GEM Generalized Exchange Macroeconomics

Chapter Two

Baseline Workplace Equilibrium

James E. Annable

CHAPTER TWO

BASELINE WORKPLACE EQUILIBRIUM

The next four chapters construct the intertemporal general workplace-equilibrium theory. In the Workplace-Marketplace Synthesis (WMS), the subject of Chapter 6, optimizing workplace exchange is combined with textbook market exchange, uniquely microfounding stabilization-relevant monetary theory.[1] The analysis begins with the baseline establishment model, which is constructed in four steps. First, wage-optimization conditions are derived for profit-seeking firms operating in the complex environment of costly, asymmetric workplace information. Next, a broad range of available evidence motivates axiomatic preferences that, in turn, inform employee utility. The third step introduces axiomatic characteristics of production that are associated with input specialization, establishment size, and the nature of jobs, all ultimately rooted in the Second Industrial Revolution. Fourth, the analysis derives the critical unbundled wage condition from optimizing workplace exchange, relocating a substantial share of rational labor pricing from the marketplace to the workplace. The circumstances that mandate the payment of downward rigid wages and chronic labor rents are identified.

I. OPTIMIZING EMPLOYERS

Assumptions. The introductory workplace-equilibrium theory posits labor to be homogeneous until hired, at which point the homogeneity becomes firm-specific; workers cannot borrow or save; and employees and employers are risk-neutral. Firms maximize expected profits; workers maximize expected utility; and both agent classes form expectations rationally. The labor market has no barriers to competition; there are no unions; and the economy is closed. Equilibrium is understood in the modern sense of a rest period in the space of optimizing decision rules.

For now, technology is fixed, with axiomatic heterogeneity introduced via a bimodal division. One class of homogeneous establishments is characterized by large scale and input specificities

---

[1] Workplace-Marketplace Synthesis, two-venue general equilibrium (TVGE), and generalized-exchange macroeconomics are used interchangeably throughout the book.

that generate costly, asymmetric workplace information and routinized jobs. The other homogeneous class demonstrates small scale, the absence of meaningful specificities, and cost-effective worker monitoring.[2] The most salient features of baseline workplace equilibrium are the definition of on-the-job behavior, the bimodal technological heterogeneity, and the specifications of employee-utility, workplace-exchange, and income-distribution functions.

On-the-job behavior. Employee behavior at work is the intuitive starting point in workplace-equilibrium theory. Define, from the perspective of establishment j, a measure of labor input (E_j) for which the production function is well defined. Variable E_j is linked to labor hours paid for (H_j) by a scalar $\acute{Z}_j$, which in TVGE modeling captures worker modal on-the-job behavior (OJB):

$$(2.1) \qquad E_j = E^Q_j + E^G_j + E^S_j(t) = \acute{Z}_j H_j, \text{ such that } \acute{Z}_j \geq 0,$$

where E^Q_j measures the contribution from unenhanced employee hours, E^G_j denotes the contribution from general human capital, and E^S_j the contribution from firm-specific human capital.[3]

Formally, there exists, given technology, a subset $\mathbf{T_j}$ of the Cartesian product of the sets to which E_j and X_j (production) belong such that every element E_j and every element X_j appear once and only once in a pair in $\mathbf{T_j}$. The 1-1 mapping of $\mathbf{E_j}$ onto $\mathbf{X_j}$ implies that the sets are equivalent and that their relation is reflexive, symmetric, and transitive. Endogenous $\acute{Z}_j$ requires the non-

---

[2] Each of the two sets of production functions is defined for a given state of technical knowledge. For LEV homogenous production ($X_j = f(E_j, K_j)$), there exists maximum real output (X^P_j) described by a capacity function ($X^P_j = f(K_j)$), such that X^P_j is increasing in capital stock K_j, providing an upper bound on production ($X_j \leq X^P_j$). Capital services (K_j) flow from the capital stock ($K^P_j = f(K_j)$, $K_j \leq K^P_j$) and are (for analytical convenience in the compact version of the TVGE model) posited to be used in fixed proportion to labor services (E_j). For the highly simplified SEV production function, homogenous output (X_k) is posited to be increasing in its single input, labor hours (H_k). More general restrictions on the overall technology space include prices are excluded from the domain of inputs, no resources imply no production, $0 \in X$ (the establishment can be idle), X_j and X_k are closed sets.

[3] Baseline E^Q_j is definitionally embodied in worker hours offered in the marketplace, helping to determine the market wage. E^G_j, representing human-capital investment that increases labor productivity in multiple firms, is also embodied in market labor hours and also influences the market labor pricing. It has become the centerpiece explanation in the economic analysis of earnings differentials. (E^G_j is more easily measured and is ubiquitously shown in empirical work to significantly influence the establishment-size wage differential. In TVGE modeling, that familiar finding is understood to additionally result from the rational practice of rent-paying firms to "cream" their excess supply of job applicants, reversing causality from the SVGE analysis.) E^S_j, denoting human capital that enhances productivity exclusively in the employing firm, is embodied in labor hours only in the jth establishment and is independent of market opportunity cost.

equivalence of $\mathbf{X_j}$ and $\mathbf{H_j}$; rejecting $\mathbf{X_j}\sim\mathbf{H_j}$ is consequential and will be shown to be ultimately destructive for most textbook descriptions of labor pricing.[4]

At one extreme, characterized by perfect workplace information, the firm wholly controls and must optimize $\acute{Z}_j$, confronting a pre-Hawthorne problem set largely involving rest pauses and other working conditions.[5] Things get more interesting once scale and specialization compromise the symmetry and cost of workplace information. Agency problems shift a significant degree of $\acute{Z}_j$ control from the employer to employees. Worker optimization on the job becomes more complex, subject to a problem set that is decidedly post-Hawthorne. Variable cooperative worker conduct results from the less restricted pursuit of employee self-interest, inviting management methods of indirect control, including wage incentives, into the optimization process.

Workplace exchange. Employers and employees exchange workplace inputs ($\mathring{I}_j$) for outcomes ($\mathring{O}_j$), respectively defined as:

$$(2.2) \qquad \mathring{I}_j = \mathring{I}^E_j + \mathring{I}^f_j \text{ and } \mathring{O}_j = \mathring{O}^W_j + \mathring{O}^N_j \text{ such that } \mathring{I}^E_j = \acute{E}_j \text{ and } \mathring{O}^W_j = W_j,$$

where $\mathring{I}^E$ stands for cooperative labor input on the job, $\mathring{I}^f$ is job-related fixed employee characteristics, $\mathring{O}^W$ denotes the pecuniary outcome from the job, and $\mathring{O}^N$ represents nonpecuniary outcome. Intra-firm analysis focuses on the most economically interesting part of the exchange, i.e., the price-mediated relationship between employee behavior ($\acute{Z}_j = \acute{E}_j/H_j$) and the hourly nominal wage (W_j) that is named the *workplace-exchange relation* (WER). It is viewed from the aggregating perspective of management but motivated by employee optimization on the job.[6]

---

[4] The 1-1 mapping of labor hours onto output, via a well-behaved production function, critically supports one of the central management decision rules in the SVGE model class: $\Delta\Pi_j/\Delta H_j=0$. From Heijdra and van der Ploeg (2002, p.3): "The decision rule is a vitally important element in the macroeconomic labour market story. It is also relatively uncontroversial: virtually all macroeconomists believe in some version of [it]. We can easily transform [it] into the demand for labour, a schedule which shows how much labour a firm wants to hire for a given real wage." This book's success will be increasing in the degree to which the commonplace decision rule is made controversial. Replacing it with $\Delta\Pi_j/\Delta(\acute{Z}_jH_j)=0$ is shown to be a necessary condition of continuous-equilibrium, policy-relevant models capable of integrating stationary and nonstationary behavior of employment, income, and productivity.

[5] Leibenstein (1963) modeled the limiting case where firms, operating in countries with low living standards, influence $\acute{Z}_j$ by directly managing the nutrition and health of their employees. For elaboration, see Annable (1984).

[6] The arrangement is best thought of as a self-enforcing implicit contract, a concept well established in economic theory. (See, for example, Okun (1981).) From Cahuc and Zylberberg (2004): "The features of labor contracts depend, to a large extent, on whether or not the results of an employee's activity can be observed and taken into

Employees' optimization of their cooperative input with respect to the wage received necessarily plays a central role in any price-focused model of workplace conduct and, ultimately, workplace exchange. Establishment-specific WERs are defined over the range of feasible labor pricing, restricting $\acute{Z}_j$ to be positive and non-decreasing in compensation paid.

In the complex work environment that results from combining employee consciousness with costly, asymmetric information and the absence of 1-1 technical correspondence between $\mathbf{X}_j$ and $\mathbf{H}_j$, the firm must identify its WER sufficiently to enable payment of the unit-cost minimizing labor price (the *efficiency wage*):[7]

$$(2.3) \qquad W^n_j = \max_W (\acute{Z}_j/W_j).$$

Given that firm's choice set permits zero production, the rational payment of the efficiency wage additionally requires that the establishment's expected nonstationary revenues be greater than or equal to its variable costs. The revenue constraint plays an important role later in the analysis.

The firm's labor-pricing decision rule critically accommodates wage rents. Indeed, absent the derivation of optimal rents, workplace modeling is uninteresting to macroeconomists. If labor compensation and use default to market solutions, theorists can properly ignore workplace exchange. A necessary condition for the rational payment of a wage greater than employees' opportunity cost is lower unit labor costs, implying an optimal wage incentive ($W_j > W^m$) exists iff:

$$(2.4) \qquad \acute{Z}_j/W_j > \acute{Z}^m_j/W^m,$$

where $\acute{Z}^m$ is the labor productivity associated with cost-effective workplace monitoring (assumed to be given) and W^m denotes nominal market opportunity cost, i.e., the discounted expected best

---

account. These results can only appear explicitly in the contract if they are verifiable. If they are not, the work relationship is governed by implicit and self-reinforcing clauses…. In reality, the labor contract most often takes the form of a relationship of subordination that simply acknowledges the employer's authority and sets out a specified amount of remuneration. It is not generally possible to know in advance what services will be supplied in return for the wage." (pp.307-8)

[7] Given costly, asymmetric information, LEV management cannot know individual $\acute{Z}_{ij}$ but know enough about group performance to identify modal $\acute{Z}_j$.

alternative wage upon separation from firm j. The WER class that is consistent with the rational payment of wage incentives is *unbundled* (i.e., $\acute{Z}_j/\acute{Z}^m{}_j > W_j/W^m$), while WERs mandating payment of market-opportunity costs are *bundled* (i.e., $W_j/W^m > \acute{Z}_j/\acute{Z}^m{}_j$).[8]

The analytic framework accommodates the familiar Solow (1979) condition, given a differentiable WER (in the range $W_j \geq W^m$) that satisfies: $d\acute{Z}_j/dW_j = \acute{Z}_j/W_j$ and $d^2\acute{Z}_j/dW_j^2 < 0$. The intriguing possibility of the (local profit-maximizing) Solow condition has always been that opportunity costs (W^m) play a nonbinding role in rational wage determination.[9]

Some preliminary implications. Bundled workplace exchange, designated by k, implies the textbook labor-input story. The firm's optimal labor price is always the market rate, determined via the interaction of households' supply of labor services (measured in hours and reflecting preferences with respect to consumption, leisure, and work conditions) and firms' demand for labor hours (reflecting profit-maximization and labor productivity ground out by establishment capital and technology). SVGE market-exchange mechanisms resolve excess demands, producing labor pricing simultaneously equal to the marginal-product value of labor and the marginal disutility of work. Rational employers and employees are market-wage takers. Effective worker oversight implies constant $\acute{Z}$, which in turn implies that optimizing workplaces generate unit labor costs that are strictly greater than $W^m/\acute{Z}^m{}_k$ for $W_k > W^m$.[10] The labor market, of course, cannot directly measure $\acute{Z}_k$. But, with bundled OJB, the limitation does not matter. The profit-seeking firm effectively gauges its cooperative labor input ($\acute{Z}^m{}_k E_k(t)$) by employee hours, which can be measured and priced in the marketplace.

