 **Prices and Employment**



employment to N_1. Rather, it is an increase in aggregate demand in the context of the real-wage-rate-induced increase in labor productivity that is responsible for an increase in the level of employment. The supply curve shifts to the right for conventional reasons, which should incorporate extraordinary increases in efficiency as well as technological change induced by increasing real wages (Altman 2002). Cutting real wages by increasing demand and prices, on the other hand, would have the effect of shifting the supply curve to the left, yielding both higher prices and less employment. Thus if S_1 is the initial supply curve and demand is increased from D_2 to D_1, the price would rise first to P_1, cutting real wages. This would shift the supply curve to S_2, yielding a lower equilibrium employment at N_0 and a price of P_3. In this context, employment would be encouraged by efforts to promote increases in labor productivity and more aggressive demand management strategies designed to accommodate the resulting increase in aggregate output, as opposed to designing policy to cut real wages.

The behavioral model is consistent with the work of Blanchflower and Oswald (1995), who find, using an international sample population, that regions with low unemployment rates are systematically correlated with high wage rates and vice versa. The level of unemployment associated with a particular level of real wages critically depends on the elasticity of labor productivity to the wage rate and the responsiveness of macroeconomic demand-side variables to increasing output. The behavioral model also suggests that a positive empirical relationship between the price level (or the rate of inflation) and employment need not be a causal one. Inflation might simply be a reflection of an accommodating macroeconomic policy that allows for the absorption of the increasing output flowing from the increase in labor productivity. Inflation would here be a product of overshooting increases in aggregate demand in a world of uncertainty. In this sense, low rates of inflation are consistent with increasing the level of employment, although not the cause of such increases.

X-EFFICIENCY: A GENERAL FRAMEWORK

What Is X-efficiency?

Simon's remarks on why and how inefficiency can persist provides an informative context within which to discuss x-efficiency theory, wherein economically inefficient firms can survive in the long run. Simon argues, in the context of discussing the evolutionary argument with regard to firm survival, that unlike what is maintained in the neoclassical theory of the firm, where only the fittest—cost-minimizing or profit-maximizing—firms can survive, there is no good reason to expect that inefficient firms should be wiped out by market forces.

> In the biological world at least, many organisms survive that are not maximizers but that operate at far less than the highest achievable efficiency. Their survival is not threatened as long as no other organisms have evolved that can challenge the possession of their specific niches. Analogously, since there is no reason to suppose that every business firm is challenged by an optimally efficient competitor, survival only requires meeting the competition. In a system in which there are innumerable rents, of long-term and short-term duration, even egregious sub-optimality may permit survival. (Simon 1987, 223)

Inefficient firms can survive if they are not challenged by relatively efficient firms. Leibenstein makes the case that inefficiency persists because inefficient firms are protected from competitive forces. Critical to this protection is the existence of imperfect product markets, which provide firms with the capacity to produce at a higher unit cost than what can be achieved if firms produced optimally. Thus firms are typically not challenged, argues Leibenstein, by optimally efficient competitors.

Leibenstein's x-efficiency theory has two critical dimensions. Much of this theory is concerned with modeling persistent economic inefficiency as a product of suboptimal managerial performance. The other dimension, which is much less structured than the first, deals with economic inefficiency, which flows from economic agents having conflicting objective functions—the roots of economic efficiency here overlap with the narrative of how the firm is managed. X-efficiency theory, as presented by Leibenstein, models an economy in the context of imperfect product markets. Unlike what is predicted in the conventional model, Leibenstein argues that such an economy is not simply characterized by allocative inefficiencies, where the latter is typically of trivial empirical importance (Altman 1990; Frantz 2004; Leibenstein 1966). Rather, firms in imperfectly competitive markets embody what he refers to as x-inefficiencies, which cannot be identified in the conventional model given its assumption of effort being maximized and invariant (see Frantz 1997 for a survey of the empirics underlying the existence of x-inefficiency). Leibenstein, basing his argument on the extant empirical literature, holds that there exists a tremendous persistent diversity in productivity across like firms. This can be attributed to analogous diversity in effort inputs across firms where in a wide array of firms there exists a suboptimal (from a neoclassical perspective) level of effort inputs. Thus many firms operate in the interior of the neoclassical production possibility frontier as a consequence of effort not being maximized. The difference between the actual production possibility frontier and the ideal is one measure of x-inefficiency wherein the outermost frontier is the benchmark for x-efficiency in production. Leibenstein makes the case, however, that from a practical point of view, the best measure of x-efficiency is the most productive or relatively most x-efficient firm where the latter might be perfectly x-efficient with respect to some ideal benchmark. One should note that Leibenstein

models x-efficiency in terms of both the quantity and quality of effort inputs. Thus productivity can vary holding the quantity dimension constant and simply varying the quality dimension—working smarter can affect productivity as much as, or even more than, working harder. Needless to say, critical to Leibenstein's model is the assumption of effort variability.

Leibenstein summarizes some of his key foundational points thusly:

> The assumptions about firms on which the theory rests are: (1) that labor contracts are vague and incomplete: (2) that detailed supervision of labor is impractical and or inefficient; hence, (3) there are normally many areas of choice open to managerial as well as other employees in determining how to fulfill their work roles. Legally, the firm makes contractual arrangements in its name, but the formal and informal contracts through which firms hire human inputs have significant gaps. The activities an individual is expected to carry out are rarely completely specified, and sometimes they are almost completely unspecified. Hence, it is necessary for individuals to interpret their jobs. (Leibenstein 1973a, 767)

Economic agents, therefore, have a considerable degree of freedom in terms of determining their effort input, where effort choice ultimately determines the extent of the firm's x-efficiency.

Unlike Leibenstein's foundational work on efficiency wage theory, x-efficiency theory applies to economies at any stage of the development process. Moreover, he delinks his x-efficiency narrative from any discussion of wages. Indeed, he assumes that the wage rate is constant and that effort input varies independently from variations in the real wage. Therefore, ceteris paribus, increases in effort input reduce unit cost, whereas reductions in effort serve to increase unit cost. This follows from equation 1b, where unit cost is determined by the wage divided by average product. Given that effort variation affects average product, effort variations drive changes in average cost. Average cost is minimized only when effort input is maximized. Unlike efficiency wage theory, there is no one wage (the efficiency wage) that minimizes unit cost or maximizes net profit, since here effort varies independently from changes in real wages. In this scenario, one measure of x-inefficiency is the difference in unit cost in the x-efficient and x-inefficient firms. X-inefficient firms are relatively high-cost producers, requiring some form of protection to survive, such as the degree of freedom with regard to pricing afforded by imperfect product markets. For this reason, Leibenstein makes the case that if one intensifies the degree of competitive pressure, firms become increasingly x-efficient. Under a perfectly competitive product market environment firms would be x-efficient. Economic agents would have no choice but to work as hard and as well as they can. Under a perfectly competitive regime the conventional and x-efficient models converge in terms of their predictions with regard to x-efficiency—firms operate along the maximum production possibility frontier, and unit cost is minimized across firms. Leibenstein argues that firm decision makers will actively invest in sheltering activities that provide protection to x-inefficient firms, such as tariffs, subsidies, cartels, and artificial monopolies. To the extent that firms can successfully shelter themselves from competitive pressures, society will be materially worse off. Indeed, Leibenstein argues that one critical determinant of per-person real incomes is a high degree of x-inefficiency and that a crucial determinant of growth, in terms of Solow's residual, is improvements in the extent of x-efficiency (Leibenstein 1966, 1973a, 1978, 1979, 1983).

Preference, Effort Variation, and X-efficiency

What drives effort variation, given imperfect product markets, are variations of managerial effort in both their quality and quantity dimensions. Leibenstein (1966, 1978) assumes that firm decision

makers have a preference for working less hard and well. Although Leibenstein objects to the use of maximization-related concepts, in his model firm decision makers are in effect maximizing their utility in terms of an objective function that includes a preference for not working as hard as neoclassical theory assumes all economic agents should and do work. This is true for all economic agents. Leibenstein maintains that all individuals have objective functions that deviate from the neoclassical ideal of effort maximization. This relates to the personality of the typical individual and the relationship of the individual (personality), her or his peers, and external constraints.

