Volume Title: Fifty Years of Economic Measurement: The Jubilee of the Conference on Research in Income and Wealth

Volume Author/Editor: Ernst R. Berndt and Jack E. Triplett, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-04384-3

Volume URL: http://www.nber.org/books/bern91-1

Conference Date: May 12-14, 1988

Publication Date: January 1991

Chapter Title: Data Difficulties in Labor Economics

Chapter Author: Daniel S. Hamermesh

Chapter URL: http://www.nber.org/chapters/c5979

Chapter pages in book: (p. 273 - 298)

9 Data Difficulties in Labor Economics

Daniel S. Hamermesh

9.1 Introduction

In the fifty years of the existence of the Conference on Research in Income and Wealth, labor economics has become a leader among subspecialties in economics in linking empirical work and theory, in acquiring large amounts of data, and in making strides in analyzing those data. Despite this distinction there are substantial imbalances in data resources in this area and in the progress in understanding labor-market phenomena that the available data have made possible. Also, areas in which we think that our knowledge has been furthered by recent studies are in fact less advanced than we believe because of problems with data. Finally, the ability to generalize our findings is in many cases limited by difficulties involving the interaction of the sets of data used and the nature of the problems under study.

In section 9.2 I present a general framework for analyzing the appropriateness of a variety of data sets to the purposes for which they are used. This approach is narrower than that of Griliches (1986), who laid special emphasis on problems of measurement error. The view implicit here is, though, both broader and different from that of Stafford (1986). He concentrated on the few major longitudinal household data sets and developed an almost Schumpeterian theory of how newly available sources of data both are called forth by and, in turn, advance theory and inform policy. Most of his attention was focused on the use of these data sets in analyzing issues in labor supply. I pay attention to labor supply in section 9.3; the bulk of the paper considers, however, three other major areas of interest to labor economists in light of this discussion of the appropriateness of data sets.

Daniel S. Hamermesh is professor of economics at Michigan State University and a research associate of the National Bureau of Economic Research.

The author is deeply indebted to Jeff Biddle for many helpful suggestions and to Steve Allen, Paul Chen, Zvi Griliches, Harry Holzer, Nick Kiefer, John Pencavel, Sherwin Rosen, Jack Triplett, and Steven Woodbury for useful comments.

Much of the discussion is of labor demand, including issues of employment adjustment, "the elasticity of labor demand," and problems of labor-labor substitution that have been addressed by very few sets of data. Of particular note in this regard are the KLEM (capital, labor, energy, and materials) data on U.S. manufacturing assembled by Berndt, Fuss, and Waverman (1979) and others. Since much of our knowledge comes from these aggregate data (see Hamermesh 1986), it is essential to analyze how well they meet the criteria presented in section 9.2. Much of the rest of what we have learned recently comes from the estimation of complex production technologies applied to data from household surveys. In section 9.4 I examine the usefulness of these studies according to the criteria I set out.

Sections 9.5, 9.6, and 9.7 present shorter discussions of labor market–wide phenomena, of trade union behavior, and of the desirability of international comparisons. In the 1940s and early 1950s, labor economists engaged in massive studies of specific local labor markets. With the exception of Rees and Shultz (1970), this type of work ceased by 1955. Today's research on labor markets must be deductive from data on samples of workers in many markets. How well suited is today's approach to analyzing how a labor market operates, as compared to the approach of nearly two generations ago? Is there a possible compromise that can meet the objections to which each might be subject under the consideration of the appropriateness of data sets? In the past ten years interest has burgeoned in analyzing what, if anything, unions attempt to maximize. Much of the work has been on one particular set of data (Dertouzos and Pencavel 1981). How representative are these data? Are the available data resources sufficient to allow us to draw any general inferences about what unions seek to do? The cultural imperialism of American empirical economics should not blind us to the possibility that the structure that describes a relationship in the United States may not be representative of some (any?) other economies. Thus it is worth considering under what circumstances the consideration of descriptions of behavior for several economies is more or less important in generalizing about behavior.

Based on the general framework for analyzing the appropriateness of data sets and its specific applications to these central issues in labor economics, I draw conclusions about the types of additional data that should be collected. Because the issues are to some extent overlapping, it should be possible to address the lacunae in the data jointly rather than treating data problems in each area separately.

9.2 A Framework for Evaluating Data Describing Workers and Employment

The general linear model describing the structure of an economic relationship at a time that the researcher wishes to characterize (usually the present) can be written as

$$y_t = bx_t + e_t, \tag{1}$$

where y is the outcome variable, x is a vector of independent variables, b is a vector of parameters describing the structure of the relationship, t is time, and e is a disturbance term. The relationship that is estimated is not this at all, but is instead

$$Y_{IT} = b_{IT}X_{IT} + e_{IT}, \tag{2}$$

where the subscript I indexes the units chosen to represent the economic relationship, and T indicates the time(s) at which they were observed. I assume throughout that (2) is estimated by best-practice technique. Thus the assumption of a linear model is merely for expositional simplicity, and the discussion applies to more complex models too. Thus $\hat{b}_{IT}$ is an unbiased estimate of b_{IT} for the particular set of data $\{Y_{IT}, X_{IT}\}$ chosen to represent the relationship between y and x, so that $E(\hat{b}_{IT}/b_{IT}) = b_{IT}$.

While I assume that $\hat{b}_{IT}$ has all the nice statistical properties we desire, it can be viewed as the best estimate of only one of a large number of vectors of parameters b_{IT} based upon possible sets of data $\{Y_{IT}, X_{IT}\}$ chosen. Essentially there is a distribution of parameter vectors b_{IT} corresponding to the distribution of the data sets. The question of interest is whether

$$E(b_{IT}|\{Y_{IT}, X_{IT}\}) = b, \tag{3}$$

where the unsubscripted b is the true value of the parameter describing the relationship of interest to the researcher. Four questions are relevant in analyzing whether (3) holds: (1) Does the particular set of variables $\{Y, X\}$ that is chosen represent the true variables $\{y, x\}$ well? This is essentially a question of measurement and specification error. Both random measurement errors and systematic errors of measurement are likely to be important problems in labor-related data. They have, though, been well covered in the econometric literature, and I pay them relatively little attention in the discussion hereafter. (2) Is the sample that underlies the set of subunits $\{I\}$ that is used to estimate (2) representative of the population to whom the theoretical relationship (1) is supposed to apply? This goes far beyond the narrow econometric issue of sample selection bias that has received so much attention from labor economists and econometricians. (3) Is the set of time periods $\{T\}$ likely to allow the researcher to draw correct inferences about the relationship between y and x that holds at time t for the typical unit (perhaps different from the representative unit in $\{I\}$)? The issue here is one of structural change. (4) Are the intervals between units in the sets $\{I\}$ and $\{T\}$ appropriate for the relationship between y and x that is being studied? This is an issue of appropriate aggregation.

Consider the second question: Is the set $\{I\}$ typical of all the units in the economy whose economic behavior we are trying to describe by (1)? If the analysis of (2) for the data set indexed by $\{I, T\}$ is to be more than an econometric case study, $\{I\}$ should cover a broad set of subunits in the economy or cover a few typical subunits. If we are confident that the data set meets all of these worries, can it be used to draw inferences about the relationship in (1) in other economies during the time period under study?