---

[8] A third class of WER bundling, $(\Delta\acute{Z}/\acute{Z})/(\Delta W/W) = 1$, microfounds perfect piece-rate compensation systems. For more on piece-rates, see Chapter 8.

[9] A more complete set of optimizing conditions is $d\acute{Z}_j/dW_j = \acute{Z}_j/W_j > \acute{Z}^m{}_j/W^m$ and $d^2\acute{Z}_j/dW_j^2 < 0$. Solow (1979) and Annable (1977, 1980), both assuming differentiable WERs, independently derived the firm's conditions for local profit maximization subject to discretionary $\acute{Z}_j$. Given self-interested workers, the firm's costs are represented as $C_j = W_j(E_j/\acute{Z}_j) + F_j$, where C is total costs and F is fixed costs. Given output, local cost-minimization with respect to the wage paid mandates unitary elasticity of effective labor input, an important original EWT conclusion: $dC_j/dW_j = (E_j/\acute{Z}_j)(1 - (d\acute{Z}_j/dW_j)(W_j/\acute{Z}_j)) = 0$; $d\acute{Z}_j/dW_j = \acute{Z}_j/W_j$; and $d^2\acute{Z}_j/dW_j^2 < 0$. Modern workplace-equilibrium analysis usefully drops the restriction that WERs must be differentiable.

[10] If the production process requires specific human capital, the firm's optimal labor pricing includes adjusting the market wage with a minimally small, constant wage premium that discourages voluntary quits. See Chapter 8.

However, if workplace exchange is unbundled, the familiar story breaks down. That breakdown attracted the attention of mid-century, on-site labor economists. They observed that, in some establishments, profits were sufficiently influenced by endogenous worker behavior to eliminate the market wage as a rational solution to the firm's labor-pricing problem. Unit labor costs were lower than $W^m/\acute{Z}^m_j$ for a range of feasible nonmarket wages ($W_j > W^m$).

Profit-seeking management becomes more complex and more recognizable. In response to the market's inability to measure unbundled $\acute{Z}_j$, labor-pricing is moved inside the firm, where W_j and $\acute{Z}_j$ are simultaneously optimized. Management rationally constructs and administers firm-specific mechanisms of workplace exchange. Those administrative mechanisms, shaped by information asymmetries, are designed to deal with the $\acute{Z}_j$ measurement problem indirectly, by playing the averages and managing to what the firm has learned about modal employee preferences. The ubiquitous goal, identified in the early Hawthorne experiments, is to encourage voluntary employee cooperation with respect to management objectives.

II. OPTIMIZING EMPLOYEES

Workplace-equilibrium theory most improves on consensus SVGE treatment of OJB by carefully motivating worker preferences, drawing on the substantial large-establishment, management, and experimental literatures. Simply positing either a 1-1 correspondence between production and hours at work or a dominating desire to loaf on the job turns out to be a badly inadequate, misleading foundation for understanding employee behavior (and pricing) in a specialized economy.

Practitioners have long understood that, to be useful, descriptions of employee behavior must include both sides of the fundamental workplace exchange: the worker-determined input to the job and the firm-determined outcome from the job. For worker i, the wage received is part of a broader array of outcomes from the job:

$$(2.5) \qquad \mathring{O}_{ij} = f(W_{ij}, \mathring{O}^N_{ij}),$$

where W denotes monetary payment per hour at work and $\mathring{O}^N$ represents nonpecuniary rewards.[11] The latter results from factors characteristic of a given job, such as the interest and challenge of the work itself, that are tend to be stable over time. Similarly, OJB is perceived as part of an array of workplace inputs:

$$(2.6) \qquad \mathring{I}_{ij} = f(\dot{Z}_{ij}, \mathring{I}^f_{ij}),$$

where $\dot{Z}_{ij}$ has been defined as cooperative labor services per hour, which can be altered quickly, and $\mathring{I}^f_{ij}$ denotes employee characteristics, such as seniority, skill, and education, that are posited to be non-decreasing in time. Managers know that cooperation is especially important, being both critical to the efficiency of specialized production and problematic to monitor.[12] The arguments in each function are assumed, for analytic convenience, to be additive.

Preferences

An employee has preferences about the workplace exchange he or she is making. Such preferences are an important feature of specialized economies, having long been a focus of academic and practitioner interest, research and experimentation. Drawing on that literature, posit that the worker prefers his or her ratio of outcomes (from the job) to inputs (to the job) to be greater than or equal to each of the following reference standards:

(a) The ratio of outcomes to inputs for his or her best alternative employment,

$$(2.7a) \qquad \mathring{O}_{ij} / \mathring{I}_{ij} \geq \mathring{O}^a / \mathring{I}^a;$$

(b) The ratio of outcomes to inputs for his or her interpersonal reference standard, i.e., other workers who perform the same or similar tasks as the individual or who work in close proximity to him or her,

$$(2.7b) \qquad \mathring{O}_{ij} / \mathring{I}_{ij} \geq \mathring{O}^b / \mathring{I}^b; \text{ and}$$

---

[11] See March and Simon (1958). Annable (1977) is an early example of the use the ratio of job outcomes to inputs to model worker behavior in the context of optimizing economic theory. For insightful applications outside economics, see Homans (1961), Adams (1963), and Blau (1964).

[12] John Stuart Mill generalized Adam Smith's division of labor to the "more fundamental" principle of worker cooperation.

(c) The ratio of outcomes to inputs for his or her intertemporal reference standard, i.e., the trend improvement of his or her outcomes-inputs ratio over time,

$$(2.7c) \qquad \mathring{O}_{ij}/\, \mathring{I}_{ij} \geq (1 + \dot{g})\, \mathring{O}_{io}/\, \mathring{I}_{io} = \mathring{O}^c/\, \mathring{I}^c.$$

In the notation, *a* represents the worker's best alternative job, *b* denotes the interpersonal reference standard, *c* is the intertemporal preference standard, *o* stands for the relevant earlier time period, and $\dot{g}$ is the trend growth rate in job outcomes that has already occurred.[13]

The three types of reference standards calibrate the worker's inherent preference for fair treatment. Formally, there exists a set of pairings of workplace outcomes and inputs $\mathbf{K_{ij}}$= $\{\mathring{O}^a/\mathring{I}^a, \mathring{O}^b/\mathring{I}^b, \mathring{O}^c/\mathring{I}^c\}$, demonstrating completeness and transitivity, for which preference relation 2.7 is satisfied by the set's least upper bound: sup $\mathbf{K_{ij}}$. The preference for equity (and the desire for redress of unfair treatment) is today understood to be axiomatic, an outcome of evolutionary biology that was embedded in neural networks as our distant ancestors adapted to survival advantages available from group cooperation. From a prominent neuroscientist: "Our instincts for sensing and responding to fair exchange evolved in a social environment where tit for tat was king. What you did to me today was coming back to you tomorrow in kind."[14]

---

[13] Scholars have been using interpersonal and intertemporal comparisons in their attempts to make economic descriptions of behavior more congruent with psychological evidence. Brenner (1983) and Frank (1985) construct utility functions that include peer-group wealth. Kahneman and Tversky (1979) include the individual's earlier wealth as the critical reference point. (See also Harford (2003).) For other notable efforts to extend utility functions to include reference standards, see Duesenberry (1949), Becker (1996), de la Croix (2001), and the important behaviorist school of economic theorists (for example, Camerer, Loewenstein, and Rabin (2004)). The original economic work on reference wages and the importance of perceived fair treatment, as has been noted, came from the workplace analysis of Kerr *et al.* (Kerr (1988, 1994).) They found that employees preferred wages that were consistent with interpersonal and intertemporal reference standards that had become ingrained (via repeated application) over time and that had been incorporated into workplace standards of acceptable treatment.

[14] Montague (2006, p.186). Modern brain research and behavioral experiments are producing critical information on proper axioms for use in economic modeling. Especially relevant to formal workplace modeling are the scientific and experimental findings suggesting that contemporary human choices are made with neural mechanisms in which natural selection has embedded a strong preference for fair treatment. Those findings are remarkably consistent across cultures, providing guidance that serious economic theorists cannot ignore. Behavioral economists are exploring, to good effect, the biochemistry of decision-making and its implications for economic postulates (for example, Camerer, Lowenstein, and Prelec (2005)). Their work is notably consistent with Gary Becker's (1996) belief that the roots of stable preferences are found in biology, not society and informed by Joel Mokyr's (2006) essay on the progress in, and difficulty of, importing concepts of evolutionary science into mainstream neoclassical theory. The workplace model's employee preferences were anticipated, as was so much else, by Adam Smith. In his *The Theory of Moral Sentiments*, important behavior motivators are the interrelated factors of status, respect, and justice. An overriding preference is fair treatment: "… we find ourselves to be under a stricter obligation to act according to justice than agreeably to friendship, charity or generosity; that the practice of these last-mentioned virtues seems to be left to some measure to our own choice, but that, somehow or other, we feel ourselves to be in a

There is little disagreement among practitioners that wage comparisons and concepts of equity and fairness pervade the labor-pricing process. For example, Milkovich and Newman, authors of a well-known business text on wage policymaking, assert in their first chapter: "Equity forms the building block, the foundation on which pay systems are designed."[15] In this introductory version of workplace-equilibrium theory, it is convenient to assume that the worker obtains information on reference standards costlessly. Loosening that assumption would strengthen the continuous-equilibrium workplace model's conclusions but at the cost of a less tractable analysis.[16]

The reference wage. The reference wage ($W^{\acute{n}}_{ij}$) is the monetary payment that is consistent with the least upper bound of $\mathbb{K}_{ij}$. When wages are the only endogenous outcome and inputs are unchanged, $\mathbb{K}_{ij}=\{W^a, W^b, W^c\}$; and the worker's preference for fair treatment is satisfied by $W^{\acute{n}}_{ij}=\sup \mathbb{K}_{ij}$. The classes of reference standards (*a, b, c*) that calibrate the employee's preference for fair treatment are named *specific workplace capital*. Investment in that category of human capital occurs during on-the-job interactions (often involving informal training) as veteran workers inculcate new hires with a particular calibration for their inherent preference for equitable treatment.[17] Given firm-specific employee homogeneity, $W^{\acute{n}}_j=\sup \mathbb{K}_j$.

---

peculiar manner tied, bound, and obliged to the observation of justice." It is perhaps additionally interesting that the importance of relativity in human satisfaction aligns nicely with its importance in physical sciences, traceable back to early insights of Galileo in considering heliocentric cosmology and much more fully exploited by Einstein.

[15] Milkovich and Newman (1984), p. 8.

[16] Lagged learning mechanisms (see Evans and Honkapohja (1999), Evans and Honkapohja (2001), and Honkapohja (2001)) could be easily, and usefully, incorporated into the workplace-equilibrium theory.

[17] It is not surprising that sociologists have thought harder about relational phenomena, like those producing specific workplace capital, than have economists. From Swedberg (2003): "Interactions between people gradually acquire an objective quality, and eventually people take them for granted. What is specific about economic institutions … is that they involve 'the mobilization of resources for collective action'." (p. 36) In workplace-equilibrium analysis, it is the dynamic variability of the reference standards – not their specific content – that is important. Much has been written about the interpersonal reference standard, indicating that the selection of comparisons is determined partly by historical happenstance. (Interested readers are referred to the large practitioner literature on the determination of standards of wage comparisons, a sampling of which is provided by footnote 19 in this chapter.) Given that the content issues of preference relation 2.7 are concentrated in the interpersonal standard, it is notable that workplace optimization could be motivated wholly with the intertemporal standard with remarkably little damage to the results.