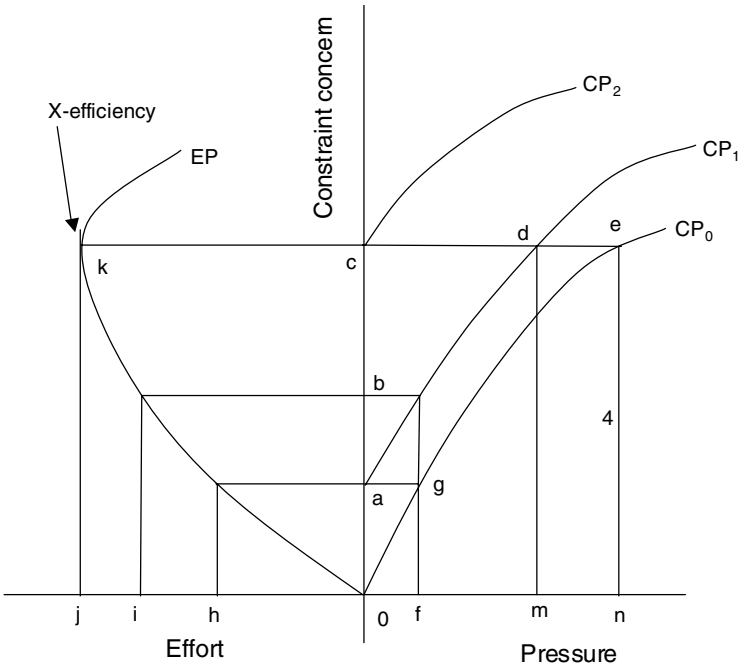
> Personality is defined in terms of (b) a taste for responsiveness to opportunities and constraints within certain standards of behavior and (c) a simultaneous taste for "irresponsible" or unconstrained behavior. The standards of behavior include moral constraints reflected in attitudes toward trust, honesty, lying, altruism, group solidarity, sacrificing for group objectives, and so on, which play a role in behavior. The compromise between (b) and (c) that the personality makes leads to a most "comfortable" degree of internal pressure [(f)] which, if everything were in the person's control, would determine the degree of maximization deviation. . . . However, the economic context (e) in which the individual finds himself may impose a higher degree of external pressures [(d)] than desired internal pressure, and hence the interaction between (d) and (f) determines the actual degree of maximization deviation. That is, pressures from peers (horizontal relations) and authorities (vertical relations) within the firm determine a degree of selective rationality different from what the person would choose in the absence of such pressures. (Leibenstein 1979, 485; see also Leibenstein 1978)

The comfortable level of internal pressures yields an equilibrium level of what Leibenstein refers to as constraint concern, which yields the utility-maximizing level of effort input, absent external pressure. The neoclassical individual's equilibrium level of constraint concern yields x-efficient effort behavior even the absence of external pressure; not so in Leibenstein's modeling.

The relationship between constraint concern, external pressure, and the level of x-efficiency is illustrated in Figure 7.7, which builds upon the work of Leibenstein (1978, 1979). Leibenstein assumes a positive relationship between constraint concern and external pressure, with pressure as the independent variable, and between constraint concern and effort input, with constraint concern being the independent variable. There are diminishing returns to external pressure with respect to constraint concern and to constraint concern with regard to effort input. X-efficient effort input (j) can only be generated with sufficient constraint concern (c). In the neoclassical narrative the segment ce would best represent an individual's constraint concern function, wherein a level of constraint concern consistent with an x-efficient effort input prevails irrespective of the level of external pressure. In Leibenstein's narrative, CP_0 or CP_1 could represent an internally derived individual's constraint concern function, both of which yield a less than x-efficient effort input at zero or low levels of external pressure. For CP_0, *n* level of constraint concern is required to generate x-efficiency, whereas for CP_1, a lesser level of effort is required to generate x-efficiency. Constraint concern functions can shift upward as an individual becomes more internally motivated to contribute more effort to the production process independent of the level of external pressure. When it shifts to point c, one has effort behavior equivalent in results to the neoclassical individual.

For Leibenstein, x-efficiency is obtained only under high levels of external pressure, given individuals' assumed preference for less than x-efficient levels of constraint concern. Once such pressure is relaxed, ceteris paribus, the extent of x-inefficiency increases as the level of constraint concern and thereby effort input is reduced. Thus, an x-efficient scenario is highly unstable given that firm decision makers are not maximizing their preferred objective function if they are

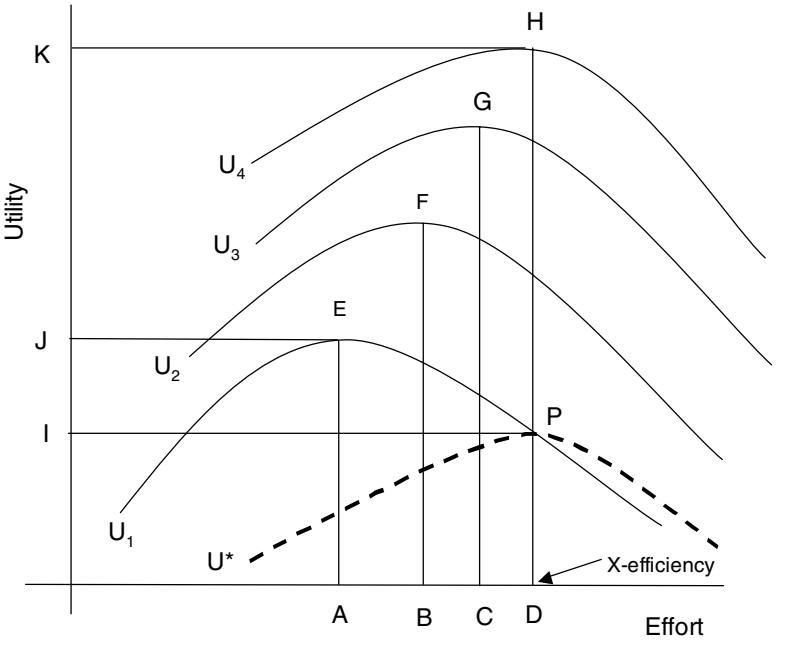
Figure 7.7 **Constraint Concern, Pressure, and Effort**



performing relatively x-efficiently. Given the assumed preference or objective function of the individual, there is an incentive for individuals to engage in sheltering activity, which provides firms with protection from competitive pressures allowing for higher and more preferred levels of x-inefficiency.

The instability of the x-efficient choice in terms of a utility-maximizing scenario is illustrated in Figure 7.8, where utility is mapped against effort input. The effort functions are concave and, as in Leibenstein's discourse, illustrate first increasing utility with increasing effort input and then diminishing utility for further increases (Leibenstein 1983, 834). X-efficiency is realized with D of effort input. Only an individual characterized by a utility function such as U_4 maximizes utility at an x-efficient level of effort input. For Leibenstein, the typical economic agent is characterized by a utility function that lies below U_4, such as U_3 and U_1. For example, if an individual with the U_1 utility function maximizes utility, only A of effort input is provided and the firm is x-inefficient. However, the individual might be forced into providing more effort input—that is, coerced into a higher level of constraint concern—by dint of competitive pressures. Indeed, these pressures might be so severe as to drive the individual into providing D of effort. In this case, the individual can be said to be maximizing her or his subutility function at P, yielding I of utility as compared to maximizing effort input at A, which would yield a higher level of utility of J. Any relaxation of market pressure provides the individual with the opportunity to move to her or his preferred equilibrium point, one consistent with x-inefficiency. Thus, in this scenario, with individuals of type U_1, x-efficiency is highly unstable in the sense that it is not a preferred equilibrium. Neoclassical theory assumes that individuals are characterized by preferences such as U_4. Relaxation of competitive pressures yields only trivial allocative inefficiencies, whereas for Leibenstein's economic agents the consequences of imperfect competitive pressures are much more severe.

Figure 7.8 **Effort Inputs and Utility**



Given such x-inefficient preferences, the secret to significantly improving the level of society's material well-being lies in making an economy more competitive. In terms of public policy, if improving material welfare is an important objective, Leibenstein's theory is supportive of more competitive product markets given that they serve to improve the level of x-efficiency. Indeed, competitiveness has much more of an effect upon material welfare than conventional theory could predict. Conventional micro theory assumes that firm decision makers (indeed, all economic agents within the firm) maximize some objective function that is consistent with the firm minimizing unit cost—in other words, producing x-efficiently. The level of unit cost has no effect on effort input. In Leibenstein's model, the competitive unit cost has a recursive effect on effort choice. In other words, once the competitive unit cost is identified, economic agents are expected to choose the *minimum* effort input required to achieve this cost. This does not preclude the firm maximizing profit in the sense of behaving in a fashion consistent with equalizing marginal cost with marginal revenue, except that in the x-inefficiency narrative unit cost and marginal cost are higher than they would be in a world where x-efficiency prevails.

The Entrepreneur and X-efficiency

As Leibenstein builds his x-efficiency modeling framework his arguments become more nuanced, wherein he delves in more detail into the micro-micro (inside the firm) reasoning underlying the raison d'être of economic agents performing x-inefficiently. One important aspect of this refinement is Leibenstein's discourse on entrepreneurship, which relates to his focus on the importance of the behavior of managers and owners to the level of x-efficiency achieved by the firm. Entrepreneurs serve the opposite role of those engaged in sheltering activities; individuals who are rent seekers of one type or another. As part of his general x-efficiency theory, Leibenstein's theory of

entrepreneurship assumes market imperfections. Given these imperfections, there is a role to be played by entrepreneurs.

> Consistent with X-efficiency theory we have to start with imperfect markets. Hence, our markets are characterized by gaps in input availabilities as well as obstacles that manifest themselves in less than equal access to inputs. Hence, in order to start a new firm, an entrepreneur must have the capacity to overcome the gaps and the obstacles in imperfect markets. He also has to be an input completer in the sense that less than the complete marshaling of inputs would not fulfill the entrepreneurial function. Obviously there are limits to the supply of gap fillers and input completers. (Leibenstein 1979, 490)

Leibenstein (1968, 74–76) adds that other important gap-filling roles of the entrepreneur include motivating firm members, providing leadership, and solving existing and potential crises. These roles are significant where economic agents have discretion over effort input. The concept of entrepreneurs as leaders and motivators speaks indirectly to seeing firms as multiagent entities, where productivity critically depends on the relationship between agents—an argument that Leibenstein eventually introduces into his modeling framework. Given a shortage of entrepreneurs that cannot be obviated by market forces alone, x-inefficiency can persist over time.