The third question should be answered on two levels. The simpler, and more frequently discussed one, is that of structural change: What do the $\hat{b}_{IT}$ tell us about the structural relationship between y and x today, or has that relation changed so drastically that the estimation of (2) has become economic history that sheds little light on today's economy? The answer to this question depends on how rapidly structural change occurs in the particular relationship and on how far in time we are removed from the observations in the set $\{T\}$. The more complex issue is a combination of structural change and misspecification: Is the relationship between Y and X no longer the same as in (2) because of the growth in importance of additional factors, denoted by Z? If data were collected on the Z, and if nothing else made the data set $\{I,T\}$ unrepresentative for current purposes, it would be a simple matter to respecify and reestimate (2) and use it to draw inferences about today's structure. If, however, data on Z were not *or could not* have been collected, there is no hope of resurrecting (2) to analyze behavior today.

The length of "the short run" varies with the problem under study. The intervals in the set $\{T\}$ should be such as to make it possible for the estimates of (2) to inform us about the speed with which equilibrium is reestablished after the system underlying (1) is shocked. Also, while we often view cross-section data as allowing us to infer equilibrium relationships, that assumption is not necessarily valid. A time series of sufficiently long T can be useful in allowing the inference of the structure of the equilibria that arise after a shock. The problem here is to use the level of temporal disaggregation appropriate to the question under study. Another difficulty arises if the units in $\{I\}$ are too large to prevent us from assuming that all underlying relationships are linear, and thus that estimation over aggregated data yields unbiased results. Are they small enough so that decision makers' nonlinear responses can be detected, or are they so highly aggregated that nonlinearities and discontinuities are all smoothed out? The problem here is that of appropriate spatial disaggregation.

This discussion is couched in terms of estimation (of the true underlying parameter vector b characterizing the relationship of interest). Clearly, though, labor economists are interested in hypothesis testing as well as in estimation. A discussion similar to that above could be developed to deal with the structuring of data appropriate to hypothesis testing. The main difference would be that, in addition to the four problems discussed above, problems that produce the equivalent of the bias term in a mean squared error based on a comparison of $\hat{b}_{IT}$ and b, we would also have to consider the statistical distributions of the $\hat{b}_{IT}$ and of the set of b_{IT} that might arise from the entire range of choices of I and T used to estimate the relationship. The problems we consider here to analyze data appropriate to estimation are a subset of those necessary to analyze data appropriate to hypothesis testing.

9.3 Labor Supply—Synergy among Data, Estimation, and Theory

During the past twenty-five years the study of labor supply has been a central focus of labor economics. We have learned something about important phenomena; major threads of microeconomic theory that had not been used in empirical work have been explicitly employed in estimation; and these applications have generated important advances in econometric theory that have been used elsewhere in economics. The data are representative of the underlying populations being studied; there is no reason to assume the results are irrelevant today because of intervening structural changes; and there is no major problem of excessive aggregation of decision-making units in what are chiefly sets of data that have households as the units of observation.

The important explorations in labor supply occurred along with a flowering of data collection. Carefully constructed cross-section sets of household data became available during the 1960s; and the computer technology, both hardware and software, to analyze them was developed simultaneously. By the mid-1970s the major longitudinal household data sets, the National Longitudinal Surveys and the Panel Study of Income Dynamics, began to be used to study labor supply. It is impossible to believe that the development of these sources of information "forestall[ed] the demise of empirical economics," as implied by the title of Stafford's (1986) essay, or even to prove that causality ran from the development of these data to advances in theory and econometrics. It is difficult enough to prove causality in well-specified econometric models dealing with hypotheses that are grounded in economic theory. One should not expect in this case to demonstrate something that historians of science have debated and about which philosophers of science have prescribed for generations. One must believe, though, that we would know a lot less, and labor economists' fascination with problems of labor supply would be less intense, if these data sources had not been constructed and the resources devoted to them had instead gone into data more readily usable in other areas of labor economics or in other subfields of economics.

I believe there have been three major advances in the empirical study of labor supply: (1) The estimation of income and substitution effects; (2) the growth in our understanding of labor-force dynamics; and (3) the recognition of the life-cycle nature of labor supply. Obviously there have been advances in understanding other supply-related phenomena, such as household production, population, and the demand for education. None of these, though, is as central to labor supply as these three main thrusts; and in none of the other areas are the links among theory, econometrics, and data so well articulated.

The first advance, spurred by Mincer (1962), led to a flowering of research using various household surveys, particularly the Survey of Economic Opportunity, that eventually gave empirical meaning to the basic results of the theory of consumer demand. It allowed better predictions about the response of labor supply to changes in the parameters of income-support programs and more careful inferences and predictions about patterns of labor supply in a growing economy. Without microdata neither of these achievements could have been attained with the same precision. As a result of the research of the 1960s our knowledge is fairly secure about the *relative* magnitudes of labor-supply elasticities of different demographic groups, and some consensus limits have been placed on the range of the absolute magnitudes of these parameters.

The second advance taught us that for many groups there is substantial mobility into and out of the labor force. The work of Heckman and Willis (1977) and others demonstrated that it is as wrong to view a 65% participation rate as reflecting participation by 65% of the population all year long as it is to view it as reflecting part-year participation by the entire population. Without longitudinal data this demonstration could not have been made. The discovery affected the course of advances in theoretical econometrics, for it generated an interest in developing econometric techniques and borrowing techniques pioneered by sociologists to analyze the determinants of mover-stayer distinctions. These techniques have been used extensively in other areas of labor economics and have been applied to other subfields of economics as well. Thus, for example, the distinction between unemployment spells and unemployment duration (e.g., Clark and Summers 1982) and issues in the burden of unemployment could not have been analyzed without longitudinal data.

The final major advance saw labor economists putting empirical meat on the bones of life-cycle theory by analyzing intertemporal substitution using longitudinal data (e.g., MaCurdy 1981; and many others). Obviously this advance was spurred by developments in macroeconomic theory; but without the microeconometric analysis by labor economists, helped along by the general implausibility of the assertion that intertemporal substitution could be very large, we would not be fairly secure in our knowledge of its relative unimportance of affecting labor supply.

The development of information useful for research in labor supply and, more generally, in studying labor-force dynamics, has proceeded from aggregated Census of Population data, to micro cross-section data on households, to longitudinal surveys of individuals and households. These developments are no longer confined to the United States. Indeed, some of the most interesting efforts in collecting these sorts of data are being made elsewhere, particularly in Australia and Sweden.

One must ask, though, where this continuing concentration of resources on collecting longitudinal household data is leading us. As a brief foray into this question I extended Stafford's (1986, table 7.1) work categorizing published studies on labor supply in six major journals. Table 9.1 presents the results of this analysis.[1] It provides some indication that professional interest in labor

Table 9.1 **Articles in Major Journals: Labor Supply Subject, by Year, 1965–87**

	1965–69	1970–74	1975–79	1980–83	1984–87
Population size and structure	7	14	19	10	5
Household production	0	11	8	8	4
Labor supply of men	2	5	7	6	6
Labor supply of women	0	3	9	6	4
Labor supply of others and income support disincentives of UI, NIT, taxes, or other	2	7	16	15	8
Retirement	0	1	4	1	4
Educational demand	3	11	9	2	2
Migration	13	9	14	4	5
Total	27	61	86	52	38

Sources: Cols. 1–4 from Stafford (1986); col. 5, author's tabulation.

supply has been slipping since the late 1970s. The only growth areas in the mid-1980s were studies of retirement and migration. Interest in retirement was probably spurred by concern about an increasingly important general economic problem and by the creation of the Longitudinal Retirement History Survey. Interest in migration stemmed from concern about policy; there was no sudden availability of data that made it possible to examine the issue. Studies in the mainstream areas of labor supply, particularly in the effects on labor supply of government programs, have been of decreasing interest.