Baseline workplace equilibrium, the focus of this chapter, can now be defined. It is the stage of two-venue general-equilibrium modeling during which $\mathbb{K}_j$ is unchanged and consistent with the payment of wage rents ($W^n_j \geq W^a$). $\mathbb{K}_j$ is permitted to rationally vary in the next chapter, when the TVGE analysis is extended beyond the stationary movement of employment and output.

Reference Standards

When wages are the only variable job outcome, the three classes of reference standards can be defined in terms of labor compensation. For now, illustrative descriptions of each will be sufficient. The optimal intertemporal calibration of workplace reference standards will be modeled in Chapter 3.

Best alternative employment. The first reference standard ($W^a = W^m$) reflects market opportunity cost, i.e., the discounted expected best alternative wage upon separation from firm *j*. This reference standard incorporates competitive labor-market behavior into workplace theory. Self-interested workers will accept no less than their market wage.

Surveys are widely used by large firms to gather information on the best-alternative reference standard. In practice, however, managements always exercise latitude in translating survey information into wage policies. Substantial wage dispersion for similar workers in the same labor market has long been characteristic of specialized economies. In an early study, Richard Lester (1946), analyzing data from U.S. wartime surveys of some sixty metropolitan areas, found compensation differentials averaging 50 percent for the same occupations in the same labor markets. Summarizing a later 1948 study, Lester (1988, p.93) concluded: "Company executives gave such reasons or objectives as the following for pursuing a high-wage policy and for trying to be early in a round of wage increases: to avoid being unionized, to preserve employee attachment to the company, to maintain good employee morale, which can facilitate supervision and increase productivity, and to assure the desired quality of labor."[18]

---

[18] Mortensen (2003) reviewed existing evidence on establishment-size wage differentials, concluding that they are substantial. Modern empirical studies have reduced Lester's estimates by accounting for human-capital differences. A problem with the more-detailed studies, however, is their assumption of a one-way causation from the human-capital differences to the firm's wage. It will be demonstrated below that, given specialized, large-scale production,

The evidence strongly indicates that, despite receiving almost exclusive attention from modern economic theorists, the labor market is only part of the worker's story. Interpersonal and intertemporal reference standards introduce (nonmarket) positional concerns into employees' perceptions of satisfaction on the job. Recognizing such relativity is a crucial step, making economic theory more consistent with what is known about employee preferences and greatly enhancing its explanatory and predictive power.

As indicated above, interpersonal and intertemporal reference standards are passed from veteran to new employees during workplace interactions, often involving on-the-job training. From Doeringer and Piore (1971, p. 31): "In any workplace, learning and teaching occur automatically, and often at little cost." Such on-the-job contact enables new workers to acquire firm-specific human capital and necessarily plays a significant role in workplace economics. The spontaneous investment process is analyzed as a dynamic optimization problem in the next chapter.

Interpersonal standards. The interpersonal reference standard (W^b) becomes established over time and is a familiar concept to even casual observers of labor behavior. Wage comparisons and concepts of equity pervade the wage-setting process and provide a critical framework for practitioner action. Arbitrators, for example, make frequent use of established standards of comparison in wage cases. Indeed, when such standards do not exist, arbitrators typically seek to create them. The literature on the establishment and maintenance of interpersonal reference standards in wage determination is large and longstanding.[19]

---

assuming one-way causation is misleading, resulting in downward-biased estimates of the wage variation rooted in establishment size. Opposite-direction causation reflects employers responding to profit-seeking forcing them to pay market-wage premiums by creaming the excess supply of job applicants, i.e., hiring the best available. TVGE modeling restores credence to the early, larger estimates of cross-sectional market wage variability.

[19] The literature includes Patchen; Livernash; Merton and Rossi; Hills; Lipset and Trow; Gartrell; Goodman; Messe and Watts; Major and Testa; Dornstein; Ross and McMillen; Ambrose and Kukik; and Heneman, Schwab, Standal, and Peterson. For more on the nature and origins of the reference standards, see Annable (1984, 2007b). Note also that the workplace model does not require the interpersonal comparison of utilities. That sort of exercise, usually found in welfare economics, is extraordinarily challenging. (See Hammond (1991).) The model being constructed here requires a much easier, two-part reference process. First, workers use available information to compare (interpersonally, intertemporally, and in the marketplace) a measurable outcome, wages paid. Compensation administrators know that assessing one's wage relative to others, or one's own history, is a commonplace activity. Second, the results of those comparisons affect worker satisfaction, another outcome well known to practitioners.

From the variety of factors generating interpersonal wage comparisons, analysts have identified three as most significant:

- Workplace proximity. Employees typically have substantial knowledge about the compensation paid to their colleagues working in the same unit, creating discontent when existing relationships are violated. The basic idea has been around a long time, largely outside economics. In 1957, in his proximity thesis, the sociologist R. K. Merton argued that individuals compare themselves with others with whom they are in "actual association, in sustained social relations." John Dunlop (1988, p.53), who among economists arguably knew the most about wage determination and labor hierarchies within firms, concluded: "Every compensation unit with more than one job classification faces the issue of internal consistency, fairness, or congruence among different job classifications and individuals…."

- Similar work. Individuals seek and typically obtain knowledge about compensation paid to workers who do the same or similar tasks as the individual. Consequently, a wage reduction relative to this reference class causes dissatisfaction. This characteristic extends comparisons beyond the worker's immediate work unit to elsewhere in the organization.[20]

- Similar fixed inputs to the job. A related criterion for assessing fair treatment used by workers is the perceived similarity of fixed inputs ($\mathring{I}^f$) to the job. As defined above, such inputs include seniority, age, education, and work skills.

Almost all large and many middle-sized firms use job-evaluation plans to help manage internal wage structures that become established over time. These plans use data that mirror employees' interpersonal reference standards, rating jobs according to such factors as education required, skill, responsibility, and working conditions. The ratings assigned to each job are then related to

---

[20] Perceptions of similar work can also extend the reach of wage comparisons to other establishments. Such external comparisons impose a greater time requirements on the dissemination of relevant information. Given space considerations, external comparisons will not be a focus of this introductory analysis.

internal pay differences, preserving customary pay relationships and reducing the likelihood that the firm's wage practices will appear arbitrary and inequitable to employees.[21]

BOX 2.1: MORE ON INTERPERSONAL REFERENCE STANDARDS

Human consciousness is understood to be organized around reference standards. That is how we make sense of a complex world. Two additional characteristics of interpersonal reference standards complete the compact workplace application.

First is the *workplace adverse-selection* phenomenon. Established internal wage structures embodied in $\mathbb{K}_j$ evolve more slowly over time than do workforce skill distributions and market opportunity costs of those skill sets, especially when general (and specific) human capital is developed on the job, job content changes, and a broad range of more productive positions are filled via promotion. It follows that pay structures always embody substantial interpersonal variation in wage rents and that the lowest-rent (most cost-efficient) employees are most at risk of being permanently lost to other firms in the aftermath of an across-the-board wage cut. (For elaboration, see Chapter 3.)

The second is the *linchpin* phenomenon, which elaborates on a particular configuration of Dunlop's job clusters. A leader of a given work group within a large organization may be successful because of his or her market relationships, reputation, and proven capacity to identify opportunities, make connections, and solve problems. He or she may be a rainmaker, generating high-value work for his or her work group. Most generally, they simply make useful things happen. If his or her ability to organize and lead a valuable operation is rooted in general human capital, the linchpin employee may earn relatively little in excess of his or her opportunity costs. However, the other members of the synergistic work group, having established reference standards ultimately calibrated to their manager's compensation, are likely paid significant rents. Linchpin employees enjoy, and can exploit, substantial loyalty from members of their work group. In particular, the group tends to adopt the linchpin's objectives as their own.

---

[21] Dunlop (1957) used the term *wage contour* to denote the end result of external comparisons. By definition, a wage contour is a relatively stable group of wage-setting establishments that are linked via similarity of product markets, similar sources for labor force, or common labor-market organization (custom). By virtue of those linkages, the contours have wage-making characteristics to which employees become accustomed. Similarly, he defined *job clusters* as a stable group of job classifications or work assignments within an establishment that are linked by (a) technology, (b) the administrative organization of the production process, or (c) social custom. The relationships within a cluster develop around key jobs and their wage rates. Those key rates provide the reference standards to peg all job rates within the job cluster.

Intertemporal standards. The final reference class (W^c) also tends to become established over time. If work on a job has yielded a fairly regular growth in the real wage, the association develops legitimacy in the mind of the employee. A reduced rate of improvement is perceived as unsatisfactory.[22] Intertemporal standards help translate the static nominal efficiency-wage associated with the workplace-exchange relation (WER) into a dynamic real efficiency wage. This third set of reference standards also introduces history into the labor-pricing process. More formally, the reference standard introduces a unit root and path dependency into rational wage dynamics.

Firms recognize that worker preferences are burdened with baggage from the past, adjusting wage and other human-resources policies accordingly. In the beginning, of course, those practices may have been consistent with labor-market circumstances. Once having become a standard for acceptable behavior, however, they take on a life of their own and become independent of contemporaneous conditions in the marketplace.[23]

Employee Utility

It is convenient to frame employee preferences using a utility function. In what follows, the familiar von Neumann-Morgenstern discounted utility framework is adapted to the axiomatic employee preferences featured in TVGE modeling:

---

[22] In practice, consistency between the intertemporal and interpersonal reference standards typically results from management's recognition of the relative importance, to employees, of both. See Chapter 8.

[23] If the model accurately describes worker behavior, then custom and tradition should be familiar concepts to wage administrators. And they are. From a text on wage administration: "To the recipient, pay serves as a symbol of status as well as a means to purchasable ends…. Understanding compensation as status values helps explain the force of custom and tradition in pay matters. The protection of present status and the desire to improve status appear to be ubiquitous human values. Protection of present status gives force to custom defined as 'what is right'. Custom and tradition require justification for change. The force of custom in compensation is conservative. Changes must not be made unless justified and when made call forth numerous other changes based on traditional relationships." (Belcher, pp.6-7) Introducing custom and tradition into economic models goes back at least to John Stuart Mill. More recently, some prominent economic theorists have recognized the significance of norms in worker behavior. A significant example is the late Sir John Hicks (1963) who, rejecting the hard-line market focus of his earlier work, eventually emphasized fairness, rooted in comparisons, from the workers' point of view, and of continuity of relations from the employers' point of view in the setting of wages. (See Chapter 4.)

(2.8) $E_{oi}\sum((1+ŗ)^{-t}Ú_j(Ć_j(t),Ĺ_j(t),W_j(t)/W^{ń}_j(t)))$,

such that $(\Delta Ú_j/\Delta(W_j/W^{ń}_j) \mid W_j \leq W^{ń}_j) > 0$.[24]

$Ć$ represents consumption at time t; $Ĺ$ is leisure (moving inversely with hours at work, H); E_o denotes the expectation of future values of the function's arguments based on the cost-effective information available at the beginning of the current period ($t=0$); $ŗ$ is the subjective discount factor ($ŗ \in (0,1)$); Δ is the change operator; and the series are summed from $t=0$ to $t=\varkappa$, the employee's desired tenure at firm j.[25] The function is temporally separable; and, to assure its existence, preference relations motivating the arguments are reasonably posited to be complete, transitive, and continuous.