However, to the extent that this shortage can be overcome, Leibenstein argues that it would be possible for economic outcomes to mimic those of the neoclassical ideal, thereby generating an x-efficient economy.

> Our theory also allows for the special case in which entrepreneurship is so vigorous that it induces a very high degree of tightness, which in turn forces a degree of pressure so that firms either minimize costs or do not survive. However, we must note that this is a very special case, one not likely to exist in most industries. Thus the micro-micro theory enables us to distinguish the environmental micro-micro conditions for perfect competition as against what is probably the normal situation of some form of imperfect competition, and hence the case in which costs are not minimized. (Leibenstein 1979, 492)

Although Leibenstein (1968, 79–83) never develops a clearly specified model of the entrepreneur, he specifies relatively higher product price—and one would presume that this reflects relatively higher profits—as well as entrepreneurial-specific education, training and experience, and culture (related in part to personality types) as key determinants of supply. In addition, Leibenstein underlines the importance of institutional variables that affect the supply of entrepreneurship. Leibenstein argues that the "sociocultural and political constraints which influence the extent to which entrepreneurs take advantage of their capacities, and the degree to which potential entrepreneurs respond to different motivational states" are critical determinants of entrepreneurial supply (Leibenstein 1968, 78–79). Thus the supply curve of entrepreneurship is a positive function of relative price, training and experience, education, and culture. He maintains that there typically remains a persistent gap between the supply and demand for entrepreneurship, which appears to a large extent to be given by the assumption that individuals typically prefer less than x-efficient behavior or less than ideal effort input. Demand is given by the potential x-efficient level of output. In Leibenstein's model, the persistent gap between supply and demand assumes that product markets are sufficiently imperfect to allow for relatively x-inefficient high-priced production. However, in his model of entrepreneurship it is not at all clear why the deficiency in entrepreneurial supply would be eliminated simply as a product of more forceful competitive pressures given

the multifarious nature of supply, so much of which is not directly related to market forces. Moreover, in his narrative on entrepreneurship, competitive pressures critically depend on there being a sufficient supply of x-efficient-type entrepreneurs. There is a recursive relationship between competitive pressures and the supply of entrepreneurship.

The Multiagent Firm and Determinants of X-efficiency

One facet of Leibenstein's modeling of entrepreneurship is its effective highlighting of motivational issues in multiagent firms when effort discretion exists. Leibenstein discusses the potential significance of conflicting preferences among economic agents for determining the level of x-efficiency (Leibenstein 1978, 1979, 1982, 1983, 1984). Such conflict, if not resolved, can result in a prisoner's dilemma (or chicken, or mutual threat) game solution to the productivity problem, yielding a minimal level of productivity. In this narrative, unlike in the traditional x-efficiency narrative, economic incentives as determinants of productivity come into play, as they do in the efficiency wage discourse. Also of significance is the role of conventions in preventing a prisoner's dilemma solution from prevailing. Unlike in traditional x-efficiency theory, in this multiagent narrative the relationship of unit cost to productivity is no longer clear since productivity and labor compensation are implicitly linked. What is articulated in the prisoner's dilemma narrative, however, is that when effort discretion exists, a multiplicity of possible sustainable equilibrium solutions are available to the productivity problem based upon the effort conventions adopted in the firm. It is implicitly assumed that whatever the resulting unit cost, the market structure is such that it will provide protection to x-inefficient firms.

Why model the firm in terms of a prisoner's dilemma scenario?

> A basic criterion is that a prisoners' dilemma occurs wherever there are possibilities for adversarial behavior between the parties, and by all parties, which reduces the joint cooperative outcome. Now, it seems reasonable to presume that adversarial behavior between employees and the firm will usually decrease productivity, while cooperative behavior will increase it. This is certainly not the case in all types of games. But this is the case for the particular "game" set where productivity is the outcome. (Leibenstein 1982, 94)

But embedded in the prisoner's dilemma game is the potential for cooperative, relatively high-productivity solutions to the productivity problem. Leibenstein (1979, 493) argues that there are two components to the productivity problem: one relates to the determination of the size of the pie, while the second relates to the division of the pie. Looked upon independently, all agents can jointly gain by increasing the pie size, but optimal pie size is determined by the division of pie size, which involves winners and losers. However, Leibenstein writes, "the situation need not be a zero-sum game. Tactics that determine pie division can affect the size of the pie. It is this latter possibility that is especially significant" (1979, 493). Productivity is not determined independently of distribution, as it is in standard theory. Thus distribution also impacts upon the level of x-efficiency achieved by the firm—a point not discussed by Leibenstein.

Modeling the firm in terms of two types of agents, employees and employers, Leibenstein argues that employees can follow at one extreme the Golden Rule option and at the other a "maximization private satisfaction" option:

> Under the Golden Rule every employee acts in the best interest of the firm. He treats the firm as he would like the firm to treat him, and puts forth effort as if the enterprise was his

> own. The alternative option is at the other extreme: the individual works as little as possible in the firm's interest and does other things (on the job) to pursue his own private interests. (Leibenstein 1982, 92–93)

The employer faces a symmetrical set of extreme options:

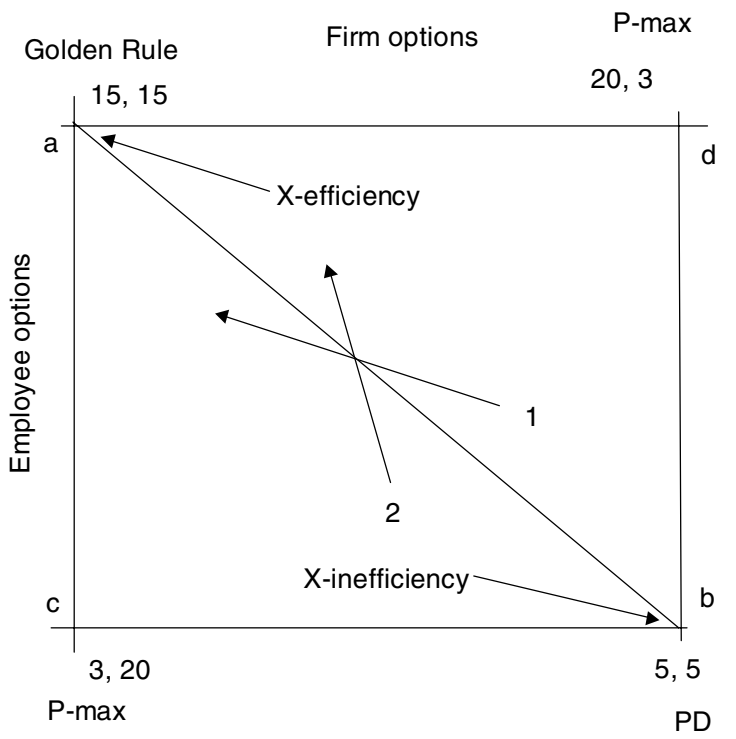
> It could behave in a Golden Rule fashion in which it provides employees with the maximal conditions, salaries, and security, consistent with "sustainable profits"; it is as if the firm operates *almost* entirely in the interest of the employees. The other alternative is parametric maximization, which implies cost minimization. That is, the firm *attempts* to minimize working conditions and wages cost while trying to get the most effort from employees. (Leibenstein 1982, 93)

In terms of Leibenstein's payoff structure, the incentive exists for employees and employers to each choose strategically the private maximization option, in the hope that the other will choose the Golden Rule option and with full knowledge that if the other party also chooses the private maximization option he/she is better off doing the same. In the latter case one is simply minimizing one's losses. Agents expect the worst from each other (private maximization) given antagonistic labor-management relations. Agents are never actually better off when one chooses the private maximization option given that one's opposite make symmetrical choices. Only if, for example, employees choose the Golden Rule, either naively or in full knowledge that employers will maximize privately, will employers be better off than if all agents cooperate to achieve the Golden Rule productivity outcome. In a repeated game scenario, which Leibenstein does not explicitly discuss, it is unlikely that employees would repeat the Golden Rule choice if employers choose to maximize privately unless employees' utility is maximized by minimizing their own material well-being. Only if the Golden Rule option is chosen by all agents simultaneously (the cooperative solution) is pie size maximized, while the individual maximization option minimizes pie size.

Diagrammatically Leibenstein posits a mutual-gain diagonal joining the Golden Rule cooperative solution with the prisoner's dilemma individual maximization solution. In Figure 7.9 this is given by line segment ab. The Golden Rule cooperative solution is given by point a and the prisoner's dilemma solution by point b. The former requires more effort input than the latter. Also, as one moves from b to a, wages and profits both increase proportionally. The cooperative solution also represents maximum x-efficiency, whereas the private maximization solution (P-max) represents minimum x-efficiency, thus the highest degree of x-inefficiency. Employees and employers are better off, given adversarial relationships, if they select the private maximization (P-max) option, since each has a claim on 5 units of output, whereas if one party selected the Golden Rule and the other chose the P-max option, the cooperator would end up with only 3 units of output. If agents engaged in private utility maximization strategies give such payoffs, firms would be producing at minimum levels of x-efficiency.