It is, of course, impossible to identify the causes of the reduced interest in the central areas of the study of labor supply. I would argue, though, that at least in part it stems from the lack of new advances in the kinds (not the quantities) of data that are available for this purpose. The rich lode of the feedback relationship between data development and the expansion of knowledge about labor supply is now yielding decidedly diminishing returns.

Clearly there are many areas that have not been well explored and many questions that can be answered with better data. We could take the position of Wagner, the dull student in *Faust,* "zwar weiss ich viel, doch moecht ich alles wissen."[2] The collection of panels that follow a cohort from school through middle age, for example, will enable us to distinguish better between the economic determinants of labor-force behavior and background effects, and between transitory economic effects and those stemming from life-cycle behavior. Additional studies of life-cycle behavior that tie labor supply to liquidity and labor-market constraints will undoubtedly be made. The potential for acquiring knowledge appears limited, though. It is not clear that the efforts of the 1970s and 1980s (since the studies in Cain and Watts [1973] that represent the initial phase of the first major achievement I discussed above) at refining our knowledge of labor-supply elasticities (or their component income and substitution effects) have done anything to narrow the agreed-upon range of estimates. There is nothing on the horizon or even imaginable that seems likely to provide the kind of spur to research in this area that it received from the development of new sets of data between 1965 and the late 1970s.

9.4 Labor Demand—a Case of Underdevelopment

Unlike the study of labor supply, in which the creation of large data sets led to tremendous strides in linking theory to empirical analysis, those of us who study labor demand have been less fortunate. The many interesting questions on the demand side have as important implications for policy as those more widely studied supply questions. Thus issues of the demand for older workers, the impacts of technical change and international competition on the distribution of employment, and the effects of mergers and acquisitions on job creation should be motivating research on the demand for labor in much the same way that interest in income maintenance programs spurred much of the research on labor supply in the 1960s and 1970s. That this has not happened—that we have made less progress in answering questions about labor demand—is largely the result of the failure to invest in the kinds of data that would allow us to obtain answers, a failure that continues today.

The questions and previous studies designed to answer them fall into two categories: those involving employment dynamics and those concerned with factor substitution. Let us consider the first group. One set of questions involves the analysis of paths of employment adjustment in response to exogenous shocks. Subsumed here is the attempt to discover the nature of the costs of adjusting employment that presumably generate observed adjustment paths. The analysis of how firms adjust employment leads to questions about how labor productivity (most simply, the output-total hours ratio) changes cyclically. These are inputs into the analysis of cost-based inflationary pressures, so that this aspect of the study of labor demand becomes a crucial macroeconomic issue. Similarly central to macroeconomics are the implications that adjustment paths have for the path of unemployment.

The study of employment adjustment should not be restricted to firms that are assumed to be infinitely lived. Rather, it should enable us to understand the economic process by which output shocks generate a continuing opening and closing of different work sites that in turn produces changes in employment. The analogy to the growth in the study of labor-force dynamics, including the study of gross flows data in the Current Population Survey (CPS), should be clear.

How well do existing studies of employment dynamics meet the criteria of appropriate aggregation, representativeness, and current structure? One broad group of studies (most recently, Morrison 1986) uses annual data on a set of factor inputs aggregated over a large number of establishments in manufacturing to analyze demand dynamics in the context of a model allowing for substitution among all pairs of inputs. (In Morrison's and many others' studies these are the KLEM data.) This strand of the literature has severe problems under two of our criteria. Almost all the available evidence (see Hamermesh 1988 for a survey) suggests that employment responds to shocks fairly rapidly. This means that annual data are inherently incapable of telling us much about the underlying path of employment adjustment: the data are too highly aggregated temporally. They are also too highly aggregated spatially. If there is any nonlinearity in the adjustment process at the microlevel, the use of aggregate data will in general fail to identify it. Aggregation should be done over the relationships estimated as characterizing the microunits, not over the microdata in a way that requires assumptions of linearity in those relationships. Since many reasonable structures of adjustment costs generate nonlinearities, this set of studies will not help much in identifying what generates the path of employment adjustment. Because the data cover only manufacturing, it is also hard to claim that the results do well on the criterion of representativeness.

Another group of studies (beginning with Nadiri and Rosen 1969 and extended by Sargent 1978) uses aggregate employment data to model how firms' expectations about product demand affect paths of employment. These allow the researcher to distinguish adjustment costs from changes in expectations; and their use of monthly or quarterly data does provide the appropriate temporal disaggregation. However, because the data cover all manufacturing employment, they suffer from excessive spatial aggregation while at the same time perhaps being unrepresentative of the entire economy.

These are not criticisms of the intellectual value of these series of studies. We do now understand more about how to model factor adjustment and how to extricate lags arising from adjustment costs from those produced by shifts in expectations. What these studies have not done is tell us much about the nature of adjustment at the plant level for the typical plant, since nonlinear adjustment may mean there is no "typical" plant, or much about the true structure of adjustment in the aggregate. Because many of them use annual data, they cannot inform us about the path of the response to exogenous demand shocks.

None of the empirical models estimated in either strand of this literature makes a serious attempt to infer anything about the level or structure of the costs that face firms when they change employment. The assumption is usually made that adjustment costs can be approximated by a quadratic, which in turn generates standard linear decision rules that are easily modeled as distributed lags. Although she offers no formal modeling of adjustment costs, Houseman (1988) does estimate lag structures for employment-output relations in basic steel production using monthly time-series data for the United States, France, and Germany. The data are not ideal, as they do not allow estimating microrelationships, but they are much closer to the ideal than even two-digit SIC data. The monthly observations guarantee that there is no problem of temporal overaggregation. The only difficulties with the results are their lack of a theoretical basis, possibly severe structural changes that have occurred and the industry's possible failure to be representative. Slightly different problems are presented by Mairesse and Dormont (1985), who use data for the same three countries based on observations of a representative group of individual (manufacturing) plants. The difficulty here is that the observations are only annual, so that no serious attempt at inferring the size or structure of adjustment costs is possible.

Two other studies are based more soundly in micro theory, but each has problems of its own that prevent us from concluding that we know much about adjustment costs. Nickell (1979) estimates standard employment-output relations on quarterly time series covering U.K. manufacturing, but he does search for structural changes induced by changes in legislation that he believes have affected the costs of hiring and dismissing workers. The difficulty here is one of spatial overaggregation, perhaps coupled with too much temporal aggregation as well. Hamermesh (1989) uses monthly plant-level data. While the results do test explicitly for alternative structures of adjustment costs that generate different paths of employment demand, the coverage of the data may not be representative, and the time series are very short.

Recently there has been a recognition in empirical studies of labor demand that plants are not infinitely lived. Dunne, Roberts, and Samuelson (1989) have assembled and performed simple statistical analyses of a file of all manufacturing plants present in any of the five Censuses of Manufactures between 1963 and 1982. The data provide the most detailed available picture of the totals of gross flows of plants into and out of existence and of the concomitant flows of employment opportunities. This is a gargantuan and praiseworthy undertaking. Nonetheless, we should recognize that what they have achieved is still not up to a level that will provide the basis for analyzing the determinants of plant closings and openings at the microlevel. The plants are observed only quinquennially. Decomposing employment changes over an observation period this long induces positive and negative biases in the estimated fraction of net employment change that is accounted for by births and deaths of plants: positive, because month-to-month or even yearly fluctuations in net employment are missed; negative, because short-lived establishments' births and deaths go unnoticed. Still more important is that this very high degree of temporal aggregation prevents one from inferring anything about the short- or even intermediate-run causes of employment change. (The absence of output or factor-price data for each plant also renders this impossible.) The restriction to manufacturing makes the data increasingly unrepresentative.