Instantaneous utility, the focus of the static analysis, is assumed to be positive and convex in consumption and leisure. The innovation here is the reference wage ($W^{ń}$). To reiterate, three classes of reference standards calibrate the additional preference for equity: W^a (his or her best alternative wage), W^b (the interpersonal reference standard), and W^c (the intertemporal reference standard). There exists a set of labor prices $\mathbb{K}_j=\{W^a,W^b,W^c\}$, for which the worker's preference for fair treatment is satisfied by the set's least upper bound: $W^{ń}_j=\sup \mathbb{K}_j$. Preference relation 2.7 implies that instantaneous utility is strictly increasing in $W_j/W^{ń}_j$ up to unity and unchanged thereafter. The axiomatic preferences used in workplace-equilibrium theory are broadly consistent with the extended-utility approach pioneered by Duesenberry (1949) and Modigliani (1949) and used more recently by a number of economists including Becker (1996) and de la Croix (2001).[26]

---

[24] Note that, for economy of presentation, the utility function holds nonpecuniary outcomes and hours-embodied inputs constant. Optimizing tradeoffs among wages, E, O^N, and I^f are interesting but play an insufficient role in the development and application of the economic model to justify complicating this introduction to coherent generalized exchange theory.

[25] An infinite horizon, greatly simplifying many applications in macro analysis, could be used. Rios-Rull (1994) produced cyclical properties similar to an infinite-horizon model by calibrating an overlapping-generations model to the U.S. population's age structure. Other restrictions on instantaneous utility are $(\Delta Ú_j/\Delta(W_j/W^{ń}_j) \mid W_j > W^{ń}_j)=0$, $\Delta Ĺ_j/\Delta H_j < 0$, and $Ć_j = H_j W_j$.

[26] For a review of the use of extended preferences in macroeconomic modeling, see de la Croix (2003). In general, care must be taken in the search for topological transformations of preferences that make one's model more consistent with the data. That practice has too often amounted to little more than the hidden use of free parameters. Such transformations are admissible in model building only if the reformulated preferences are both intuitively plausible and strongly supported by the evidence, i.e., they must be axiomatic. The real-world importance that

Workers still prefer higher to lower wages, but they are now dissatisfied (experiencing a reduction in utility) when their compensation falls below the established reference standard. Moreover, given consumption and leisure, wages in excess of the reference standard produce nil gain in satisfaction.[27] As has been emphasized, positing reference-dependent utility is consistent with the findings of economists who have been investigating the psychological foundations of economic behavior. In reviewing that literature, Rabin concluded: "Overwhelming evidence shows that humans are often more sensitive to how their current situation differs from some reference level than to the absolute characteristics of the situation."[28]

More fundamentally, as was also noted above, evidence is accumulating that, over the course of human evolution, an elemental concern about relative position has been hard-wired in the human brain. The naturally-selected adaption facilitates cooperation and improved survival probabilities that accompany the capacity to specialize by working together. Arthur Robson (2001), for example, came to the hard-wired conclusion in his review of the literature on the biological basis of economic behavior; and researchers in the emergent field of neuroeconomics have provided consistent support for the idea. (See also Frank (1999, 2005), Fehr *et al.* (2005), Zak *et al.* (2005), and the extensive work on "ultimatum" and "social dilemma" games in experimental economics.) At some level, we all know this. H. L. Mencken once defined a rich man as someone who earns a hundred dollars a year more than his wife's sister's husband.

---

employees assign to fair treatment by management easily satisfies those requirements. No other worker-preference set widely used in economic models today (e.g., the dominant urge to shirk) can make that claim.

[27] The greater consumption and leisure that result from higher wages do, of course, increase satisfaction.

[28] Rabin (1998), p.13. Rabin's survey also emphasized two characteristics that are compatible with the optimizing workplace theory. First, loss aversion is ubiquitously important. "Researchers have identified a pervasive feature of reference dependence: In a wide variety of domains, people are significantly more averse to losses than they are attracted to same-sized gains." (p. 14) Experimental evidence shows that most people experience an asymmetrically strong dislike for even modest financial setbacks. The static workplace model, in which relative wage losses push the employee out of equilibrium while same-sized relative gains do not, is consistent with that psychological characteristic. Second, marginal utility diminishes most sharply in the immediate neighborhood of established reference standards. "The marginal effects in perceived well-being are greater for changes close to one's reference level than for changes further away… diminishing sensitivity is a pervasive pattern of human perception." (p.15)

The employee preference for fair treatment by management is stable. He or she always wants to be treated equitably. But, in application, that preference must be calibrated by the choice of reference standards, which are given in the baseline analysis but made endogenous in Chapter 3. Endogenous calibration of the reference wage ($W^{\acute{n}}_{ij}(t)$) is motivated by intertemporal substitution of perceived fair treatment for consumption, the analysis of which will complete the microeconomic modeling of optimizing behavior on the job.

Momentary utility and labor rents. The existence and circumstances of $W_j=W^{\acute{n}}_j>W^m$ are derived later in this chapter. For now, simply note that the payment of wage rents usefully constrains worker optimization. Preferred jobs are rationed, and employment by a rent-paying establishment implies being pushed off one's market labor-supply schedule, suppressing work-leisure choice.[29] From the perspective of employees, hours on such jobs are exogenously determined, limiting the active pursuit of instantaneous utility maximization to adjusting on-the-job behavior:

$$(2.9) \qquad \max_{\acute{Z}} \acute{U}_j(C_j,L^o_j,W_j/W^{\acute{n}}_j).$$

In baseline analysis, the reference wage is also given, implying fixed consumption (C^o) unless the firm attempts to reduce rents by cutting compensation from established $W^{\acute{n}}_j$.[30] Employee dissatisfaction with the inequitable change would then be reinforced by his or her preference for more to less consumption. Workers maximize momentary utility on the job at $W_j/W^{\acute{n}}_j=1$.

The employee, occupying an inherently subordinate position in the establishment's hierarchy of authority, prefers fair treatment, defines that preference along three dimensions, and formulates optimizing workplace decision rules accordingly. His or her equilibrium is understood as a rest period in the space of those decision rules. If preference relation 2.7 is violated, the now dissatisfied employee is in disequilibrium with respect to on-the-job behavior. Moreover, if the nature of the workplace does not permit the establishment of interpersonal and intertemporal

---

[29] Gordon (1990, p.1138) makes a similar point but provides no derivation.

[30] It will be demonstrated in the next section that workers cannot spontaneously organize workplace behavior that pushes the firm's unit-cost minimizing labor price above the established reference wage. Absent help, effective employee action is shown to be confined to defending $W^{\acute{n}}_j$.

standards, the operative reference system collapses to the best-alternative-job comparison, making the reference wage equivalent to the worker's market opportunity cost (sup $\mathcal{K}=$ {W^a}=W^m). Employees, no longer earning wage rents from endogenous OJB, return to their market labor-supply schedule and their optimization of work-leisure choice.

In the remainder of this chapter, baseline analysis critically derives how employee satisfaction with his or her relative position (relation 2.7) is related to profit-seeking employer's rational efficiency wage (condition 2.2). Before that, however, the roles of establishment size and the routinized nature of production and nonsupervisory jobs are examined in useful detail.

III. TECHNOLOGICAL CONSTRAINTS

This section elaborates on two central technological constraints of the workplace-equilibrium model. The first bifurcates production with respect to establishment scale; the second bifurcates jobs, in part to capture their differing capacity to generate nonpecuniary satisfaction. Both separations are rooted in the Second Industrial Revolution and the specificities that characterize the hugely altered production landscape. Caballero and Hammour (2003, p.173) have carefully analyzed the effects of input specificity on market exchange, providing a useful definition: "… a factor is specific with respect to a production arrangement – its current production relationship with other factors using a given technology – when it would lose part of its value [productivity] if used outside this arrangement."[31]

Labor specificity is associated with economic rents, agency problems, and differing capacities of jobs to generate nonpecuniary satisfaction, considerably complicating optimizing workplace behavior. In particular, it implies interrelated limitations, also associated with establishment size, on the firm's capacity to monitor worker behavior and the market's capacity to price labor, introducing critical structural heterogeneities into profit-seeking wage policymaking.[32]

---

[31] Caballero and Hammour (1998 and 2001) powerfully organize their analyses of macroeconomic dynamics, integrating markets and relevant institutions, around input specificities.

[32] Vertical integration, which substitutes managerial for market coordination, has been extensively studied since the pioneering efforts of Coase (1937). Economists have identified a variety of interrelated reasons for rational firms to make rather than buy an input, including coordination costs (Coase), synergies resulting from team production

Heterogeneous Work Establishments

In workplace modeling, firms become large to exploit the productivity gains inherent in factor specialization, defined to include work-task transformations that occur with the development and progressive deepening of capital-intensive production. Increasing establishment size and specialization has two companion effects that are significant in the analysis of workplace behavior: (a) economic rents generated by firm-specific physical and human capital, and (b) rising cost and asymmetric nature of management information on employee activities, rooted in team production as well as the deteriorating ratio of signal to noise as intra-firm information chains grow longer.[33]

In large establishments, workers typically know how to do their jobs and how they are doing them better than does management. Employers must manage and price labor resources, and employees must pursue their own best interests, in the context of that fundamental asymmetry, which is further complicated by mutually limited trust rooted in conflicting employer-employee objectives. The labor market prices worker hours (H), not labor cooperative input (E). Given sufficient $\acute{Z}_j$ endogeneity, rational labor pricing necessarily migrates from the marketplace to the workplace. Moreover, the relatively substantial investment in specific human capital characteristic of large establishments generates significant costs of job switching, further reducing the capacity of markets to price labor inputs in line with their productivities and encouraging long-tenured employment. In small work establishments, labor-input specificities, job-mobility costs, and monitoring problems are insufficient to support meaningfully endogenous $\acute{Z}$.

(Alchian and Demsetz (1972)), and asset specificity (Williamson (1975)). Chandler (1996, pp.80-81) provides a useful summary analysis constructed on two propositions: "The first proposition is that modern multiunit business enterprise replaced small traditional enterprise when administrative coordination permitted greater productivity, lower cost, and higher profits than coordination by market mechanisms.... The second proposition is simply that the advantages of internalizing the activities of many business units within a single enterprise could not be realized until a managerial hierarchy had been created."

[33] Williamson, Wachter and Harris (1975) tackled this issue theoretically, arguing that firms design workplaces and internal job ladders to orient employee attention toward longer-term rewards. The strategy is employed in the hope of reducing costly haggling over the disposition of rents. There is much more on the hold-up problem, an exceptionally important concept that has been largely ignored in market-centric modeling, later in the analysis.

An important building block in workplace-equilibrium modeling is the separation of work establishments by size. *Large establishments* are posited to have relatively high ratios of specific to general human capital, with specificity implying team production, economic rents, and significant turnover costs.[34] Relevant workplace information is compromised by production that is not additively separable by individual employee input, creating asymmetries and agency problems. A related characteristic is long information chains from workplace behavior to that level of management whose objectives are well synchronized with the goals of the organization as a whole. Meanwhile, *small establishments* are posited to have nil investment in specific capital, short (efficient) internal information chains, separable production functions, and insignificant turnover costs. The intuitive size heterogeneity results in markedly different labor-pricing processes in large versus small firms.[35]

About a quarter of paid employment in the United States occurs in firms with fewer than 25 employees; at the other extreme, about a third of all employees work in firms with more than 1,000 employees. The evidence consistently shows that workers in large establishments are paid more than similar small-firm employees. In one careful example, James Pearce (1990) reported that, after adjusting for human-capital characteristics, large-firm workers earned some 12 per cent more than their small-firm counterparts.[36] For overviews of the abundant evidence, see Mortensen (2003) and Chapter 10.

In small establishments, internal information systems need not be complex. Typically, the owner-manager directly supervises workers as a by-product of his or her other duties, limiting employee latitude to vary behavior. In workplace-equilibrium theory, such cost-effective

---

[34] Within large establishments, there will typically be pockets of employees whose functions are such that their productivity is readily monitored by management. Many workers in sales fit that description and, therefore, have pay packages that include pay-for-performance formulas. Such wage arrangements produce significant selection effects, as ambition or skill induces individuals into employment that can be effectively monitored and rewarded. See Lazear (1996). In the TVGE model, of course, employee venue-homogeneity eliminates such variation.