But Leibenstein maintains that the prisoner's dilemma solution does not obtain for institutional reasons. Building upon the extant game-theoretical literature, conventions represent a potential solution to prisoner's dilemma problems, where the possibility exists of multiple equilibria that yield a range of outcomes, from one just greater than the prisoner's dilemma solution to the cooperative solution. Any of an array of possible conventions yields a productivity solution greater than what could be achieved in the absence of conventions when rational agents

Figure 7.9 **Prisoner's Dilemma and X-inefficiency**



engage in private maximization. Conventions can yield solutions along the mutual benefit curve ab, where both employees and employers gain equally from productivity increases. But solutions can also be found along curve 1 or 2, where output increases but income is redistributed to labor along 1 and to the firm along 2. In both of these scenarios, both employees and employers are better off in absolute terms than under a less cooperative and less effort-intensive regime. But in the absence of conventions, employees and employers will revert to a private maximization prisoner's dilemma option.

> At any specific time the latent prisoners' dilemma possibilities are held in abeyance by conventions, institutions, and laws, involving trust, enforcement of contracts, etc. . . . If the adversarial portions are absent, then the mutual choice is the optimal position on the cooperative diagonal. Adversarial portions of the payoff table may be made essentially nonaccessible through nonmarket conventions, such as trust, honesty, fairness, legal recourse for misunderstanding or fraud, emphasis on reputation for fair dealings, etc. Thus, a convention of honesty in contractual relations eliminates adversarial behavior in which both sides attempt to cheat the other. Similarly, an effective low-cost system of laws which enforces contracts may minimize the inducement to use other types of adversarial behavior. (Leibenstein 1982, 96–97)

For this reason, "two countries with identical inputs, identical knowledge, identical capital accumulation, and the same level of employment, may yet produce significantly different out-

puts because production takes place on the basis of very different effort conventions" (Leibenstein 1982, 96).

Cooperative Determinants of X-efficiency

Related to this discourse, Leibenstein raises the importance of voluntary increases in effort input, as opposed to monitored or coerced increases in effort input in achieving x-efficient levels of output. He stresses that monitoring alone can backfire in terms of its productivity effect, given the existence of effort discretion:

> What fairly extensive monitoring essentially does is to set up something that approximates a mutual threat game. The firm stresses that if a minimal (or agreed upon) level of performance is not achieved, then the individual will be punished; for example, fired. The individual employee, and employees as a group, on the other hand, are in position to impose some damage to the firm by lowering effort, or strikes, and/or by contributing to labor turnover costs. (Leibenstein 1983, 838)

Leibenstein argues that x-efficiency can best be realized when firm members "interpret their jobs in such a way that they made effort choices which involved cooperation with peers, superiors, and subordinates, in such a way as to maximize their contribution to output" (1978, 206). Moreover:

> The main general point is that merely obtaining an acquiescent nonshirking effort is of limited value. Freely offered effort, inclusive of attentiveness and caring about the quality of effort, in return for what is viewed as a good deal (in the long run) is likely to result in higher productivity. (Leibenstein 1983, 838)

This involves a limited use of monitoring, "and instead resorts to other motivating forces, which in essence involve higher levels of trust and lower implicit adversarial relations. But such relations are likely to involve quite a bit of discretion on all sides" (Leibenstein 1983, 838). Thus, given effort discretion, effort input can move beyond prisoner's-dilemma-type outcomes only if agents develop conventions that are based on trust and which are therefore self-enforcing via peer pressure or constraint concern emanating from the individual her or himself. Just as Costa Nostra detainees need not succumb to the prisoner's dilemma option for reasons of fear and loyalty, which yield sufficient trust that the other party won't talk, cooperative labor-management relationships can have the same effect in the firm, keeping employees and employers from choosing relatively x-inefficient productivity options.

Leibenstein's narrative of the multiagent firm raises significant issues with regard to the determinants of x-inefficiency. Importantly, it raises the importance of incentives as a determinant of effort input, which is key to the various extant models of efficiency wages. But the connection between wages, effort input, unit cost, and net profit remains quite vague. Moreover, unlike in his foundational x-efficiency narrative, the connection between individual preferences, effort inputs, and the extent of competitiveness or environmental tightness is amorphous. What is clear from his multiagent narrative is that market forces remain significant in the background as increasing competitive pressures force agents to behave more x-efficiently—in this scenario by revising effort conventions—so as to maintain unit cost at a competitive level. But imperfect product markets provide the shelter behind which x-inefficiency can persist as equilibrium (albeit unstable equilibrium) behavior.

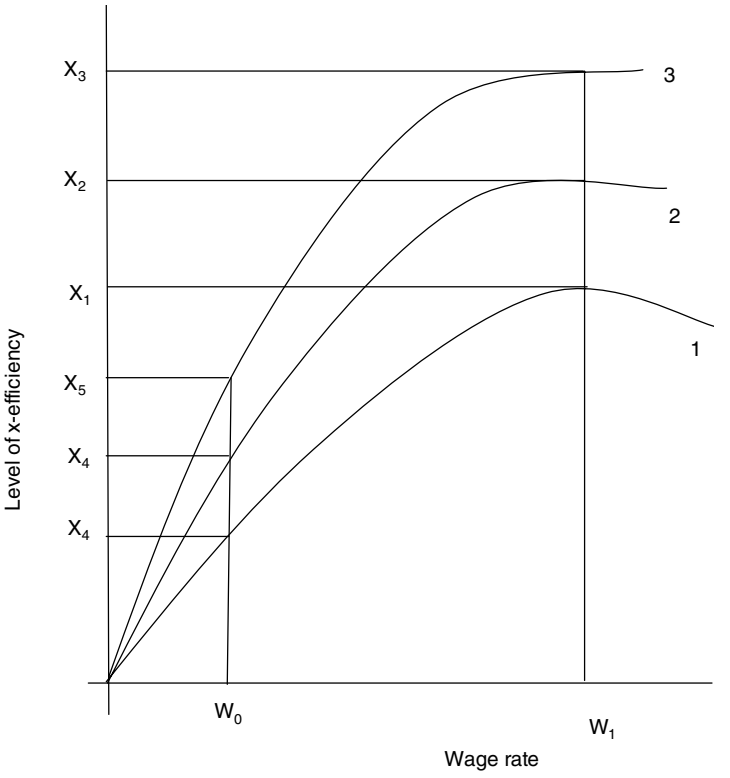
A BEHAVIORAL MODEL OF THE FIRM AND X-EFFICIENCY THEORY

The behavioral model of the firm developed in Altman (1992, 1996, 1998, 1999, 2001b, 2002) and discussed above in the context of efficiency wage theory helps to fill some of the gaps contained in x-efficiency theory, providing a modeling framework to address significant paradoxes related both to the conventional wisdom and to x-efficiency theory itself. Common to the behavioral model, conventional x-efficiency, and efficiency wage theories, effort is assumed to be a variable input in the production function. Critical to the behavioral model is its joining of the concepts of the causal relationship between incentives and productivity contained in efficiency wage theories and that of the determinants of the level of x-efficiency.

Unlike in x-efficiency theory, where effort variability is discussed in the context of fixed levels of labor compensation, in the behavioral model, although effort can vary independent of material incentives, the latter can play a critical role in determining the extent of x-efficiency. Thus, unlike in x-efficiency theory, a direct causal link is explicitly posited between labor compensation, effort input, and labor productivity. This point is illustrated in Figure 7.10, where curves 1, 2, and 3 are wage x-efficiency curves that are assumed to be subject to diminishing returns. Wages are taken to represent the overall material incentive package of the firm. The level of x-efficiency is a positive function of the real wage. For each curve, the level of x-efficiency is given by the wage rate, such as W_0 or W_1. In contrast, in x-efficiency theory the level of x-efficiency is determined by the preferences of managers/owners. The level of x-efficiency can vary independent of the wage (though this argument is somewhat refined, but only vaguely in terms of modeling specificities, in Leibenstein's prisoner's dilemma scenarios). In the behavioral model the level of x-efficiency can be maximized only by providing economic agents (who can include managers) with an appropriate material incentive package, given by W_1 and wage x-efficiency curve 3. We have three such curves to reflect the reality, captured in the overall x-efficiency narrative, that nonmaterial incentives can affect the level of x-efficiency. Material incentives alone won't do the trick. Thus, maximizing the level of x-efficiency requires the appropriate mix of material and nonmaterial incentives, although material incentives are typically quite important. In the neoclassical framework the level of x-efficiency is always at a maximum, irrespective of the wages and preferences of economic agents.