Jacobson (1986) and Leonard (1987) assembled similar sets of longitudinal data that have the advantages of covering all private nonfarm establishments and of being available annually. This mitigates some of the problems of temporal overaggregation. (Spatial aggregation is obviously not a problem.) The only difficulties are that the data are available starting only recently; each data set covers only one state, Pennsylvania and Wisconsin, respectively; and output data are not available (though payroll and, by calculation, earnings data are in Jacobson's data). Thus far these data have also been used only to decompose aggregate net changes in employment into births, deaths, and expansions/contractions of plants.

The interesting questions in factor substitution have to do with the effects of imposed changes in factor prices on the quantity of labor employed and with the effects of changes in the supply of labor on wage rates. These questions are of interest at the aggregate level and for various disaggregations of the work force. In the former case the crucial issue is the aggregate elasticity of labor demand; in the latter case it is one of substitution among workers of different types. In both cases, though, the question can be discussed by analyzing how firms' employment of different groups of workers responds to exogenous changes in their wage rates.

Research on labor-demand elasticities and labor-labor substitution can be divided into pure time-series studies and cross-section or pooled time-series, cross-section varieties. The former group consists mostly of analyses of annual aggregate data in which labor is treated as homogeneous or is disaggregated into production and nonproduction workers. (Berndt and Christensen 1974 was the first, McElroy 1987 the most recent strand of this literature.) As noted above, the underlying data suffer from problems of representativeness. Though their high degree of temporal aggregation is not a severe problem for measuring labor-demand elasticities, the aggregation of all workers into at most two groups limits their applicability to questions of labor-labor substitution. Their excessive spatial aggregation poses especially severe problems. The relationships that are estimated involve nonlinear transformations of the underlying data. There is no reason to assume that the aggregation of these relations for the underlying establishments would produce the same estimates as the aggregate data. Without simulation studies of the effects of aggregation of the establishment data (ignoring issues of aggregation of labor into one or two homogeneous groups), we cannot be sure how much is learned about the essentially microeconomic question being asked. Similar problems exist in the vast set of time-series analyses based on data aggregated over all establishments within an industry (see Hamermesh 1986).

Recently there have been a few efforts to answer questions about factor substitution using pooled time-series, cross-section data based on establishments. Barbour, Berger, and Garen (1987) examine four years of quarterly observations on nearly 1,000 coal mines in Kentucky; Hart and Wilson (1988) use five annual observations on nearly 50 metal-working establishments in the United Kingdom. Both data sets allow the authors to infer labor-demand elasticities (for homogeneous labor) at the appropriate levels of spatial and temporal disaggregation. The only difficulties with the data used in these commendable studies are that they are clearly unrepresentative of anything other than their particular industries and locales, and their coverage of very short time periods makes it unlikely that they capture (average out) any short-term fluctuations in the b, the parameters describing the structure of the underlying economic relationship.

The estimation of labor-demand elasticities using cross-section data has been a growth industry since the late 1970s. Unfortunately, these studies, all of which estimate flexible approximations to cost or production functions, have been based on data from widely available household surveys rather than on establishment data. Thus some of the work (e.g., Grant and Hamermesh 1981; Grossman 1982) uses Census of Population data aggregated to the standard metropolitan statistical area (SMSA) level and linked to data on the capital stock by SMSA. Another (Berger 1983) uses time series of CPS data for states in a similar manner.

The main problem in this set of studies is the general inappropriateness of household data for the purpose of estimating demand relationships, basically a problem of potentially severe unrepresentativeness. Essentially each worker in the household survey represents the establishment that employs him or her; many plants have none, while others have several representatives in the survey. There is no reason to expect biases due to this unusual sampling procedure, but it is hardly designed to minimize sampling error in data used to describe the behavior of plants. The spatial aggregation (to the SMSA level) is also excessive, due to the nonlinearity of the relationships that we noted above in discussing time-series estimates on aggregated data. The only virtue of these studies is that they do allow the authors to draw inferences about substitution among groups of workers, as they disaggregate the work force into a substantial number of potentially interesting categories.

An alternative approach, exemplified by Borjas (1986), avoids the spatial aggregation in the studies cited above by using individuals' wages from the Census of Population as dependent variables in a generalized Leontief model of production. The major problem with the other studies is not obviated here, though: the household data used are representative of employers' demand for the individuals, but they are likely to have very large errors in their role of measuring firms' behavior.

Two studies (Sosin and Fairchild 1984; and Allen 1986) use plant-level data to estimate production relationships. Those studies satisfy all the criteria of appropriate spatial disaggregation required to estimate the relevant parameters, assuming we believe the cross-section data reflect equilibria. (Temporal disaggregation is not a major problem in this area.) The data are representative of the structures (several industries in Latin America, and school and office building construction in the United States, both in the early 1970s), though clearly not of all industry or of other economies. These studies should be models of how appropriate data can be assembled and used to estimate parameters describing a particular production technology.

Unlike the study of labor supply, which was rejuvenated by the development of large longitudinal household surveys, no similar advance has occurred in data on establishments that might produce a renascence in the study of labor demand. We do have the Longitudinal Research Database (LRD) on establishments, an annual establishment-based file constructed from the same sources that generate the published data in the *Annual Survey of Manufactures*. Though this set of data does overcome problems of spatial aggregation, annual observations are too infrequent to capture many of the labor-demand phenomena of interest. Also, the restriction to manufacturing plants means the data suffer from serious problems of unrepresentativeness. Also available are the Employment Opportunity Pilot Project (EOPP) data, a panel with two observations on each of a large number of establishments in 28 sites. The difficulties with this set of data are that it is no longer an on-going data collection effort, the sites are not representative, and only limited information is provided on sales in the participating establishments.

What is needed is a quarterly, or even better, a monthly longitudinal survey of an appropriately stratified sample of establishments that is representative of all private nonfarm business. This survey should be establishment based, should replace defunct establishments with appropriate substitutes, and should be benchmarked at regular intervals to available censuses of business, manufactures, mining, and so on. Given the frequency of observations that is required, only a small sample is feasible, but with careful sampling this can be reasonably representative. The survey should contain data on total employment and on employment disaggregated into several meaningful skill categories, on hours worked by each group of workers, on the payroll for each type of worker, on other labor costs (to allow the very much needed study of the effects of nonwage costs on labor demand) and on total sales and production. These latter two series are especially important if the empirical study of labor demand at the appropriate microlevel is to have any basis in microeconomic theory.

The data collection effort I am proposing is mostly an extension and rationalization of what already exists. The monthly BLS-790 data that form the published series on disaggregated weekly earnings, hours, establishment-based employment, and so on, cover a much larger sample than is needed for this proposed survey. The OSHA sampling frame is similar and has the virtue of mandatory reporting requirements that the BLS-790 data set lacks. What the survey requires is expanding one of these or some other existing sampling frame, requiring mandatory sampling if possible, to obtain information on nonwage labor costs and output/sales, to include a more meaningful disaggregation of skills, to reduce the sample size tremendously while enhancing its representativeness, and to develop a means of building an appropriately constructed longitudinal format in which to handle the data files. The additional data to be collected—nonwage labor costs—are already collected through the mechanisms that produce the employment cost index (ECI). It should be possible to use the procedures that generate the inputs into the ECI in constructing the proposed longitudinal file. The only new information is that on output and sales, and the new skill classifications, and, at least for manufacturing establishments, all but the skill classifications are already collected on an annual basis.