[35] The workplace model is a particularly powerful example of the "contested-exchange" class of information problems – i.e., the complications rooted in the costs of monitoring and enforcing contractual relationships.

[36] Also using U.S. data, Kevin Reilly (1995) developed evidence that large establishments offer more specific-training opportunities than their small counterparts. As a result, they have greater incentive to use wage policy to encourage longer job tenure of their employees; or they are simply rationally responding to already established wage rents. As noted above, in the TVGE model class, the use of human-capital differences in estimating wage premiums associated with establishment size imparts a downward bias to the reported effects.

monitoring sharply restricts WER sensitivity to changes in the wage paid. This is not a new idea. Alfred Marshall (1891, p.284) observed: "The small employer has advantages of his own. The master's eye is everywhere; there is no shirking by his foremen or workmen, no divided responsibility, no sending half-understood messages backward and forward from one department to another."

As implied by Marshall, large establishments are a different story. It is difficult and costly to design and operate a long internal information system that does not lose substantial data, especially on what workers actually do on the job. Information systems are further complicated by synergistic employee tasks and team production. Long experience demonstrates that it is particularly problematic to monitor cooperative effort.[37] From Herbert Simon (1991, p.32): "For the organization to work well, it is not enough for employees to accept commands literally. In fact, obeying operating rules literally is a favorite method of work slowdown during labor-management disputes, as visitors to airports when controllers are unhappy can attest. What is required is that employees take initiative and apply all their skill and knowledge to advance the achievement of the organization's objectives."[38]

Moral hazard has motivated a metamorphosis of management from hands-on, direct supervision with an emphasis on worker sanctions in small establishments to management by a web of (formal and informal) rules, rooted in the belief that fair treatment will encourage voluntarily cooperative behavior, in large establishments. The most critical management task in large organizations is convincing workers to accept the firm's objectives as their own.

---

[37] From Demsetz (1995, p.18): "This synergistic interaction makes it difficult, even impossible, to isolate the contributions to the value of output purely attributable to a single input." Management control decreasing in establishment size has occasionally been featured by economic theorists, e.g., Stigler (1962) and Williamson (1975).

[38] The effect of size on the capacity of management to control worker behavior has been studied extensively by sociologists. Richard Simpson (1985), for example, insightfully described and contrasted systems of direct and bureaucratic control. In the first, found in small offices and shops, the boss supervises employees directly, issuing orders without the intervention of formal rules governing the interaction. As workplaces grow larger, however, direct control gives way to a hierarchy of supervision and bureaucratic control. In the United States, nearly one in ten employees works as a manager, creating the need in big establishments for managers to manage managers. This bureaucratic approach is characterized by formal rules and procedures that define appropriate conduct of workers, managers, and their managers, making use of formal incentives and sanctions to help compensate for lost information on workplace behavior.

BOX 2.2: SCIENTIFIC MANAGEMENT AND WORKPLACE EXCHANGE

The transformation of workplaces into the particular environment that most efficiently accommodates the specialization and scale required after the Second Industrial Revolution is perhaps the greatest socioeconomic undertaking of the past 150 years. Understanding that well-documented evolution would help macro theorists better understand the complex nature of exchange in the modern economies they are attempting to stabilize. The following elaborates on the colorful role of Frederick ("Speedy") Taylor in the crucial process.

Scientific Management

The generalization of exchange helps the formal economic method accommodate "scientific management," developed by Frederick Taylor a century ago. It was a revolutionary approach to managing labor as tasks became increasingly specialized and workplaces more complex in the aftermath of the Second Industrial Revolution. Peter Drucker assessed Taylor's work as perhaps "… the most powerful as well as the most lasting contribution America has made to Western thought since the Federalist Papers." (Kanigel (1997), p.10) Taylorism illustrates how the functions of management were altered and expanded as establishments grew in size, and no historical analysis of worker on-the-job behavior is complete without showing where it fits in.

In scientific management, systematic measurement is applied to the problem of increasing efficiency in routinized work. Taylor conducted his experiments with a stopwatch, timing workers' movements and rest pauses. Selection criteria were established to place workers in jobs for which they are most suited. They are then instructed in the most efficient methods to use in their work tasks. He introduces time-and-motion studies to manual employment and helped design the devolution of complex labor processes into simple, repetitive tasks. Taylorism provided early blueprints for Class-I jobs, characterized as monotonous and routine.

Fair Treatment

Taylor, however, opened the door to a broader understanding of the specialized workplace when he argued that, even with optimal technical procedures, cost-efficient production requires worker cooperation. From Stessin (1960, p.5): "Taylor called for a 'mental revolution' by management in its attitude towards the workforce. He argued that if employees were to be won over to scientific management, they must have the assurance of fair treatment by the employer. This meant that management must voluntarily adopt new concepts of boss-worker relations. [For example] the supervisory must abandon his prerogative of 'instant dismissal' and substitute 'just cause' as the standard for separation from a job."

Even the stopwatch-carrying developer of time-and-motion studies recognized, as a result of his own experiences in a range of workplaces, that labor contracts are inherently incomplete and that employees ubiquitously prefer fair treatment from management, making mainstream economic theorists lonely adherents to the worker-as-robot view.

Big corporations produce a substantial share of GDP in modern economies. Imperfect information forces large establishments to find substitutes for direct supervision, developing administrative mechanisms that support workplace exchange. From John Dunlop (1988, p.47): "Every workplace of size that persists over time develops and in turn is governed by a complex of rules created by its industrial relations system. Such a 'web of rules' emerges irrespective of labor organization or collective bargaining; no continuing workplace is ever truly unorganized."[39]

Heterogeneous Jobs

Workplace modeling must additionally account for the nature of jobs, further introducing useful structure into its description of labor pricing and use. One of Adam Smith's best-known injunctions is that rational wage determination cannot be independent of work-task characteristics. WERs for scientists conducting original research likely differ markedly from exchange functions of production workers in a large factory. Generalized-exchange theory uses a bimodal separation to introduce fundamental job heterogeneity. The employees themselves retain useful point-of-hire homogeneity.

Complications arising from employee consciousness become more significant as work tasks become increasingly specialized and routinized. Work is divorced from end products, limiting the range of tasks performed and requiring navigation of a detailed hierarchy of authority. Routinized jobs are denoted *Class-I*. Using the earlier notation, such employment imposes two significant restrictions on formal workplace exchange:

- Nonpecuniary outcomes are an inherently minor component of total employee outcomes; in Class-I exchange, $\mathring{O}^N/\mathring{O}$ is relatively small.

---

[39] It is interesting to consider Dunlop's experience-based observation about spontaneous workplace organization in the more general context of the Austrian school's high-theory concept of spontaneous order. Hayek (1944) argued that any institution emerging in the absence of intentional manipulation is the unintended (spontaneous) consequence of the interaction of self-interested individuals, endowing it with all the information dispersed among the various agents.

- Firm personnel decisionmaking, whenever confronting substantial workplace information costs and asymmetries, is more informed by worker fixed inputs, particularly seniority, than by inherently imperfect measurements of on-the-job behavior. In Class-I workplace exchange, $\mathring{I}^f/\mathring{I}$ is relatively large.[40]

The formal definition of Class-I jobs is:

$$(2.10) \qquad \mathring{O}^N/\mathring{O} < \mathring{I}^f/\mathring{I}.$$

The remainder of the employed workforce holds *Class-II* jobs, defined as:

$$(2.11) \qquad \mathring{O}^N/\mathring{O} \geq \mathring{I}^f/\mathring{I}.$$

This second job class significantly, but not exclusively, includes employment that yields substantial personal satisfaction from the performance of the work tasks themselves.

The bimodal employment separation is easily recognizable in modern specialized economies. Class-I jobs are shaped by the specialization imposed by large-scale, high-volume (goods and services) production technologies, occurring in workplaces characterized by costly, asymmetric information. Such routinized employment is "unpleasant mainly because it fails to stimulate the worker yet prevents him from seeking stimulation elsewhere." (Scitovsky (1977, p.92)) Adam Smith drew early attention to the progressive simplification of work tasks, concluding that a significantly restricted range of on-the-job activity is a source of substantial boredom.[41]

---

[40] Class-I employees also prefer that management, in constructing and implementing incentives, emphasize fixed inputs as opposed to imperfect measures of OJB. The assignment of greater significance to $\mathring{I}^f_j$ than to I^E_j makes workplace exchange less arbitrary and, consequently, less likely to generate discontent. It should additionally be noted that, although fixed inputs such as education are trivialized by the assumption of point-of-hire worker homogeneity, their inclusion would enrich, not damage, the analysis.

[41] Philosophers would describe Class-I jobs as "instrumental", i.e. its value is not intrinsic but results from what it can be used to achieve. The employment class is vividly illustrated in Michael Brawoy's study (1979) of life on a factory floor in Chicago, where he documented the various ways devised by workers to make the time pass and help them put up with the monotony of the work. Also noteworthy here, in a survey conducted in the mid-1990's, Freeman and Rogers (1999) asked a large sample of employees: "On an average day, what best describes your feeling about going to work? Would you say you usually look forward to it, wish you didn't have to go, don't care one way or the other." (p.44) They found that "about one-third of the workforce are not eager to do their job." (p.43)

Moreover, large bureaucratic establishments make no secret of their general practice of rewarding employees in routinized jobs more on the basis of their fixed characteristics, particularly seniority, than on inherently flawed measures of their performance on the job.[42] Extensive evidence demonstrates the ubiquity of awarding wage increases, eligibility for desirable job transfers, and ranking in a layoff queue on the basis of employee seniority.

Meanwhile, Class-II jobs tend to be more of a hodgepodge, variously reflecting employment characteristics that variously include substantial satisfaction from the work itself (pushing up $\mathring{O}^{N}/\mathring{O}$) and cost-effective OJB supervision (pushing down $\mathring{I}^{f}/\mathring{I}$). Using longitudinal surveys, Kohn and Schooler (1982) identified "occupational self-direction" as especially significant in shaping an employee's response to his or her job. Work tends to be intrinsically rewarding when it requires individual initiative and independent judgment, implying complex workplace tasks, lack of routine, and remote supervision.[43] They found, not surprisingly, that jobs with such characteristics tended to be higher-status positions in their organizations. Richard Florida (2002) has named workers who enjoy relatively high non-wage benefits from their jobs "the creative class". Including scientists, engineers, artists, musicians, designers, and knowledge-based professionals, he estimates this creative class accounted for roughly 30 percent of total employment in the United States in 2000, up from 20 percent in 1980.[44]

---

[42] The Industrial Relations Research Association has examined this issue in some depth. In a summary of the relevant IRRA studies, Kleiner *et al.* (1987, p.108) concluded: "Using data collected within firms, they all reached similar conclusions – that pay level is weakly or not significantly correlated with performance rating and is more strongly related to seniority and education." Indeed, wage-setting procedures in large firms resemble a system of entitlements, with their emphasis on across-the-board increases, catch-up to consumer price inflation, and seniority-based hierarchies used in job transfers and layoffs.

[43] Tournament theory, for example, helps explain the behavior of employees with intrinsically satisfying jobs who are seeking advancement in an organization. (See Lazear, 1995.) It is not surprising that academic economists, themselves in an inherently self-directed occupation initially focused on the achievement of tenure, would empathize with the tournament model as a workplace description. That model, however, provides little predictive power for the behavior of workers in Class-I jobs that demonstrate fundamentally different outcome-input characteristics than does university research and teaching. See also the theory of occupational choice constructed by Annable and Fruitman (1973) for evidence on satisfaction derived from performing a job and its influence on wage determination for high general human-capital employees.

[44] At the other extreme (Class-II jobs), many immigrants, especially those who are undocumented, assign relatively low importance to their fixed characteristics. Recall that, in this introductory presentation of workplace economics, workers have been assumed to be market-homogeneous, limiting sources of heterogeneity to the nature of jobs and the size of work establishments.