Another feature of the behavioral model is illustrated in Figure 7.11, where the potential relationships between the level of x-efficiency and the level of unit cost are mapped. In the behavioral model, the possibility of the level of x-efficiency not affecting the level of unit cost is allowed for. This is given by the linear unit cost x-efficiency curves BM_0 and BM_1. Underlying differences in the level of x-efficiency are differences in the level of labor productivity, which in turn are driven by differences in the level and quality of effort input. X_M represents the maximum level of x-efficiency. Bear in mind that the level of x-efficiency is determined by the level of productivity. Thus, maximum x-efficiency is a product of productivity maximized with regard to effort input. Unit cost is inelastic with respect to changes in the level of x-efficiency if, in the background, productivity increases are just offset by changes in the level of input costs such as labor compensation. Therefore, a wide array of levels of labor compensation is consistent with some unique unit cost of production. Increasing wages, for example, need not increase unit cost, while cutting wages need not reduce unit cost. This refers back to the discussion surrounding equations 1a and 1b and Figure 7.4, above. In this case, neither maximizing nor minimizing the level of x-efficiency affects the level of unit cost and thereby the competitive position of the firm. The two unit cost x-efficiency curves reflect the possibility that firms can be characterized by different firm

Figure 7.10 **Wages and X-efficiency**

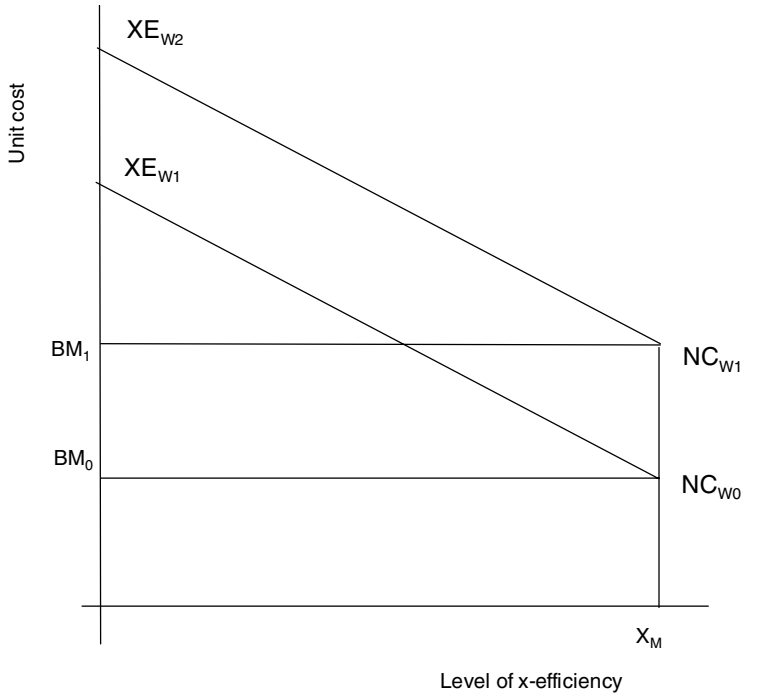


cultures such that the ratio of productivity to labor cost is greater in one firm than in another. At one extreme, at maximum x-efficiency, one firm might produce at a lower unit cost than another because firm members are more productive in one firm than in another at a lower wage. But this is sustainable only if the higher-cost firm is sheltered from competitive pressures.

The prediction derived from the traditional x-efficiency model that levels of x-efficiency are negatively related to unit cost is illustrated by unit cost x-efficiency curves XE_{W1} and XE_{W2}, where it assumed that the level of x-efficiency varies independently from wages. Relatively x-inefficient firms are less cost-competitive and can survive only when sheltered from competitive pressures. In the conventional neoclassical model, since all firms are assumed to be x-efficient at all times, it is differences in the rate of labor compensation that can have a determining effect on the firm's competitiveness. This is illustrated by points NC_{W0} and NC_{W1} in Figure 7.11, where firms are producing at maximum x-efficiency—there is no x-inefficiency—and differences in unit cost are driven by differences in the rate of labor compensation. In the behavioral model, unlike in the x-efficiency model, different levels of x-efficiency need not affect the state of a firm's relative competitiveness. And unlike in the conventional neoclassical model, in the behavioral model differences in the rate of labor compensation need not affect a firm's capacity to compete. High-wage firms need not be less competitive than low-wage firms.

A critical distinction between the behavioral model and x-efficiency theory is that in the behavioral model x-inefficiency can persist in a variety of product market structures, inclusive of a highly competitive one, since lower levels of x-efficiency need not translate into higher unit cost

Figure 7.11 **Unit Cost and X-efficiency**



and higher levels of x-efficiency need not translate into lower unit cost. Firms characterized by high and low levels of x-efficiency can all be cost-competitive. Therefore, even perfect product market competition need not yield an x-efficient economy. On the other hand, x-efficiency can prevail even with highly imperfect product markets. In addition, in the behavioral model, the level of x-efficiency can be driven by production costs, of which labor costs are a critical component. Given market structure, the level of x-efficiency can be driven by differences in the level of labor compensation. Thus, high-wage firms or economies can be relatively x-efficient even if contained in a relatively sheltered product market environment. Also, unlike in efficiency wage theory, in the behavioral model there is no one wage (the efficiency wage) that is an equilibrium wage, one that minimizes cost and maximizes net profit. The behavioral model can also better explain the existence of multiple equilibria in terms of different levels of x-efficiency and also provide the underlying rationale for why superior x-efficient regimes can be rejected by firms even in light of competitive pressures—relatively x-efficient firms need not have a competitive advantage over relatively x-inefficient firms.

By explicitly linking material incentives to the determinations of levels of x-efficiency the behavioral model helps explain potential behaviors of economic agents and the incentives for producing more or less x-efficiently. In Leibenstein's multiagent modeling framework we are left with the impression that, given imperfect product markets, anything goes with regard to the level of x-efficiency, contingent upon the firms' work-related conventions. In the behavioral model, material incentives and work environment drive the level of x-efficiency, with explicit implications for cost and profit. In the behavioral model, to the extent that the Golden Rule and the prisoner's dilemma options yield the same unit cost and the same profit, both options are sustain-

able in either a competitive or noncompetitive environment. There is no Darwinian imperative driving the economy to any one option. Moreover, in this case, only labor benefits as one moves from the prisoner's dilemma to the Golden Rule option, since higher wages are required to enhance the level of x-efficiency. In this scenario, conventions can play an important role in determining the direction in which, with regard to the level of x-efficiency, the firm will move. Moreover, to the extent that firm decision makers prefer the prisoner's dilemma option—they don't materially benefit from the Golden Rule option and it requires more effort to achieve it—higher levels of x-efficiency are driven by the capacity of labor to achieve higher wages and improved working conditions in the context of the market economy. Once this capacity is diminished, one can expect that the firm would revert toward the prisoner's dilemma option.

This takes us back to our earlier discussion of the relationship between preferences and the level of x-efficiencies (Figure 7.8). When employees and employers have different objective functions—in this case, employees are willing to work relatively x-efficiently given the appropriate work environment, while employers would prefer less x-efficiency given workers' preferred environment—employers are maximizing only their subutility functions, as opposed to their preferred utility in the high-wage environment. This is an unstable Golden Rule equilibrium unless employer preferences change over time toward a preference for the higher-wage Golden Rule option. Thus, when employers do not clearly benefit from more x-efficiency—even if there is no zero-sum game involved—the behavioral model suggests that the extent of x-efficiency critically depends upon the power relations between employees and employers, the extent to which employers are sympathetic to the material and spiritual well-being of their employees, and, related to this, the capacity of the firm to build cooperative relationships between employees and employers. One should note that in the behavioral model there is no need to rely on notions of choice irrationality to generate any of its substantive analytical predictions. Agents can be maximizing their respective utility or subutility functions. But if individuals have conflicting objective functions or face severe constraints on their choice options, individuals' constrained utility-maximizing choices can be highly unstable, changing quite dramatically when individuals are afforded the opportunity. What is critical to this narrative of effort discretion are the determination and analysis of the preferred effort input under various circumstances, and the set of circumstances under which utility-maximizing individuals or groups of individuals would have stable preferences for x-efficient behavior.

In many of Leibenstein's multiagent scenarios (see Figure 7.9), all parties benefit as one moves toward the Golden Rule option, although it is possible for the increasing income to become more unequally distributed toward employees or employers. In the behavioral model, two distinct behavioral possibilities are suggested. One would be in the context of improvements in x-efficiencies, yielding no changes in unit cost. The other context would be when increases in x-efficiency yield not only increases in income to employers and employees, but also a reduction in unit cost. In both cases, increases in the level of x-efficiency can be provoked by increases in wages and improvements in working conditions that pressure firms into becoming more x-efficient. In the first case, however, rejection of or struggling against the Golden Rule option would incur a clear opportunity cost to employers—a loss of income or profit. Such a sacrifice could be made, although economists would expect that it would be a negative function of the level of increased income expected to accrue to employers with increases in the level of x-efficiency. As this expected income increases, one would expect the probability of engaging in such self-sacrifice to diminish, or the percentage of employers choosing such a sacrifice to diminish. But in this scenario competitive pressures per se cannot force firms into becoming more x-efficient. The preferences of employers and the bargaining power of employees play key roles in determining the level of x-efficiency.

In the second context, not only are employers sacrificing income, but firms are also producing at a higher unit cost when producing x-inefficiently. This scenario is most in line with Leibenstein's origin x-efficiency model. Economic agents—in particular, employers (inclusive of managers)—have a strong leisure preference. This yields x-inefficiency along with higher unit costs. Such firms can survive only when sheltered from competitive pressures. It is only this particular scenario that is consistent with the notion that increasing competitive pressures will drive firms into becoming more x-efficient. The behavioral model suggests that this scenario is only one of a variety of possible x-efficiency scenarios, especially when one gives proper attention to the connection between material incentives and overall working conditions and the level of x-efficiency.