The collection of data on a quarterly or monthly basis would enable us to characterize adjustment paths more satisfactorily. Its basis in establishments would also provide most of the information that would allow us to obviate the more important difficulties with the data underlying studies of labor-labor substitution. The only additional requirement on this proposed data set is that employment in each establishment (and hours and payroll too) be disaggregated by various cuts of the labor force. At the very least, disaggregation by sex, race, three age groups, and several skill categories would be useful in answering questions most relevant to issues of policy involving the distribution of job creation and the effects on wages of changes in relative supplies of workers in different demographic groups.

Strides in constructing complex models for inferring the nature of error structures in factor adjustment and the nature of the technology of factor substitution have neither been matched nor motivated by similar strides in the collection of data appropriate to the estimation. We are piling more theoretical and econometric structure upon the same sets of unsuitable data. Until we create the kind of data set outlined here, the situation is likely to become worse.

9.5 Labor-Market Studies—Can the Past Be Recaptured?

During the late 1940s two major studies of local labor markets were undertaken, Reynolds (1951) for New Haven, and Myers and Shultz (1951) for Nashua, New Hampshire. (Segal 1986 presents an excellent discussion and evaluation of these studies' lasting importance.) While not the first studies of entire labor markets, these did carry the genre to its peak. Similar work, which advanced the literature by using more complex methods of analyzing the data, was carried out by Rees and Shultz (1970) in the early 1960s for the Chicago labor market. The general approach of these studies was to combine household and establishment data. In each case questions were asked of employers in a number of plants and of substantial numbers of workers in those plants. In essence the studies can be described as cross-section combined establishment-household surveys. This combined approach has not been repeated.[3] We have been relegated to using increasingly complex sets of household data and fairly paltry sets of establishment data.

Is their absence a loss? That is, did we learn anything from the labor-market studies, and are there questions of interest today that could be answered better if we had data like those collected in the labor-market studies? One set of analyses that was novel at the time was of the role of spatial differences in wage rates among workers with similar characteristics in identical jobs at different plants (stressed especially by Rees and Shultz 1970). To some extent this information is now duplicated with journey-to-work information from the Censuses of Population, though the level of occupational detail is not so great as in the labor-market studies. Nonetheless, the studies were the first to stress the importance of distance and the relative locations of workers and jobs in producing large wage differences among otherwise identical workers.

The labor-market studies were ahead of their time in their focus on how job vacancies are filled, on how workers search for jobs, and on wage structures. It is true, particularly in the two studies from the late 1940s, that much of the research is based on the attitudinal questions that we economists abhor. In those same studies, though, there is much discussion of the role of unemployment insurance benefits in job search; of reports on how workers acquire knowledge about alternative jobs when they become unemployed; and of the nature of jobs, including trade-offs between wages and job characteristics, that affect workers' search behavior. Even the best empirical studies of job search of the last 20 years based on large household data sets (e.g., Holzer 1987) could have proceeded better using data from the labor-market studies of the 1940s if the same theoretical issues had been posed and the statistical techniques now in use had been widespread then.

The labor-market studies meet most of the criteria in section 9.2 fairly well. Their particular distinction is the appropriateness of the level of aggregation—to individual firms and households. In some ways they fail on the criterion of representativeness, in that the labor markets studied are not representative of anything but themselves. In another sense, though, the data are quite representative: they provide the best possible way of describing the inherently market phenomena that the authors were trying to examine and that still interest us today. Obviously the industrial structures of the local labor markets have changed over forty years (consider Nashua [Myers and Shultz 1951] in particular). Whether this means that the labor-market facts that these studies demonstrated remain valid is unclear, but their approach to thinking about labor markets is surely still useful.

The kinds of data collected in the labor-market studies would provide much better answers to some of the questions about which labor economists are most concerned today. Consider first the notion of efficiency wages, the idea that there are substantial wage differentials that arise from firms' attempts to elicit effort from workers. Much of the "evidence" on this consists of demonstrating the existence of unexplained wage differentials in household data across narrowly defined industries (e.g., Krueger and Summers 1988). While the concept was not addressed in today's terms, the role of efficiency wages and unemployment as a discipline device was recognized in the labor-market studies, "The change from a balanced to a loose local labor market unquestionably brought with it a tightening up of plant discipline" (Myers and Shultz 1951, 144).

Analyzing the combined establishment-household data using today's techniques and concepts could shed far more light on the importance of efficiency wages. (Beginning efforts in this direction were made by Osberg, Apostle, and Clairmont 1986; and Groshen 1986.) For example, with wage data on individuals in the same specific occupations one could easily measure the importance of firm-specific effects. This would provide two substantial advances over current studies of wage differentials, in that it would allow us to examine wages within very detailed occupations at the level of individual establishments. Longitudinal data from labor-market studies would also allow one to examine how occupation-specific wage differences across plants affect turnover, a manifestation of worker dissatisfaction and the obverse side of the extra effort that efficiency wages are alleged to elicit.

The second area of current research is on the relative roles of job-matching and on-the-job training in producing observed patterns of wage growth with job tenure (see Abraham and Farber 1987). The question is whether wage growth results from firm-specific training or whether it just reflects sorting of workers so that more senior workers are those who have remained with the employers with whom they are well matched. The kinds of data produced in the old labor-market studies would not add much to this discussion because of their limitation to single cross sections. If such data were collected longitudinally, though, these questions could be answered as definitively as is possible in empirical work. With combined longitudinal establishment-household data we could follow workers in specific jobs as output and productivity vary in the plants where they work. That would enable us to observe more closely the effects of actual investments in training (if any) that are taking place and contributing to wage growth. Similarly, examining the detailed characteristics of job vacancies in relation to the characteristics of current and new workers would allow us to study the matching hypothesis directly rather than infer it from complicated modelling of the error structures of wage equations.

A revival of the kinds of data collection that underlay the labor-market studies would yield very high returns in instructing us about how labor markets function. One method is to replicate the early studies in specific labor markets using modern sampling techniques and collecting data that we now obtain in household and establishment surveys. An approach that will probably yield more information at lower cost would combine the longitudinal establishment survey proposed in section 9.4 with a linked survey of substantial samples of individuals employed in the establishments. This approach has the virtue of increasing spatial representativeness and providing the desired combined longitudinal establishment-household data. Still another method, though one that will not provide the monthly or quarterly data that are necessary for some purposes, is to use the establishment data underlying the Area Wage Surveys as a starting place for the construction of the kinds of data needed for the purposes of this section. This approach to data collection will also allow for the easy acquisition of detailed product- and labor-market characteristics that would be useful both for the specific topics discussed above and for other market-based issues.

9.6 Union Goals

For many years economists and industrial relations specialists have discussed what unions try to maximize. Developing the pioneering work of Dunlop (1944), economists have recently specified models designed to allow the estimation of the parameters of "union utility functions" on microdata. In simple models particular forms of these functions are combined with loglinear labor-demand equations to infer the parameters. More complex models test whether the union's marginal rate of substitution between employment and wages equals the slope of the labor-demand curve, or whether unions and employers move off the demand curve to a Pareto-superior point on the contract curve.