It is noteworthy that economists at research-oriented universities hold class-II jobs. Their personal experiences, focused on work satisfaction, idiosyncratic hiring processes, evaluations based heavily on publications and other easily observable measures of performance, and tenure, poorly correspond to the circumstances and behavior of workers in class-I jobs. Economists must recognize that their personal experience cannot be generalized to most of the labor force. The many attempts to do so have resulted in badly misleading models, especially the ever more elaborate analysis of frictional unemployment by modern search theorists.[45] (See Chapter 5.)

IV. UNBUNDLED WORKPLACE EXCHANGE

From baseline equilibrium, disturb the worker with a reduction from the reference wage. If job inputs are not altered, he or she now confronts disequilibrium workplace exchange:

$$(2.12) \qquad (\Delta\mathring{O}+\mathring{O})/\mathring{I} < \mathring{O}/\mathring{I},$$

where ΔO represents the negative departure from the preferred job outcome (caused by the wage reduction from the reference rate, $W^{\mathring{n}}$).

If the best-alternative-job reference standard is violated (sup $\mathbb{K}=W^a$), the rational employee quits, moving to the better position. If only the interpersonal or intertemporal reference standard is violated and other outcome components are unchanged, the worker would remain on the job but be dissatisfied. Truman Bewley (1999, p.43) concluded the following from his intensive interviews with firms on their pay policies: "Within a company, pay inequality offends (indeed, sometimes outrages) employees and destroys trust…." A production input with the capacity to be

---

[45] Given the final word in this introduction to formal job-content analysis, Edmund Phelps provides a relatively optimistic assessment of the nature of employment compatible with economic progress. Unlike classical economists who foresaw the rising incidence of class-I jobs as an inevitable result of specialization, Phelps (2009) in his work-in-progress on economic growth moves class-II employment to center stage: "… if we are going to have any possibility of intellectual development we're going to have to have jobs offering stimulating and challenging opportunities for problem solving, discovery, exploration, and so on. And capitalism, like it or not, has so far been an extraordinary engine for generating creative workplaces in which that sort of personal growth and personal development is possible; perhaps not for everybody but for an appreciable number of people, so if you think that it is a human right to have that kind of life, then you have on the face of it a justification for capitalism." Quoted in Vane and Mulhearn (2009, p.121).

distrustful, offended, and even outraged about compensation substantially complicates factor pricing.

The now dissatisfied employee rationally alters his or her input to the job – the only available route to restoring equilibrium. Such workplace behavior has been long understood by practitioners. In his text on wage administration, Patten concluded: "Individual employees are 'bargaining' every day when they decide to remain with, quit, or reduce their contribution to an employer."[46] Steps to redress perceived inequitable treatment by management via diminished $\acute{Z}$ can include actions beyond reduced cooperative effort on the job: absenteeism, tardiness, petty sabotage, organized slowdowns (often strict compliance to management rules), and work stoppages.[47]

Here is the fundamental, real-world fact upon which workplace-equilibrium theory is constructed. *Employers have long believed that dissatisfied workers, given the latitude, adversely alter their behavior on the job*.[48] That commonplace fact and its profound implications motivate a more powerful, policy-relevant macroeconomics.

If working in a large establishment, the dissatisfied employee exploits the latitude provided by limited monitoring to modify his or her input to the job:

---

[46] Patten (1977), p.13.

[47] Illustrating the variety of actions available to damage an employer, Dickens, Katz, Lang, and Summers (1989), document the significance of worker theft. Chapter 7 usefully elaborates on organized work stoppages, demonstrating that they are a critical class of profit-adverse behavior by dissatisfied workers. Moreover, in nonunion circumstances, management always fears that dissatisfied workers will demand a NBER union-recognition vote. Putting it all together, employers need employees' cooperation; and they rationally invest in equity-based wages and workplace rules in the rational attempt to obtain it.

[48] Large-establishment employers tell a consistent story. (Recall Chapter 1.) They believe neither that worker OJB is wholly determined by technology nor that employees are inveterate shirkers. Instead, over the past century, they learned that workers prefer fair treatment, that workplace labor behavior is discretionary, and that policies designed to positively influence employee willingness to adopt management goals are useful in the pursuit of profit. There should be no dilemma here for economic theorists. If employers believe that OJB is variable, adversely influenced by worker dissatisfaction, and economists do not, then economists must change. It should also be noted here that rational worker behavior is broadly consistent with reciprocity theory, especially in that people devote resources to punish others for what they perceive to be hostile acts. Workplace reciprocity could be crudely incorporated into a game-theoretic framework, producing a continuous-strategy prisoner's dilemma. See Rabin (1993).

(2.13) $\qquad (\Delta\mathring{O}_j+\mathring{O}_j)/(\Delta\mathring{I}_j+\mathring{I}_j) \sim \mathring{O}_j/\mathring{I}_j.$

The worker cannot decrease fixed inputs to the job, implying that cooperative behavior is altered. Given that the change in outcomes is a reduction from the reference wage, preference relation 2.13 can be restated:

$$\Delta W_j/\Delta(\acute{Z}_jH_j) \sim \mathring{O}_j/\mathring{I}_j; \text{ and}$$

(2.14) $\qquad (\Delta(\acute{Z}_jH_j)/(\acute{Z}_jH_j))/(\Delta W_j/W^{\acute{n}}_j) \sim (W^{\acute{n}}_j/\mathring{O}_j)/((\acute{Z}_jH_j)/\mathring{I}_j).$

For workers in Class-I jobs, relation 2.14 simplifies to:

(2.15) $\qquad (\Delta(\acute{Z}_jH_j)/(\acute{Z}_jH_j))/(\Delta W_j/W^{\acute{n}}_j)>1,$

which is equivalent to:

(2.16) $\qquad (\Delta\acute{Z}_j/\acute{Z}_j)/(\Delta W_j/W^{\acute{n}}_j)>1.$

Relation 2.16 is named the *unbundled wage condition* and is illustrated in Figure 2.1. Employee OJB remains unbundled throughout the range of feasible reductions from the reference wage ($W^m \leq W_j < W^{\acute{n}}_j$). The $\acute{Z}_j$ contraction ($\Delta\acute{Z}_j$) is reasonably posited to decelerate relative to labor-price cuts as the input of employee cooperative services approaches the technological minimum associated with cost-effective workplace monitoring ($\acute{Z}^m_j$). Rational employee behavior implies that price-mediated exchange in the large-establishment workplace is a nonconvex set. Given that wage setting demonstrating local minimization of unit labor costs is insufficient to assure profit maximization, firm labor pricing must be additionally constrained by labor's market opportunity cost such that:

(2.17a) $\qquad (W_j \mid W^{\acute{n}}_j/\acute{Z}^{\acute{n}}_j < W^m/\acute{Z}^m_j)=W^{\acute{n}}_j; \text{ and}$

(2.17b) $\qquad (W_j \mid W^{\acute{n}}_j/\acute{Z}^{\acute{n}}_j \geq W^m/\acute{Z}^m_j)=W^m_j.$

In Figure 2.1, the WER tracks rational employee OJB ($\acute{Z}$) as the wage paid (W) moves relative to the reference wage ($W^{\acute{n}}$); the radius vector tracks the firm's efficiency wage ($W^n=\max(\acute{Z}/W)$). The conditions for which the reference wage ($W^{\acute{n}}$), consistent with worker utility maximization,

equals the efficiency wage (W^n), consistent with employer profit maximization, have been identified: large, specialized work establishments offering Class-I jobs. The efficiency wage is used by LEV firms to minimize unit labor costs, while the reference wage is rooted in interpersonal, intertemporal, and best-alternative-job reference standards used by workers to assess their satisfaction with the firm's compensation policies and their own workplace conduct. Neither wage concept is well defined absent the other.

The relevant WER is, furthermore, discontinuous at $W^{\acute{n}}=W^n$, implying a severely truncated labor-input schedule. Once aggregated, LEV labor supply has powerful macrodynamic consequences. In the baseline model, profit-seeking firms confronting unbundled $\acute{Z}$ cannot pay more or less than W^n, microfounding the crucial meaningful wage rigidity that was defined in the first chapter.[49] Generalized-exchange analysis has derived the holy grail of original efficiency-wage theory – a profoundly important outcome.

FIGURE 2.1. MODERN GEOMETRY OF EFFICIENCY-WAGE OPTIMIZATION

---
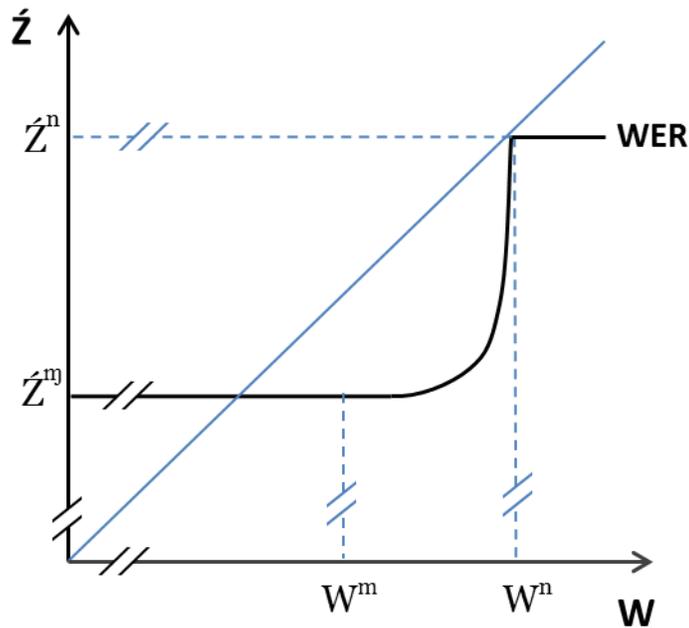
[49] Nominal wages are downward rigid absent $\mathbb{K}_j$ recalibration. Chapter 3 demonstrates $\mathbb{K}_j$ to be rationally unchanged in the context of stationary employment fluctuations.

BOX 2.3: FAIR TREATMENT AND THE HUMAN BRAIN

Brain imaging laboratories, employing advances in magnetic resonance imaging technology, have been coming up with startling insights into how human brains are organized and function. The core idea of functional imaging is to associate a detailed picture of brain activity with the moment a thought or activity occurs. Researchers have discovered that each thought or activity activates a particular constellation of brain regions and have been mapping a number of those functional networks. One such network, for example, is associated with falsehoods. Telling a lie triggers neurons in different regions of the brain than does telling the truth.

The exploration of economic issues is a logical application of neuroimaging research. Researchers at Princeton University have been mapping brain activity while economic choices are being made. In one particularly relevant experiment, subjects play a familiar game (which has come to be known as the "ultimatum game"). The first player is given $10 and a choice. He or she can give any part of the money to a second player, who then also has a choice. If the offered cash is accepted, both players keep the allocated money. If rejected, each get nothing. Any subgame perfect equilibrium of the ultimatum game mandates that the minimum ($1) be offered and accepted. Economists know that something is preferable to nothing. Yet, in the experiments, many players reject low-ball offers; and the modal proposal is to split the money. What is going on?

Using MRI scanners, it has been discovered that stingy offers stimulated a part of the brain known to be associated with feelings of disgust. From Colin Camerer: "They can predict with good reliability, from looking at the brain, what a person will do. People whose brains are showing lots of disgust will reject offers."* The research suggests what wage administrators already know. The perceived equity of economic outcomes matters. Moreover, unfair treatment motivates a desire to punish the perpetrator, even if the punishment involves personal cost.

Generalized-exchange theory asks less of workers than the ultimatum game. Equilibrium workplace dynamics is derived from employee preference for fair treatment by management as well as the urge to retaliate when he or she perceives inequity. In large establishments, however, retaliation need not incur significant personal cost to the employee. A large part of the job of personnel managers is well understood to be concerned with designing policies that do not trigger the "disgust" part of employee brains.