The above scenario, which is a special case of the general framework of the behavioral model, is consistent in unit cost results with scenarios wherein effort varies irrespective of material incentives, which is the chief concern of Leibenstein's foundational x-efficiency narrative. Management drives the level of x-efficiency through its preference functions that lean toward x-inefficient behavior. It is assumed that the quantity and quality of effort input among all firm members is a function of managerial effort input or that variations in managerial effort input suffice to drive firm productivity. Thus, irrespective of the incentive system in place (inclusive of the system of industrial relations), firms can achieve x-efficiency in production as long as management has the will and the wherewithal to do so. In this case, ceteris paribus, one would expect that low-wage economies should dominate high-wage economies if there is a supply of sufficiently x-efficient managers and entrepreneurs, unless the relatively high-wage economies are sheltered from the competitive threat posed by the low-wage x-efficient economies. This must be the case if low-wage economies can achieve x-efficiency irrespective of how employees are rewarded and treated. It is here implicitly assumed, as it is in conventional neoclassical theory, that workers will work as hard and as well as they can, irrespective of wages and working conditions. Get rid of shelters and firms will be forced into becoming more x-efficient. However, even given Leibenstein's assumptions, competitive pressures yield x-efficiency only if the supply of x-efficient management and entrepreneurs is sufficient to generate the competitive pressures required to make all firms x-efficient, including those that have management with x-inefficient preferences. Competitive pressures alone are not enough.

As opposed to Leibenstein's foundational x-efficiency narrative, in the behavioral model, as is the case in Leibenstein's prisoner's dilemma narrative, at least implicitly, the level of effort input is a function of the incentive structure within the firm, inclusive of material rewards. In this scenario, increasing effort input and productivity are tied to higher input costs. Thus, increasing the level of x-efficiency might have no effect on unit cost, while reducing the level of x-efficiency, accompanied by lower input cost, might also have no effect on unit cost. Entrepreneurs, as gap fillers and input completers, can play a critical role in generating x-efficiency, in part by affecting the motivational structure within the firm. What is required for x-efficiency here is entrepreneurs who can contribute toward developing Golden Rule conventions in high-wage environments. But when increasing the level of x-efficiency is not accompanied by lower unit cost, increasing the supply of x-efficient entrepreneurs need not make the economy more competitive. Rather, they critically contribute to making firms and economies relatively more x-efficient irrespective of the extent of competitive pressures.

The behavioral model of the firm, which builds on the extant x-efficiency and efficiency wage literature, makes explicit the relationship between material incentives, x-efficiency, and unit cost. It thereby specifies the opportunity cost at different levels involved in remaining or becoming relatively x-inefficient. The behavioral model provides an analytical framework that allows for persistence of x-efficiency under different competitive environments as well for multiple equilib-

ria in terms of levels of x-efficiency and in-firm conventions. Unlike in x-efficiency theory, in the behavioral model the level of x-efficiency is not largely driven by market imperfections and managerial preferences. And, unlike efficiency wage theory, it does not predict a high-wage efficiency wage equilibrium. Rather, the behavioral model of the firm provides a broad analytical framework that allows for conventional x-efficiency and efficiency predictions as special-case outcomes. More attention is paid to alternative scenarios that appear to be much consistent with realities of everyday economic life, past and present.

SOME CONTRIBUTIONS OF A BEHAVIORAL MODEL OF THE FIRM

I focus on five public-policy-related issues to highlight some of the potentially significant analytical contributions of a behavioral model. One relates to the question of convergence between low- and high-wage economies. Another issue relates to why superior work environments are not dominant. Connected to the above is the contention that globalization, in terms of more international trade, yields a race to the bottom with regard to wages and working conditions. I also briefly discuss the theory of the reciprocal firm, which predicts the dominance of relatively high-wage economies when effort discretion exists. Finally, I discuss the possibility of firms behaving ethically in competitive markets in the context of the behavioral model. The connection between levels of x-efficiency and technological change is touched upon.

Neoclassical theory predicts that convergence should take place between high- and low-wage economies in part given the cost advantage that low-wage firms have over high-wage firms in a world where all firms are assumed to be x-efficient. Such convergence has not taken place (Altman 2002; Baumol and Wolff 1988; DeLong 1988; Pritchett 1997). Rather, we've had convergence among the already high-wage economies, with some exceptions such as Japan, South Korea, and Taiwan. In the behavioral model, competitive pressures need not result in firms and economies converging in terms of wages and productivity, since differentials in wages and productivity can be offset by differences in productivity and wages, respectively. From the perspective of the behavioral model, one reason for the absence of convergence would be the low-level capacity of employees (workers and peasants) in low-wage economies to bargain for improved material benefits and standards of work that would, in turn, pressure firms into becoming more x-efficient. This raises the issue of the importance of power and institutions in the determination of economic results (Altman 2000; Rothschild 2002). Another reason for the absence of convergence consistent with the behavioral model as well as with both neoclassical and conventional x-efficiency theory is the sheltering of x-efficient firms. But the behavioral model suggests that reducing sheltering per se will not induce more x-efficiency in the absence of increased bargaining power on the part of employees and a sufficient number of x-efficiency-prone entrepreneurs. Moreover, in the behavioral model convergence would tend to take place dynamically, with the low-wage economies becoming high-wage economies as the low-wage economies catch up with the high-wage economies in terms of x-efficiency, as opposed to the high-wage economies converging downward in terms of wages and levels of x-efficiency. This model also has significant implications for understanding the implications of unions, minimum wages, unemployment insurance, and other institutional interventions on the market for economic efficiency and development (Altman 1992, 2001a, 2004).

There exist many alternative forms of industrial relationship with relatively high levels of x-efficiency, yet they do not dominate the marketplace. Neoclassical theory would predict that superior firms should rule the roost, whereas x-efficiency theory would predict that such firms should be producing at a lower unit cost and have a competitive advantage. The behavioral model makes

analytical space for the reasonable scenarios wherein firms are superior in terms of their level of x-efficiency and in terms of providing employees with a preferred set of material benefits and work environment, yet fail to dominate their product markets. To the extent that superior firms are more costly to operate, their higher level of x-efficiency might serve to keep them competitive with the relatively low-input-cost x-inefficient "inferior" firms. In this case, there would be no reason to expect the superior firms to dominate, for they have no clear competitive advantage (Altman 2002). The extent to which superior firms make their presence felt becomes more a product of the bargaining power of employees and the preferences of employees and employers. Being superior need not provide firms with a competitive advantage (see also Ichniowski et al. 1996; Tomer 2001).

Both advocates and detractors of economic globalization often suggest a race to the bottom. For the former, economies have little choice but to make labor costs compatible with the new harsher competitive realities. This will yield higher levels of per capita output, which will somehow and sometime trickle down to the society at large. Opponents of economic globalization concur with the argument that in the first instance globalization will cause many to suffer economically and socially, and such suffering will persist into the future. There will be clear winners and even more losers in this zero-sum game. The behavioral model suggests that economic globalization generates significant benefits to employees under the appropriate institutional environment. To the extent that economic globalization increases the demand for labor and thereby the bargaining power of employees, this can have the effect of increasing wages and improving working conditions and thereby pushing economies into becoming more x-efficient. As labor market pressure intensifies along with product market competition, firms must become more x-efficient to remain competitive. In the absence of economic globalization, wherein the demand for labor is less, workers' bargaining power is reduced, reducing the pressure to improve wages and other labor benefits. Globalization can therefore serve to benefit labor and need not result to a race to the bottom (and can even have the opposite effect) if the basic democratic and related labor rights are in place, providing employees with the capacity to take advantage of increasing labor demand, which typically coincides with increased economic globalization. Globalization yields no improvement to workers' well-being only in the absence of such capacity-facilitating and -building rights, and if capabilities are created that allow employers to neutralize the advantages that increasing economic globalization tends to provide workers. Moreover, the behavioral model suggests that success in an increasingly global economy does not require a race to the bottom. A race to the top is quite consistent with increasing globalization (2006b).

A model of the reciprocal firm has been developed by Fehr and Gatcher (2000a, 2000b). Their model clearly and explicitly builds on efficiency wage theory and the ultimatum game experimental literature (Güth 1995). This model overlaps significantly with Leibenstein's prisoner's dilemma narratives. A key point made in this model that has received much critical acclaim is its prediction that, given imperfect labor contract and effort discretion, firms tend to pay workers relatively high wages, wages that exceed the opportunity cost of workers. Indeed, workers are paid an efficiency wage. This is a wage determined by the capacity of workers to retaliate (by reducing their effort input) against employers who pay them wages that are deemed to be unfair. In this context it is profitable for the firm to pay employees what would appear to be an excessively high wage. One would expect that firms should tend to pay workers such relatively high wages in both the short and long terms.

Both efficiency wage theories and the behavioral model are compatible with short downward inflexibility of real wage. In the behavioral model, although cutting real wages need not result in reduced net profit or increased unit cost, as it would in efficiency wage models, it also need not provide any persistent benefits in terms of higher net profit or reduced unit cost.