The main strand of research (Dertouzos and Pencavel 1981; Pencavel 1984; Brown and Ashenfelter 1986; MaCurdy and Pencavel 1986) is entirely based on pooled annual time-series, cross-section data describing wages and employment in locals of the International Typographical Union (ITU). The studies proceed from the simple labor-demand model to various tests of whether the bargain is constrained by demand. The second strand, Farber (1978) for the United States, and Carruth and Oswald (1985) for the United Kingdom, uses annual time series on mine workers to estimate the degree of relative risk aversion in union utility functions.

The authors are very aware of problems with specifying a single utility function for the union. Pencavel in particular argues that the ITU is well suited to finessing the problem of internal union decision-making, because (he argues) the workers are homogeneous and the union is very democratic. (Thus he implies that the median and average voters are identical.) No one would make these claims about miners' unions in the United States or the United Kingdom, so that one wonders whether the idea of estimating a union utility function makes sense for them. A similar problem exists with Eberts and Stone's (1986) cross section of teachers in New York state school districts.

The difficulty in all these studies, but particularly in the strand of work by Pencavel and his colleagues, is the limitation to what is essentially one small and remarkably atypical (see Lipset, Trow, and Coleman 1956) segment of the union sector. Here is a case where tremendous resources have been devoted to building and testing ever more complex models on what is essentially the same set of data. Assuming the model is relevant beyond the ITU, it is difficult to believe that additional effort at collecting a new set of data on another union would not add more to our understanding of what unions do than introducing yet more complexities to the basic model.

9.7 Is There a Need for Validation Using International Data?

In the discussion in section 9.2 I set out as the desideratum the acquisition of data that will provide the best estimates of the vector of structural parameters *b* describing the underlying economic relationship. Do these parameters describe behavior generally, or are we only concerned with characterizing agents' actions in one particular economy? If the former we must be especially careful to consider whether, even if our data meet all the criteria for appropriateness that I have laid out, the results they generate can be used to draw inferences that apply beyond this country's borders. The issue is basically one of representativeness of the data, except that too often we think that the universe we are trying to represent is the economy of the particular country where we reside. The obverse question involves the uses to which studies of other countries' labor markets can be put by American economists. These are basically (1) To provide additional laboratories for the estimation of parameters describing economic behavior generally and (2) to provide contrasts to our own labor market.

Whether such generalization is possible depends to a large extent on whether (1) there are sufficient similarities in consumers' tastes across countries that we should expect similar behavioral responses to various stimuli; (2) markets are sufficiently interconnected and technology diffuses sufficiently rapidly that competition eliminates much of the international differences in behavior that would otherwise arise; and (3) the institutions that regulate behavior are sufficiently similar so that the similar behavior inherent in economic agents is not altered by nonmarket forces. Since technology flows more freely across borders than does labor, these considerations suggest that generalizing about supply behavior from studies on data characterizing only one economy is likely to be more risky than drawing inferences about labor demand. Institutional differences do inhibit generalization; they also provide opportunities to predict the effects of altering domestic institutions and to obtain data that allow for independent replication of estimates of their impacts (assuming international differences in tastes and technology are not too great).

Killingsworth's (1983) monumental study of labor supply summarizes a vast array of research and (among other contributions) tries to determine the reasons for the disturbingly wide range of estimates of supply parameters. While different estimation techniques, data sets, and measurement difficulties undoubtedly contribute to the problem, one wonders how much of the range results from underlying differences among the different populations being sampled. Although, as I noted in section 9.4, the data are not very satisfactory, we have obtained a number of stylized facts about labor demand (see Hamermesh 1986). Given the sorry quality of the data, even the minimal knowledge we have obtained about labor-demand behavior generally would not be possible without the accretion of demand studies from several economies.

In the area of predicting the effects of institutional change American economists can learn much from studies of other economies. An excellent example is in inferring the effects of imposing comparable worth, where comparative studies (e.g., Gregory, Daly, and Ho, 1986) can tell us at least as much as generalizations based on the existing structure of the domestic labor market. In other cases our institutions are similar to those of other countries, but our federal system imposes such uniformity that it is difficult to have much confidence about estimates of labor-market effects. A particularly good example is the evaluation of the employment and labor-force effects of the federal minimum wage (Mincer 1976). A study for Canada (Swidinsky 1980), where provincial laws produce greater cross-section variation in effective minimum wages, substantially increases one's confidence in the results obtained for the United States.

The answer to the titular question of this section—whether we should validate our work on data from countries beyond the United States—is a resounding yes. We will never be able to make universally applicable statements about all aspects of labor-market behavior; but with more attention to studies that use data from countries other than the United States, we will at least avoid the embarrassing ethnocentricity that often characterizes our attempt to generalize empirical results. At the same time, such attention will improve our understanding of the domestic labor market.

9.8 What Is to Be Done?[4]

Doing applied economics properly is an art—and the data used in practicing this art must meet the criteria of appropriate aggregation, representativeness and current structure. Too often we empirically oriented labor economists have the lazy person's habit of tailoring our methods of analysis, and sometimes even the basic questions we ask, to fit the available data. In the case of analyzing labor supply, where the available data are representative, offer the appropriate degree of disaggregation and capture current structures well, this is an excellent approach. In other cases it is not. Studies of labor-demand phenomena and of the interaction of supply and demand in the labor market have been based on data that are often inappropriately disaggregated, unrepresentative, or uncharacteristic of current structures. Indeed, the tremendous resources devoted to collecting data that are best suited for analyzing labor supply, and the consequent availability of those data, have reduced incentives to collect data that are more suitable for these other questions.

This is not a condemnation of recent empirical research on issues other than labor supply. We have learned a lot; but what we can possibly learn about these issues is severely limited by the lack of appropriate data. Rather than rely on inappropriate data, those of us interested in empirical research in labor economics outside the narrow and decreasingly fertile area of labor supply must adopt some of the sociologists' willingness to generate new sets of data (though, one would hope, without abandoning our willingness to construct models to organize the analysis of those data). Also, given the limited resources available for collecting data, we must urge public officials responsible for funding data collection to get out of the rut of concentrating on ever-larger and ever-longer sets of household data and redirect resources toward the kinds of data that are more likely to yield new basic insights into the operation of labor markets. The individual data-collection efforts implied by such a redirection of public and private activities cannot take place without expending large amounts of time and money. If coupled with some curtailing of the increasing tendency to spend energy and budget resources on accumulating additional longer household-based longitudinal studies, they need not add to the share of public resources devoted to the collection of data in labor-related areas.

The major area toward which resources should be shifted is the collection of longitudinal, monthly or quarterly, establishment data to which household data on workers in the sampled establishments are linked. This data set should contain the information now collected by the BLS in its immense monthly surveys of establishments as well as information on output and sales. The sample of establishments need not be large, but it must be representative of the entire economy, not merely the over-studied manufacturing sector. Simultaneous sampling of panels of workers in these establishments that provides information like that now available in the NLS and the PSID, or even in the CPS supplements, should also be undertaken. In ten years we would thus have in hand at little extra cost a tool that would allow us to understand increasingly important phenomena that have been heretofore either relatively neglected or studied using inappropriate data.

Without the kind of endeavor proposed here the only progress possible in these areas of research and public policy will come through the continued efforts of individuals who collect small, usually unrepresentative, and incomplete sets of establishment-household data. This catch-as-catch-can approach has been and can continue to be important. It is unlikely to provide sufficient additional knowledge to save the study of labor economics from increasingly sterile empirical work using the existing massive sets of household data and from the growth of "labor theory" that is increasingly detached from the analysis of empirical phenomena.

Notes

1. The six journals are the *American Economic Review, Econometrica, Journal of Political Economy, Quarterly Journal of Economics, Review of Economics and Statistics,* and *International Economic Review.* I tried to follow the same sets of exclusions as Stafford (1986).