*Virginia Postrel, "Economic Scene," *The New York Times* (February 27, 2003), p.C2.

V. MORE IMPLICATIONS OF WORKPLACE EQUILIBRIUM

Two-venue WERs. With axiomatic preferences and technological constraints providing the foundation, unbundled workplace exchange has been derived from the rational interaction of employers and employees in large establishments offering routinized jobs. The workplace-exchange relations for large and small establishments respectively are:[50]

$$(2.18a) \qquad \acute{Z}_j = f(W_j, W^n_j, \acute{Z}^m_j, W^m), \text{ such that}$$
$$\text{if } W_j \in [W^n_j, W^m], (\Delta \acute{Z}_j / \acute{Z}^n_j)/(\Delta W_j / W^n_j) > 1;[51]$$

$$(2.18b) \qquad \acute{Z}_k = \acute{Z}^m_k, \text{ such that}$$
$$\text{if } W_k > W^m, (\Delta \acute{Z}_k / \acute{Z}^m_k)/(\Delta W_k / W^m) < 1.$$

The unbundled WER of equation 2.18a restricts the simultaneous optimization of firm labor pricing and worker cooperative behavior sufficiently to produce workplace equilibrium at $W_j = W^n_j = \max(\acute{Z}/W) = W^{\acute{n}} = \sup \mathbb{K}_j > W_m$ and $\acute{Z}_j = \acute{Z}^n_j = \acute{Z}^{\acute{n}}_j$. Such equilibrium persists until $\mathbb{K}_j$ is rationally recalibrated. (See Chapter 3.) Rational labor-pricing reconciles profit- and utility-maximization, necessarily occurs in the workplace rather than the marketplace, and powerfully generalizes unbundled $\acute{Z}_j$ to all large establishments offering Class-I jobs.

In a critical innovation, the baseline WER supply schedule determines the optimal wage on its own, requiring no interaction with labor demand. Optimizing workplace exchange equilibrates the employers' efficiency wage, minimizing unit labor cost, and employees' reference wage,

---

[50] The unbundled-bundled WER venues are variously and equivalently referred to as large versus small establishment, workplace versus marketplace, high versus low productivity, high versus low specialization, high versus low wage, rent versus no-rent, and good versus bad jobs. Each of the unbundled-venue names strictly refers to the workplace venue (equation 2.18a) that offers Class-I jobs and is sufficiently restricted by costly, asymmetric employer-employee information to support rational payment of labor rents. Each of the bundled-venue names refers to the remaining workplaces (equation 2.18b).

[51] If $W_j > W^n_j$, $(\Delta \acute{Z}_j / \acute{Z}^n_j)/(\Delta W_j / W^n_j) = 0$. Unbundled worker OJB is not a new idea in economics, having been recognized and analyzed by middle 20th-century labor economists. In a noteworthy example, Richard Lester (1951, pp.46-48) provided an insightful description of the sensitivity of worker OJB to wage cuts, distinguishing between time spent on the job and effort and "diligence" produced by employees. He argued that, by cutting wages, employers were unilaterally altering the terms of the employment arrangement, making them vulnerable to retaliation in form of reduced effort, inattention to quality, etc. While circumstance could produce $W^n_j = \sup \mathbb{K}_j = W^m$, that outcome is unlikely to exist and extremely unlikely to persist. It can in the interests of model tractability be ignored in this introductory exposition of the generalized-exchange model class, implying $W^n_j = \sup \mathbb{K}_j > W^m$.

satisfying their axiomatic preference for equitable treatment, capturing the essential properties broadly known to govern labor management and pricing in large establishments. (For elaboration, see Chapter 8.) Demand-independent labor pricing produces, over the business cycle, rational downward rigidity of nominal wages. It also allows firms' rational expectations of product demand to play an explicit, direct role in determining the level of employment. (See Chapter 6.) The essential WER discontinuity orients baseline labor-pricing dynamics around preventing cuts from the efficiency/reference wage.

In constructing the model so far, $\mathbb{K}_j$ has been posited to be unchanged, an assumption that is dropped in the next chapter. Two related definitions are used throughout the analysis. First, the calibration of $\mathbb{K}_j$ (and, therefore, W^n_j) is durable for a period t to t+ḳ such that ḳ>0 if unchanged $\mathbb{K}_j$ remains consistent with rational behavior. Second, simultaneous employer-employee rest periods in the space of intra-firm decision rules, if consistent with durable $\mathbb{K}_j$, are named *baseline workplace equilibrium*. The capacity of such equilibrium to persist for extended periods of time, perhaps reaching into decades and accommodating a great deal of stabilization-relevant dynamics, will be demonstrated in Chapter 3.

Meanwhile, small firms, effectively monitoring workplace behavior, also simultaneously optimize labor pricing ($W_k=W^m$) and employee behavior ($\acute{Z}_k=\acute{Z}^m_k$), providing the conditions of workplace equilibrium when worker OJB is bundled. Profit-seeking small employers must pay the market wage. Their employees, lacking exploitable workplace-information imperfections, cannot establish interpersonal and intertemporal reference standards, truncating $\mathbb{K}_k$ and refocusing worker attention on optimizing work-leisure choice and searching for good jobs.[52]

Implicit contracts. Equation 2.18a provides a port of entry into coherent macro modeling for the powerful, albeit largely lost to modern theorists, literature on implicit contracts. Rational

---

[52] Two-venue modeling microfounds Okun's (1981) "career" and "casual" employment and Clark Kerr's (1954) "balkanized" labor sectors, an earlier version of the basic idea. Rational activities in Kerr's "secondary" sector (here named the *large-establishment venue* characterized by unbundled worker OJB) differ fundamentally from optimization in his "primary" sector (the *small-firm venue* characterized by bundled OJB). The derivation of unbundled WERs also microfounds the Kerr *et al.* interrelated assertions that secondary-sector money wages are downward rigid and are contemporaneously independent of the labor market. From Dunlop (1988, p.67): "… everything we know about labor markets and wage determination explicitly precludes any simple or direct or stable relationship between the general level of wage, salary, or benefit schedules and aggregate unemployment…."

workplace exchange in the unbundled venue is transacted via that contract class. TVGE modeling has improved on textbook descriptions of the phenomenon by generalizing its circumstances beyond worker risk aversion. Modern workplace analysis roots implicit contracts in more universal preferences, identifies the necessary role played by costly, asymmetric intra-establishment information, and specifies implicit enforcement mechanisms that are recognized, and respected, by large employers. From Hallock (2009, p.69): "Most workers have one employment contract that is explicit and another one that is implicit. The explicit contract specifies working hours, compensation, and job tasks. The implicit contract involves expectations about the extent to which the employment relationship is not just a payment for labor on the spot market but instead is likely to continue over time."

Stability. Baseline optimization, in the context of unbundled $\acute{Z}_j$, produces an important OJB asymmetry. All disturbances begin from equilibrium ($W_j = W^n_j = W^{\acute{n}}$). Preference relations 2.7 imply that positive compensation shocks (i.e., W_j increasing from $W^{\acute{n}}_j$) do not disturb workplace equilibrium and, consequently, have no effect on $\acute{Z}_j$.[53] TVGE macroeconomics neither requires nor accommodates problematic gift-exchange.

More interesting action occurs when W_j decreases from $W^{\acute{n}}_j$, falling increasingly below the established reference rate. Such wage cutting causes dissatisfaction, and employees restore workplace equilibrium by altering their on-the-job behavior. In large establishments offering Class-I jobs, durable $\mathbb{K}_j$ implies that pay cuts are inconsistent with profit maximization. Asymmetric labor behavior in the neighborhood of W^n_j combines with firm profit-seeking to establish LEV workplace-equilibrium stability at $W_j = W^n_j$, $\acute{Z}_j = \acute{Z}^n_j$, satisfying Samuelson's correspondence principle and validating the earlier use of comparative statics.

Dynamic substitution effects. Optimization implies agent willingness to substitute inter-temporally among arguments in utility functions. Labor-related tradeoffs, however, have proved analytically challenging. In dynamic Walrasian (SVGE) equilibrium, it is typically not possible

---

[53] From human-resources expert Frederick Herzberg (2003, p.91): "… the opposite of job dissatisfaction is not job satisfaction, but *no* job dissatisfaction."

to derive substitution theorems that are valid for all parameter values, requiring theorist priors to make market-centric models capable of yielding operational theorems.

BOX 2.4: ROOTS OF IMPLICIT CONTRACT MODELING

Implicit contracts are the nonexplicit subset of self-enforcing, long-term agreements that govern the exchange of goods or services. Implicit contracts are effective because violation makes both parties worse off. Such contracts are important in specialized economies, accounting for a significant share of the pricing and allocation of scarce resources. They also have deep roots in macroeconomic theory, where Okun's "invisible handshake" is probably the best known, and certainly the best named, application of implicit contracting.

In a relatively early phase of the enduring quest to microfound Keynesian macroeconomics, Martin Baily (1974), Donald Gordon (1974), and Costas Azariadis (1975) innovatively utilized the implicit-contract concept in their independent analyses of involuntary job loss that motivates the lion's share of employment fluctuations over business cycles. In particular, they were interested in large, specialized establishments, inherently restricted by costly, asymmetric information, that became ubiquitous in the aftermath of the Second Industrial Revolution. Those firms necessarily generate and rely on specific human capital, implying relatively long-tenured employment. Requisite longer-term work relationships in turn inform reputational capital. The Keynesian theorists recognized that, in such circumstances, employers and employees are poorly understood as buyers and sellers in the spot labor market. Unfortunately, as a result of their working wholly within the mainstream SVGE model class, they confined the application of implicit contracts to elucidating risk sharing that rationally affects market labor pricing. The limited-application modeling could neither suppress wage recontracting nor fit the available evidence. For example, employees "buying" Baily's insurance against economic uncertainty would be paid a wage that is, on average, lower than their market opportunity cost. The data, by contrast, indicate that wage rents are paid in large, specialized firms.

TVGE modeling features implicit contracting in its most powerful form, directly governing the optimizing price-mediated exchange that occurs in large-establishment workplaces. In particular, implicit contracting enables the introduction of critical, deeply rooted worker preferences for fair treatment as well as retribution for equity violations into rational workplace behavior restricted by costly, asymmetric information and routinized jobs. Such preferences are axiomatic; their broad recognition by practitioners is readily apparent in the globally familiar equity-oriented personnel rules and practices that provide content to the workplace exchange mechanisms constructed by profit-seeking managements. Economic exchange is properly generalized.

Workplace-equilibrium theory alters workers' dynamic maximization problem. On the one hand, its difficulty is aggravated by employees' preference for fair treatment. Substitution may now

occur simultaneously at three margins: consumption and leisure; leisure and equity; and consumption and equity. On the other hand, as indicated earlier, the enriched constraints on optimization facilitate modeling. Most significantly, Keynes's Second Classical Postulate has been rejected, i.e., rational labor pricing in the specialized workplace chronically exceeds opportunity costs ground out in the marketplace. Being employed in rationed jobs considerably simplifies the intertemporal tradeoffs available to workers.

The *consumption-leisure tradeoff* is effectively suppressed by chronic wage rents. Management sets the workweek, and workers (preferring not to lose their rents) accept what is offered. From the perspective of employees, hours exogeneity (H^o) implies exogenous away-from-work leisure time (L^o), eliminating substitution between consumption and leisure: $\Delta L^o_{ij}(t)/\Delta C_{ij}(t)=0$. In generalized-exchange modeling, suppressing this first trade-off helps establish disturbances in aggregate nominal demand as the critical causal determinant of fluctuations in total employment and production.