However, if firm decision makers view their capacity to reduce real wages as simply a product of the downward segment of the business cycle, which they expect will not last, cutting real wages might be viewed as engendering short-term costs in terms of loss of trust, increased labor-management conflict, and restructuring, which are not economically viable if real wages will have to be restored as economic conditions improve (Bewley 1999). The behavioral model, however, suggests that firms need not converge toward some high wage equilibrium, as is suggested by the model of the efficiency-wage reciprocal firm. In the behavioral model the possibility of multiple equilibria with regard to real wage rates, effort input, and levels of x-efficiency is highlighted. A low-wage equilibrium is a very real possibility even when effort discretion exists and workers thereby have the capacity for retaliating against employers for paying them an unfair wage. This is especially true when employers believe that their capacity to drive real wages down is not cyclical but is a product of long-run fundamentals. Thus if employers have a preference for low-wage workplace environments, they can be expected to pursue this option as long as it is consistent with competitive unit costs.

This modeling perspective is consistent with a positive empirical relationship between higher wages and higher levels of x-efficiency as well as with the simultaneous existence of low- and high-wage firms and economies plus a wide range of firm types in between these two extremes. This type of scenario is also much more consistent with the stylized facts of economic life. The existence of effort discretion and the capacity of workers to retaliate against unfair wages and working conditions need not result in employees being paid relatively high wages. Rather, in terms of the behavioral model, a key determinant of whether firms and economies are relatively high-wage is the institutional capacity of workers and peasants to bargain and pressure for higher wages in a world of effort discretion, where the latter allows firms to remain competitive in a high-wage environment. A similar point is made succinctly by Domar in his discussion of the evolution of serfdom and slavery (1970). If firms are constrained into paying an efficiency wage, one should expect convergence to take place between high- and low-wage firms as workers exercise their capacity to force employers to pay them at a fairer rate. Some optimal efficiency wage should be expected to dominate by dint of circumstances. Such does not appear to be the case. The determination of firm types is much more complex than either neoclassical or efficiency wage theories would have it, and this complexity is captured in part by the behavioral model.

Can firms behave ethically in a market economy? The standard argument is that to the extent that behaving ethically incurs higher unit costs, ethical behavior is not sustainable in a competitive economy. At best ethical firms must be sheltered and society at large must bear the economic cost of some economic agents realizing their particular ethical preferences. Even if economic agents (and firm decision makers in particular) prefer that their firms be ethical, such a preference cannot be realized over the long term within the framework of a competitive market economy. Ethical can mean different things to different folks, but there is some consensus that an ethical firm is consistent with a work environment where working conditions and benefits are relatively high and the industrial relations environment is relatively cooperative. Take this as an example of an ethical firm (see Altman 2001a for a discussion of sustainable ethical production from the perspective of "green" economies). Contrary to the conventional wisdom, the behavioral model suggests that the ethical firm need not have to produce at a higher unit cost, since higher input cost can be compensated for by a higher level and quality of effort input. There is an ethical dividend that can be expected when firms behave ethically. However, this dividend need not be large enough to lower unit cost or increase net profit. Rather, the ethical firm can be as cost-competitive and as profitable as the relatively unethical firm. For this reason, the behavioral model suggests that when effort discretion exists, firm

decision makers actually do have some choice as to how ethical their firm will be. Market forces per se do not determine this choice for them, which would be the case from the perspective of the conventional wisdom. The behavioral model is therefore consistent with the simultaneous existence of both ethical and unethical firms (Altman 2005).

The focus of our discussion has been that of a world where technological change is being held constant so as to isolate the importance of effort variability for various economic issues. But technological change itself can be significantly affected by the level of x-efficiency (Leibenstein 1973b). And technological change is a critical driver of economic growth—as would be improvements in the level of x-efficiency both in the x-efficiency narrative and in the more general behavioral narratives. Technical change involves transformations to the process of production such that more output, controlling for quality, can be produced with fewer inputs. But as Leibenstein points out, new technologies need not be adopted if the associated increases in productivity require higher, or even x-efficient, levels of effort input and this is not within the preference domain of decision makers. In addition, if technological change requires increased effort inputs from employees who will not benefit from this change, the required effort inputs will not be forthcoming. "Best practice" technology from the perspective of conventional theory will not be adopted. More specifically, I develop elsewhere an induced theory of technological change, where technological change is induced by increases in factor input costs (Altman 2001b, ch. 6; 2005). High-wage firms, for example, are induced to adopt new technology that serves to keep unit cost competitive with the relatively low-wage firms. The high wages also induce economic agents to input the necessary effort required to make the new technology cost-effective. The relatively low-wage x-inefficient firms need not adopt the new technology if it will not serve to lower unit cost or enhance profit. This would be the case when the new technology requires higher levels of effort input, which in turn requires higher wage rates. This model helps explain why more productive technologies are often not adopted, especially in low-wage economies, even when competitive pressures are severe.

CONCLUSION

The introduction of effort discretion into the modeling of economic agency has had a profound impact upon economic theory, although the conventional wisdom remains dominated by theories that pay little to effort discretion. Once effort discretion is introduced as a choice variable, economic narratives that deal with the determination of employment, real wages, productivity, unit cost, and competitiveness take on a different meaning. Such narratives are not at all dependent upon assumptions of irrationality or the errors and biases narrative, which has received so much recent attention in the social science and public policy literature, often under the banner of behavioral economics. The assumption of effort discretion has given rise to a rich literature that falls most generally under the rubric of efficiency wage and x-efficiency theory. The focus of the latter is the relationship of material incentives to productivity as determined by the quality and quantity of effort input. The former focuses upon how preferences with regard to effort input, in the context of different sets of incentives, affect the level of economic efficiency.

Efficiency wage and x-efficiency theory are discussed in some detail, with special focus on their underlying assumptions and the implications of these theories for understanding the micro-micro economics of firm behavior and public policy. Critical attention is paid to the foundational work of Akerlof and Leibenstein. The efficiency and wage and x-efficiency narratives are critically appraised, and extensions and revisions are suggested in terms of a behavioral model of the firm developed by Altman as it bridges and extends the foundational argument of both theories. Fundamental to the behavioral model is that it allows for multiple equilibria with regard to the

wage-effort-productivity relationship. Unlike efficiency wage theory, there need not be one unit-cost-minimizing and net-profit-maximizing wage, and involuntary unemployment need not be a product of real wages being too high. In other words, unemployment might be too high given demand-side problems and productivity being too low for marginal workers—such workers being too x-inefficient. Unlike x-efficiency theory, the behavioral model clearly links levels of x-efficiency to the material incentives that are one of the key determinants of the level of x-efficiency. The behavioral model suggests that key causal determinants of the degree of x-efficiency are real wages and working conditions. X-efficiency is not simply a function of preferences and the degree of competitiveness. The extent of institutional space provided to labor to bargain for improvements in their conditions is highlighted, as is the notion that given effort discretion, even under severely competitive markets, x-inefficiency is sustainable, as are both high- and low-wage firms and ethical and unethical firms.

In terms of the behavioral model, the existence of effort discretion provides individuals with a wide array of choices with regard to how firms and economies evolve. Choice is significantly affected by the institutional parameters within which choices are made, as well as how such parameters differentially affect the degree of freedom afforded to individuals and groups of individuals. Economic theory should be constructed to make allowance for the degree of choice that individuals actually do have in a market economy. Models incorporating effort discretion as a choice variable play an important role in this endeavor.

Future research can be expected to provide more nuanced models that incorporate the assumption of effort discretion and multiagent firms with preference functions, which deviate from neoclassical norms for good rational reasons. Such models can serve to better explain the economies in which we live. Needless to say, such models and their underlying assumptions must be tested using traditional data sets as well as through experiments and surveys. In the tradition of behavioral economics, theories should contribute to explaining the facts, with regard to both causality and analytical prediction. And this often involves introducing behavioral and institutional assumptions that are given little play in the conventional modeling narratives.

REFERENCES

Akerlof, George A. 1980. "A Theory of Social Custom, of Which Unemployment May be One Consequence." *Quarterly Journal of Economics* 94: 749–75.

———. 1982. "Labor Contracts as Partial Gift Exchange." *Quarterly Journal of Economics* 97: 543–69.

———. 1984. "Gift Exchange and Efficiency Wage Theory: Four Views." *American Economic Review, Papers and Proceedings* 74: 79–83.

———. 2002. "Behavioral Macroeconomics and Macroeconomic Behavior." *American Economic Review* 92: 411–33.

Akerlof, George A., William T. Dickens, and George L. Perry. 1996. "The Macroeconomics of Low Inflation." *Brookings Papers on Economic Activity* 1996, 1*:* 1–59.

———. 2000. "Near-Rational Wage and Price Setting and the Long-Run Phillips Curve." *Brookings Papers on Economic Activity* 2000, 1: 1–44.

Akerlof, George A., and Janet L. Yellen, eds. 1986. *Efficiency Wage Models of the Labor Market.* Cambridge: Cambridge University Press.

———. 1988. "Fairness and Unemployment." *American Economic Review, Papers and Proceedings* 78: 44–49.

———. 1990. "The Fair Wage Hypothesis and Unemployment." *Quarterly Journal of Economics* 105: 255–83.

Altman, Morris. 1990. "A Critical Appraisal of Corporate Bigness and the Transactions Cost Economizing Paradigm." In Roger Frantz, ed., *Handbook on Behavioral Economics*, 2A:217–32. London: JAI Press.