2. Johann Wolfgang Goethe, *Faust:* "To be sure [we] know a lot, but would like to know everything" (Pt. 1, scene 1).

3. The EOPP data set does combine establishments and households. However, employers are asked questions only about the characteristics of their most recent hire, so that very little is made of the combined nature of the data.

4. Apologies to V. I. Lenin.

References

Abraham, Katharine, and Henry Farber. 1987. Job Duration, Seniority and Earnings. *American Economic Review* 77:278–97.

Allen, Steve. 1986. Union Work Rules and Efficiency in the Building Trades. *Journal of Labor Economics* 4:212–42.

Barbour, James, Mark Berger, and John Garen. 1987. The Short-Run Demand for Labor in the Coal Industry. University of Kentucky. Typescript.

Berger, Mark. 1983. Changes in Labor Force Composition and Male Earnings: A Production Approach. *Journal of Human Resources* 17:177–96.

Berndt, Ernst, and Laurits Christensen. 1974. Testing for the Existence of a Consistent Aggregate Index of Labor Inputs. *American Economic Review* 64:391–404.

Berndt, Ernst, Melvyn Fuss, and Leonard Waverman. 1979. A Dynamic Model of Cost Adjustment and Interrelated Factor Demands. Discussion paper no. 79-30. University of British Columbia.

Borjas, George. 1986. The Demographic Determinants of the Demand for Black Labor. In *The Black Youth Employment Crisis,* ed. Richard Freeman and Harry Holzer. Chicago: University of Chicago Press.

Brown, James, and Orley Ashenfelter. 1986. Testing the Efficiency of Employment Contracts. *Journal of Political Economy* 94:S40–S87.

Cain, Glen, and Harold Watts. 1973. *Income Maintenance and Labor Supply.* Chicago: Markham.

Carruth, A. A., and A. J. Oswald. 1985. Miners' Wages in Post-war Britain: An Application of a Model of Trade Union Behavior. *Economic Journal* 95:1003–20.

Clark, Kim, and Lawrence Summers. 1982. The Dynamics of Youth Unemployment. In *The Youth Labor Market Problem,* ed. Richard Freeman and David Wise. Chicago: University of Chicago Press.

Dertouzos, James, and John Pencavel. 1981. Wage and Employment Determination under Trade Unionism: The International Typographical Union. *Journal of Political Economy* 89:1162–81.

Dunlop, John. 1944. *Wage Determination under Trade Unionism.* New York: Macmillan.

Dunne, Timothy, Mark Roberts, and Lawrence Samuelson. 1989. Plant Turnover and Gross Employment Flows. *Journal of Labor Economics* 7:48–71.

Eberts, Randall, and Joe Stone. 1986. On the Contract Curve: A Test of Alternative Models of Collective Bargaining. *Journal of Labor Economics* 4:66–81.

Farber, Henry. 1978. Individual Preferences and Union Wage Determination: The Case of the United Mine Workers. *Journal of Political Economy* 86:923–42.

Grant, James, and Daniel Hamermesh. 1981. Labor Market Competition among Youths, White Women and Others. *Review of Economics and Statistics* 63:354–60.

Gregory, Robert, A. Daly, and V. Ho. 1986. A Tale of Two Countries: Equal Pay for

Women in Australia and Britain. Discussion paper no. 147. Centre for Economic Policy Research, Australian National University.

Griliches, Zvi. 1986. Economic Data Issues. In *Handbook of Econometrics*, ed. Zvi Griliches and Michael Intriligator. Amsterdam: North-Holland.

Groshen, Erica. 1986. Sources of Wage Dispersion: How Much Do Employers Matter? Ph.D. diss., Harvard University.

Grossman, Jean. 1982. The Substitutability of Natives and Immigrants in Production. *Review of Economics and Statistics* 64 (1982): 596–603.

Hamermesh, Daniel. 1986. The Demand for Labor in the Long Run. In *Handbook of Labor Economics*, ed. Orley Ashenfelter and Richard Layard. Amsterdam: North-Holland.

———. 1988. The Demand for Workers and Hours and the Effects of Job Security Policies. In *Employment, Unemployment and Labor Utilization*, ed. Robert Hart. London: Unwin Hyman.

———. 1989. Labor Demand and the Structure of Adjustment Costs. *American Economic Review* 79:674–89.

Hart, Robert, and Nicholas Wilson. 1988. The Demand for Workers and Hours: Micro Evidence from the UK Metal Working Industry. In *Employment, Unemployment, and Labor Utilization*, ed. Robert Hart. London: Unwin Hyman.

Heckman, James, and Robert Willis. 1977. A Beta-logistic Model for the Analysis of Sequential Labor Force Participation by Married Women. *Journal of Political Economy* 85:27–58.

Holzer, Harry. 1987. Informal Job Search and Black Youth Unemployment. *American Economic Review* 77:446–52.

Houseman, Susan. 1988. Shorter Working Time and Job Security: Labor Adjustment in the European Steel Industry. In *Employment, Unemployment and Labor Utilization*, ed. Robert Hart. London: Unwin Hyman.

Jacobson, Louis. 1986. Job Creation and Destruction in Pennsylvania, 1975–85. W. E. Upjohn Institute, Kalamazoo, Mich. Typescript.

Killingsworth, Mark. 1983. *Labor Supply*. Cambridge: Cambridge University Press.

Krueger, Alan, and Lawrence Summers. 1988. Efficiency Wages and the Wage Structure. *Econometrica* 56:259–94.

Leonard, Jonathan. 1987. In the Wrong Place at the Wrong Time: The Extent of Frictional and Structural Unemployment. In *Unemployment and the Structure of Labor Markets*, ed. Kevin Lang and Jonathan Leonard. New York: Basil Blackwell.

Lipset, Seymour, Martin Trow, and James Coleman. 1956. *Union Democracy: The Internal Politics of the International Typographical Union*. Glencoe, Ill.: Free Press.

MaCurdy, Thomas. 1981. An Empirical Model of Labor Supply in a Life-Cycle Setting. *Journal of Political Economy* 89:1059–85.

MaCurdy, Thomas, and John Pencavel. 1986. Testing between Competing Models of Wage and Employment Determination in Unionized Markets. *Journal of Political Economy* 94:S3–S39.

Mairesse, Jacques, and Brigitte Dormont. 1985. Labor and Investment Demand at the Firm Level: A Comparison of French, German and U.S. Manufacturing, 1970–79. *European Economic Review* 28:201–31.

McElroy, Marjorie. 1987. Additive General Error Models for Production, Cost and Derived Demand or Share Systems. *Journal of Political Economy* 95:737–57.

Mincer, Jacob. 1962. Labor Force Participation of Married Women. In *Aspects of Labor Economics*, ed. H. Gregg Lewis. Princeton, N.J.: Princeton University Press.

———. 1976. Unemployment Effects of Minimum Wages. *Journal of Political Economy* 84:S87–S104.

Morrison, Catherine. 1986. Structural Models of Dynamic Factor Demands with Nonstatic Expectations. *International Economic Review* 27:365–86.