Meanwhile, the *tradeoff between leisure on the job* (L^w) *and perceptions of fair treatment* ($W_j/W^{\acute{n}}_{ij}$) is suppressed in the specialized venue by the nature of productive OJB. Given the importance of cooperation and acceptance of management's goals in the circumstances of unbundled $\acute{Z}_{ij}$, discretionary OJB changes typically have little to do with distasteful physical effort on the job. On-the-job leisure (L^w) is broadly understood to be unaffected, within a zone of independence, by reductions in $\acute{Z}_{ij}$: $\Delta L^w_{ij}/\Delta \acute{Z}_{ij}=0$, $\Delta \acute{Z}_{ij}/\Delta(W_j/W^{\acute{n}}_j)\geq0$, $\Delta L^w_{ij}(t)/\Delta(W_j/W^{\acute{n}}_j)=0$.[54]

---

[54] In workplace-equilibrium theory, the "zone of independence" is restricted to reductions in cooperative labor input on the job ($\Delta \acute{Z}_{ij}<0$), in response to inequitable treatment, that are independent of leisure at work ($\Delta L^w_{ij}=0$). In practice, however, the zone of independence is likely more substantial. Even physical-effort shirking is different from, and less readily motivated than, the employee's preference to be somewhere other than on the job. Practitioners argue that, once at work, loafing may make the worker's experience less satisfactory and his or her time on the job go more slowly. Moreover, depending on the firm's supervisory practices, avoiding detection while physically shirking can be difficult and unpleasant – in other words, hard work. There is, on the other hand, ample documentation of employee dislike of significant work speed-ups. Insightful analysts have suggested a middle ground. Leibenstein (1976) argued that workers have a "comfort effort range" within which they can vary $\acute{Z}_{ij}$ without affecting their utility. Simon (1951) defined a "zone of acceptance", which he later summarized: "An employment contract contains all sorts of implicit (and explicit) limitations that set boundaries to the range of actions the employee will be directed to perform. These boundaries define the 'zone of acceptance' within which an employee can be expected to obey orders. The zone of acceptance is also sometimes called a 'zone of indifference', for the choice among alternative behaviors, while of major importance to the employer, may be of little or no concern to the employee." (Simon, 1991, p.31) Practitioners understand that, if work speed-ups push employees outside their indifference zones, latent preferences to shirk are activated.

The third tradeoff, between *consumption and fair treatment*, motivates the interesting optimizing employee behavior along with the intertemporal malleability of the WER. More specifically, *C*-for-$W^{\hat{n}}$ substitution, simplified by the assumption that employees strictly prefer consumption to fair treatment, animates the worker's intertemporal investment in specific workplace capital ($\mathbb{K}_j$), making the reference wage ($W^{\hat{n}}$) endogenous. Active-$\mathbb{K}_j$ equilibrium, modeled in the next chapter, is motivated by the third, surprisingly tractable, tradeoff.

MWR channel. Baseline wage recontracting, i.e., a worker is offered and rationally accepts (in lieu of losing his or her job) a reduced wage that does not violate opportunity costs, is now understood as a class of workplace exchange that requires either intra-firm information sufficient to support cost-effective employee supervision or a Class-II job. Costly, asymmetric workplace information and routinized jobs, inherently associated with Chandler's "new corporate forms" characteristic of the Second Industrial Revolution, rationally inhibit labor-price recontracting, a restriction from which crucial macroeconomic content follows.

Most important, the MWR channel *uniquely* microfounds the capacity of adverse nominal demand disturbances to induce involuntary job loss. Forced job separation has been shown to be the empirical engine of employment fluctuations in the U.S. economy. In an outcome vigorously sought since Keynes and the 1930s depression, generalized exchange has microfounded a robust, stabilization-relevant monetary theory of production and employment. (For elaboration, see Chapter 6.) Moreover, the tightly specified MWR channel, i.e., limited to large, specialized establishments offering routinized jobs and creating downward inflexible wages over the business cycle as well as chronic, time-varying wage rents, leaves numerous, particular, testable tracks in the economic data. Evidence consistent with the TVGE narrative is all over the place, from equity-based corporate personnel policies to the existence of forced job loss (both layoffs and downsizing), providing sharp contrast to mainstream SVGE thinking that must ignore much of the data generated by highly specialized economies.

VI. CONCLUSION

Like all theories, the labor pricing modeled in this chapter simplifies reality. Despite that, a great deal of the nature of wage determination in large, complex establishments has been captured. Employees, working in a hierarchy of authority, want to be treated equitably. Employers, confronting the inherent asymmetry and costs of workplace information, learned long ago that being attentive to that elemental preference in their labor-management practices, centrally including compensation, is consistent with profit seeking. Labor pricing is relocated to inside the firm, and its tight relation with market equilibrium, described in Keynes's Second Classical Postulate, is severed. Macroeconomics that is both coherent and stabilization-relevant becomes possible.

Misunderstood efficiency wages. Efficiency wages, particularly the original morale-centric formulation that Solow and I pioneered and that outlined how theorists could go about coherently scrapping the Second Classical Postulate, are today badly understood by economists. In their otherwise exemplary history of economic thought, Scepanti and Zamagni (2005, p.369) provide a typically inadequate description: "The theory of efficiency wages is based on three principal ideas. The first is that the intensity of work effort of each employee, and therefore the marginal productivity of labour, increases with an increase in wages. The second is that the workers' effort is also influenced by the level of unemployment, in that the fear of being dismissed for inefficiency increases with an increase in the probability of not immediately finding another job with the same pay – a probability that rises with the level of unemployment. The third hypothesis is that there is a type of asymmetric information, as firms are not able directly to ascertain the intensity of effort of hired worker or the ability of those to be hired. In these conditions, it is in the interest of firms to pay high wages to encourage workers' effort."[55]

---

[55] The first two characteristics erroneously describe morale-centric efficiency wages. The Scepanti-Zamagni depiction is relevant only to the Shapiro-Stiglitz shirking variant of efficiency wages. Despite being best known of all the EWT branches, Shapiro-Stiglitz generates flexible labor pricing and confines job loss to dismissal for cause, a trivial category that is incorrectly described, does not track the business cycle, and has no macro content. Indeed, business historians recognize shirking theory's central role for firing and fear to be a nineteenth-century idea that in large, specialized enterprises is destructive of profits and, consequently, has had little place in best-practices LEV labor-management relations for generations. For elaboration on the curious Shapiro-Stiglitz model, especially its irrelevance to understanding and management of macro instability, see Chapter 9.

BOX 2.5: META-EXTERNALITY

Generalized-exchange modeling greatly enriches A.C. Pigou's (1920) analysis of economic externalities, a powerful concept earlier noted by Henry Sedgwick and Alfred Marshall. Most relevant to macro theorists are production externalities, in which optimizing firm behavior can induce market inefficiency. From David Kreps (1990, p.289): "The idea in a production externality is that a firm, by its choice of a production plan changes the feasible set of production plans for other firms." In the present context, production plans are understood to rationally govern the pricing and use of scarce resources. To Pigou, economists have a moral responsibility to identify negative externalities and design government interventions that ameliorate their effects.

The TVGE reconfiguration of mainstream coherent SVGE macroeconomics has identified the most costly class of mismatched private and social costs, here named *meta-externality*. Profit and utility maximization in large, specialized establishments microfounds the MWR Channel that translates nominal demand disturbances into same-direction changes in production, employment, income, and profit. Adverse shifts in total spending broadly upset production plans, imposing periodic, widespread involuntary job loss and underutilized capital on specialized economies. In Pigou's language, the private costs of continuous, general decision-rule equilibrium are intermittently much less than the social cost of recession or, worse, depression. A neoclassical justification for the discretionary management of aggregate demand has been provided.

In the more general narrative, large establishments created by the Second Industrial Revolution rationally suppress wage recontracting. Such suppression follows from the generalization of optimizing exchange from the marketplace to the specialized workplace. Once true axioms are used to inform worker preferences and model intra-firm behavior as continuous-equilibrium phenomena, involuntary job loss and its associated market failures are best understood as a powerful set of meta-externalities that imply the need for aggregate-demand intervention. Coherent macro theory is made stabilization-relevant and the fruitful process of synthesizing competitive-market and Keynesian modeling, with their respective incentives and externalities, can begin in earnest.

A powerful feature of TVGE meta-externality is its practical immunity with respect to the Coase Theorem. Parties involved in continuous-equilibrium market failure rooted in the nonmarket MWR Channel cannot bargain macro instability into a Pareto-optimal outcome.

This chapter indicates, and the evidence broadly supports, that LEV employers rationally believe that cooperative OJB is damaged by *cuts* from $W^{\acute{n}}$, which is rooted in employee established reference standards $\mathbb{K}_j$. The next chapter demonstrates that rational workers must experience significant job downsizing to justify accepting wage reductions absent a tit-for-tat response. Labor productivity that is simply increasing in labor pricing is a very different, nonintuitive

process, badly complicating model stability and, outside of the limiting case of nutrition and health, supported by neither practitioner testimony nor the evidence. In the TVGE model class, the rational efficiency wage (W^n) is independent of market unemployment as well as market opportunity costs.[56] Profit-seeking LEV firms pay downward rigid wages over the business cycle and wage rents at all times; as a result, their employees understand the very low probability of "immediately finding another job with the same pay" no matter what the overall jobless rate.

Deep roots of original efficiency-wage theory. As indicated above, the original EWT literature has deep roots in actual behavior. In his study of the transition of the Lancashire cotton-spinning industry from at-home piecework to factories employing an on-site workforce, historian Michael Huberman traces the roots of modern OJB management to the early nineteenth century. The Second Industrial Revolution, motivating new corporate forms in order to effectively manage high-volume, rapid through-put production, began in earnest later in the nineteenth century. But it was anticipated by a small vanguard of larger-establishment, specialized industries. If employees do indeed have an axiomatic preference for fair treatment which influences their OJB, there should be supportive evidence prior to the broad large-firm reorganization of production.

From Huberman (1996, p.6, pp.13-14): "Workers entered the factories with customary notions of a fair day's pay for a fair day's work [from their experience with at-home production] and would expend optimal effort only if they were assured this rate…. Cooperation between firms and workers was not immediate. Pressed by competitive forces, firms initially challenged workers' control. But changes in work organization and constant attempts to cut rates of pay only raised

---

[56] The following elaboration anticipates, perhaps usefully, Chapter 9. With respect to unemployment, Scepanti and Zamagni are describing a particular EWT subset of follow-up models pioneered by Shapiro and Stiglitz (1984). Reflecting mainstream preference for analyzing market activity, subset theorists attempt to model workplace exchange within the coherent SVGE framework, identifying dismissal for cause as the forced job loss of interest and using fear generated by the high incidence of market unemployment to motivate cooperative labor input. A central role for firing and fear is a nineteenth-century idea that has had, for good reason, little place in best-practices LEV labor-management relations for generations. Moreover, Shapiro-Stiglitz labor pricing is downward flexible over the business cycle and cannot inform the channel through which nominal disturbances induce (temporary) layoffs and (permanent) job downsizing. Such models fail to justify discretionary demand management. The substitution of a tiny-incidence category of joblessness (dismissal for cause) for the vastly more important components of high- and low-frequency unemployment (layoffs and downsizing) is an indicative tip-off that SVGE efficiency-wage modeling is fundamentally different from the original workplace-exchange literature. That the two model classes share the same rubric has damaged the progress toward a policy-relevant formal macroeconomic theory.

unit costs. It took one generation for firms and workers to recognize the benefits of cooperation…. Fair wages took hold in large firms in the leading industrial centers of Manchester and Bolton. By 1850, firms in other areas adopted these policies, although there were important exceptions. Firms needed to show their commitment to the fair wage, and to this end they introduced work-sharing and a layoff strategy that protected senior workers. These arrangements were initially rules of thumb, but standard piece rates themselves became codified in written lists…. Yet all rules, whether formally stipulated or not, were governed by community standards of what was just and fair."