———. 1992. "The Economics of Exogenous Increases in Wage Rates in a Behavioral/X-efficiency Model of the Firm." *Review of Social Economy* 50: 163–92.

———. 1996. *Human Agency and Material Welfare: Revisions in Microeconomics and Their Implications for Public Policy*. Boston: Kluwer Academic Publishers.

———. 1998. "High Path to Economic Growth and Development." *Challenge: The Magazine of Economic Affairs* 41: 91–104.
———. 1999. "The Methodology of Economics and the Survivor Principle Revisited and Revised: Some Welfare and Public Policy Implications of Modeling the Economic Agent." *Review of Social Economics* 57: 427–49.
———. 2000. "Labor Rights and Labor Power and Welfare Maximization in a Market Economy: Revising the Conventional Wisdom." *International Journal of Social Economics* 27: 1252–69.
———. 2001a. "When Green Isn't Mean: Economic Theory and the Heuristics of the Impact of Environmental Regulations on Competitiveness and Opportunity Cost." *Ecological Economics* 36: 31–44.
———. 2001b. *Worker Satisfaction and Economic Performance*. Armonk, NY: M.E. Sharpe Publishers.
———. 2002. "Economic Theory, Public Policy and the Challenge of Innovative Work Practices." *Economic and Industrial Democracy: An International Journal* 23: 271–90.
———. 2003a. "Economic Growth and Income Equality: Implications of a Behavioral Model of Economic Growth for Public Policy." *Canadian Public Policy* 24: S87–S118.
———. 2003b. "The Nobel Prize in Behavioral and Experimental Economics: A Contextual and Critical Appraisal of the Contributions of Daniel Kahneman and Vernon Smith." *Review of Political Economy* 16: 3–41.
———. 2004. "Why Unemployment Insurance Might Not Only Be Good for the Soul, It Might Also Be Good for the Economy." *Review for Social Economy* 62: 517–41.
———. 2005. "Reconciling Altruistic, Moralistic, and Ethical Behavior with the Rational Economic Agent and Competitive Markets." *Journal of Economic Psychology* 26: 232–57.
———. 2006a. "Involuntary Unemployment, Macroeconomic Policy, and a Behavioral Model of the Firm: Why High Real Wages Need Not Cause High Unemployment." *Research in Economics* 60: forthcoming.
———. 2006b. "Economic Growth, 'Globalization' and Labor Power." *Global Business and Economics Review* 8: forthcoming.
Becker, Gary S. 1996. *Accounting for Tastes*. Cambridge, MA: Harvard University Press.
Bewley, Truman. 1999. *Why Wages Don't Fall During a Recession*. Cambridge, MA: Harvard University Press.
Baumol, William J., and Edward N. Wolff. 1988. "Productivity Growth, Convergence, and Welfare: Reply." *American Economic Review* 78: 1155–9.
Blanchflower, David G., and Andrew J. Oswald. 1995. "An Introduction to the Wage Curve." *Journal of Economic Perspectives* 9: 153–67.
Cyert, Richard M., and James C. March. 1963. *A Behavioral Theory of the Firm.* Englewood Cliffs, NJ: Prentice-Hall.
DeLong, Bradford J. 1988. "Productivity Growth, Convergence, and Welfare: Comment." *American Economic Review* 78: 1138–54.
Domar, Evsey D. 1970. "The Causes of Slavery or Serfdom: A Hypothesis." *Journal of Economic History* 30: 18–32.
Fehr, Ernst, and Simon Gachter. 2000a. "Reciprocal Fairness, Cooperation, and Limits to Cooperation." In Erich Holzl, ed., *Conference Proceedings: Fairness and Cooperation, XXV Annual Colloquium on Research in Economic Psychology and SABE 2000 Conference*, 1–7. Vienna: IAREP.
———. 2000b. "Fairness and Retaliation: The Economics of Reciprocity." *Journal of Economic Perspectives* 14: 159–81.
Fehr, Ernst, and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics* 114: 817–68.
Frantz, Roger S. 1997. *X-efficiency Theory, Evidence and Applications.* Topics in Regulatory Economics and Policy 23. Boston: Kluwer Academic Publishers.
———. 2004. "The Behavioral Economics of George Akerlof and Harvey Leibenstein." *Journal of Socio-Economics* 33: 29–44.
Friedman, M. 1953. "The Methodology of Positive Economics." In *Essays in Positive Economics,* 3–43. Chicago: University of Chicago Press.
Gigerenzer, Gerd. 2002. *Adaptive Thinking: Rationality in the Real World*. Oxford: Oxford University Press.
Gigerenzer, Gerd, and Reinhard Selten. 2001. *Bounded Rationality: The Adaptive Toolbox*. Cambridge, MA: MIT Press.
Goldstein, Daniel G., and Gerd Gigerenzer. 2002. "Models of Ecological Rationality: The Recognition Heuristic." *Psychological Review* 109: 75–90.
Güth, Werner. 1995. "On Ultimatum Bargaining Experiments—A Personal Review." *Journal of Economic Organization and Behavior* 27: 329–44.
Ichniowski, Casey, Thomas A. Kochan, David Levine, Craig Olson, and George Strauss. 1996. "What Works at Work: Overview and Assessment." *Industrial Relations* 35: 299–333.

Kahneman, Daniel. 2003. "Maps of Bounded Rationality: Psychology for Behavioral Economics." *American Economic Review* 93: 1449–75.
Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decisions Under Risk." *Econometrica* 47: 263–91.
Keynes, John M. 1936. *The General Theory of Employment, Interest, and Money*. New York: Harcourt, Brace and Company.
Leibenstein, Harvey. 1957a. *Economic Backwardness and Economic Growth*. New York: J. Wiley and Sons.
———. 1957b. "The Theory of Underemployment in Backward Economies." *Journal of Political Economy* 65: 91–103.
———. 1966. "Allocative Efficiency vs. 'X-efficiency.'" *American Economic Review* 56: 392–415.
———. 1968. "Entrepreneurship and Development." *American Economic Review*, *Papers and Proceedings* 58: 72–78.
_____. 1973a. "Competition and X-efficiency: Reply." *Journal of Political Economy* 81: 765–77.
_____. 1973b. "Notes on X-efficiency and Technical Change." In E.B. Ayal, ed., *Micro-Aspects of Development*, 18–38. New York: Praeger.
———. 1978. "On the Basic Proposition of X-efficiency Theory." *American Economic Review, Papers and Proceedings* 68: 328–32.
———. 1979. "A Branch of Economics Is Missing: Micro-Micro Theory." *Journal of Economic Literature* 17: 477–502.
———. 1982. "The Prisoner's Dilemma in the Invisible Hand: An Analysis of Intrafirm Productivity." *American Economic Review* 72: 92–97.
———. 1983. "Property Rights and X-efficiency: Comment." *American Economic Review*. 73: 831–42.
———. 1984. "On the Economics of Conventions and Institutions: An Exploratory Essay." *Journal of Institutional and Theoretical Economics* 140: 74–89.
Marsh, James G. 1978. "Bounded Rationality, Ambiguity, and the Engineering of Choice." *Bell Journal of Economics* 9: 587–608.
Pritchett, Lant. 1997. "Divergence, Big Time." *Journal of Economic Perspectives* 11: 3–17.
Reder, Melvin W. 1982. "Chicago Economics: Permanence and Change." *Journal of Economic Literature* 20: 1–38.
Rothschild, Kurt W. 2002. "The Absence of Power in Contemporary Economic Theory." *Journal of Socio-Economics* 31: 433–42.
Shafir, Eldar, Peter Diamond, and Amos Tversky. 1997. "Money Illusion." *Quarterly Journal of Economics* 112: 341–74.
Shapiro, Carl, and Joseph E. Stiglitz. 1984. "Equilibrium Unemployment as a Worker Discipline Device." *American Economic Review* 74: 433–44.
Simon, Herbert A. 1959. "Theories of Decision Making in Economics and Behavioral Science." *American Economic Review* 49: 252–83.
———. 1978. "Rationality as a Process and as a Product of Thought." *American Economic Review* 70: 1–16.
———. 1987. "Behavioral Economics." In John Eatwell, Murray Millgate, and Peter Newman, eds., *The New Palgrave: A Dictionary of Economics*. London: Macmillan.
Smith, Vernon L. 2003. "Constructivist and Ecological Rationality in Economics." *American Economic Review* 93: 465–508.
———. 2005. "Behavioral Economics Research and the Foundations of Economics." *Journal of Socio-Economics* 34: 135–50.
Stiglitz, Joseph E. 1987. "The Causes and Consequences of the Dependence of Quantity on Price." *Journal of Economic Literature* 25: 1–48.
Thaler, Richard H. 1992. *The Winner's Curse: Paradoxes and Anomalies of Economic Life*. New York: Free Press.
Todd, Peter M., and Gerd Gigerenzer. 2003. "Bounding Rationality to the World." *Journal of Economic Psychology* 24: 143–65.
Tomer, John F. 2001. "Understanding High-Performance Work Systems: The Joint Contribution of Economics and Human Resource Management." *Journal of Socio-Economics* 30: 63–73.
Tversky, Amos, and Daniel Kahneman. 1981. "The Framing of Decisions and the Psychology of Choice." *Science* 211: 453–58.