Myers, Charles, and George Shultz. 1951. *The Dynamics of a Labor Market.* New York: Prentice-Hall.
Nadiri, M. I., and Sherwin Rosen. 1969. Interrelated Factor Demand Functions. *American Economic Review* 59:457–71.
Nickell, Steven. 1979. Unemployment and the Structure of Labor Costs. *Journal of Monetary Economics* 11:187–222.
Osberg, Lars, Richard Apostle, and Don Clairmont. 1986. The Incidence and Duration of Individual Unemployment: Supply Side or Demand Side? *Cambridge Journal of Economics* 10:13–33.
Pencavel, John. 1984. The Trade-off between Wages and Employment in Trade Union Objectives. *Quarterly Journal of Economics* 99:215–32.
Rees, Albert, and George Shultz. 1970. *Workers and Wages in an Urban Labor Market.* Chicago: University of Chicago Press.
Reynolds, Lloyd. 1951. *The Structure of Labor Markets.* New York: Harper & Bros.
Sargent, Thomas. 1978. Estimation of Dynamic Labor Demand Schedules under Rational Expectations. *Journal of Political Economy* 86:1009–45.
Segal, Martin. 1986. Post-institutionalism in Labor Economics: The Forties and Fifties Revisited. *Industrial and Labor Relations Review* 39:388–403.
Sosin, Kim, and Loretta Fairchild. 1984. Nonhomotheticity and Technological Bias in Production. *Review of Economics and Statistics* 66:44–50.
Stafford, Frank. 1986. Forestalling the Demise of Empirical Economics: The Role of Microdata in Labor Economics Research. In *Handbook of Labor Economics,* ed. Orley Ashenfelter and Richard Layard. Amsterdam: North-Holland.
Swidinsky, Robert. 1980. Minimum Wages and Teenage Unemployment. *Canadian Journal of Economics* 13:158–71.

Comment Sherwin Rosen

Hamermesh's paper reflects on how new data can improve our understanding of several problems in labor economics. He reaches two main conclusions. First, highly disaggregated (monthly firm or even establishment) time-series data are necessary to make further progress on the demand side of labor markets. Second, labor-market surveys of the kind done by Reynolds, Rees, and Shultz in the 1950s and 1960s are the next logical step in the evolution of empirical labor economics. I disagree with the first conclusion, not because such data will be without interest, but rather because they will not solve the main difficulty with existing work. I agree with the thrust of the second conclusion (though not those particular survey instruments) because matched firm and worker samples hold promise for resolving empirical problems with existing theories and for advancing concepts and models to incorporate broader and more complex forms of behavior.

Modern labor economics has forged some of the closest connections between data and theory in all of applied economics. It is difficult to even imag-

Sherwin Rosen is the Edwin A. and Betty L. Bergman Professor of Economics at the University of Chicago and a research associate of the National Bureau of Economic Research.

ine the field without the existence of the one-in-one-thousand 1960 census tape and the early computers used to process those numbers. The gradual emergence of panel data expanded the range of questions that could be examined. These more demanding questions stimulated development of more sophisticated theories and statistical methods to deal with them. It is perhaps worth reminding ourselves from time to time that progress in any branch of empirical economics must involve such interactions between data and theory. Constructing models and organizing our thoughts about a problem is greatly influenced by available data and existing empirical regularities. A good model illuminates the main trends and facts these data reveal, but it is limited and must flounder at some level of empirical detail. Failures of this kind are, however, most interesting if they suggest how an expanded conception of the problem might remedy the situation and incorporate a broader class of phenomena.

Hamermesh points to the problem of labor supply as central to the field of labor economics and suggests that progress in this area has been so large that not much more can be expected from it. Now much has been learned about labor-supply behavior, but few would argue that we are anywhere close to full understanding. If there are diminishing returns, it is probably because the year-to-year changes in labor hours worked by individuals in panel data are poorly explained by current theories. The fact is that it has been very difficult to find wage and other pure, unilateral labor-supply price effects in the time-series behavior of individuals, at least at business cycles frequencies where recent interest in the subject came from. Is this because the data are inadequate for the theory or, as seems likely to me, is it the theory that needs overhauling to account for the data? Evidently more attention must be focused on the conceptual issues of joint decisions concerning hours and layoffs by both firms and workers.

The bulk of Hamermesh's paper is devoted to questions of labor demand. I am skeptical that monthly panels of establishments and firms will greatly advance our knowledge of this subject for the following reasons. First, is the unfortunate statistical constraint of measurement error that reduces the signal-to-noise ratio to intolerable proportions in very disaggregated data. For instance, problems of estimating meaningful capital stocks at the firm or establishment level are well known. And firms differ greatly in the qualities of their outputs and inputs. None of that fits neatly into our models but is more or less averaged away in more aggregate data.

Second, technological differences across firms and productivity change over time are very important for understanding detailed micro behavior, but are basically nuisance issues from the point of view of labor-demand elasticity questions. Their immense importance at this level of disaggregation reduces one's ability to study the primary question. To be sure, technical change affects factor demand studies at any level of aggregation, but industry-level data renders this problem less important. And more pragmatically, the great hope for disaggregation expressed by many economists in the 1950s and 1960s has not exactly been fulfilled in the intervening years. That situation is not going to change any time soon.

Third, there are compelling conceptual reasons to be wary of highly disaggregated firm-level data for these questions. The most useful concept of labor demand by far is the *market or industry demand.* That is why the standard theory is built upon constant-returns-to-scale production functions at the industry (aggregate) level. This is a wonderful finesse of such complicated issues as the size distributions of firms, the extensive margin, and firm-specific rents that do not have to be explicitly considered for a market concept of factor demand. Furthermore, looking only at the firm ignores scale effects—the tendency for the industry to contract or expand through output demand efforts in response to variations in output prices provoked by factor price changes. This important part of the problem is best studied at the industry level even if data are perfect.

Fourth, I doubt if monthly establishment data will contribute much toward understanding the dynamics of factor demand. Monthly data are ideal for studying seasonal variations, and those are important for some industries such as fishing and construction. But what can they tell us about the longer term movements associated with business cycles that motivate much of the interest in this subject? Quarterly data are better for that purpose. However, even quarterly data do not show much sensitivity of factor demand to quarter-to-quarter movements in factor prices, as Nadiri and I showed in an early NBER study.[1] While methodological improvements in our methods have been proposed, no convincing explanation for this finding has been found. So if substitution effects cannot be found in quarterly data, how in the world will seasonally dominated monthly data improve the situation?

Neoclassical theory of labor supply and demand works quite well on a longer time frame and in cross sections. However, the empirical finding that wages do not matter much for explaining quarter-to-quarter variations in either supply or demand calls for both new data and some new conceptions of these problems. Many of us think that the conception somehow lies in the economics of contracts suggested by the long-term relations observed between workers and employers. This makes me very enthusiastic about Hamermesh's proposal for renewed interest in matched worker-firm surveys because having matched data is the only way in which the marriage aspects of labor-market exchange can be thoroughly studied. This is bound to yield many interesting results whether or not the economic theory of contracts proves useful. Existing matched records from different administrative sources (e.g., SSA, CPS, and IRS) do not provide nearly enough information for this.

Two important barriers must be overcome in this endeavor. First, we must

1. M. Ishaq Nadiri and Sherwin Rosen, *A Disequilibrium Model of Demand for Factors of Production,* General Series no. 99. (New York: National Bureau of Economic Research, 1973.)

convince decision makers at the appropriate statistical agencies that the effort will be worthwhile. This must include mechanisms for preserving confidentiality of firm records. Second, serious thought and resources must be put into designing the survey instrument, pretesting it, and implementing it statistically. The fact is that economists such as Hamermesh and myself who are interested in using such data are not well equipped to carry out these important production matters. However, if we keep waiting for someone else to do all of that hard work we are likely to be waiting for an awfully long time.