Supply and Demand in the Labor Market*

David Card
Princeton University

November 1987

Supply and Demand in the Labor Market

ABSTRACT

This paper presents a survey of recent microeconometric studies of the labor market, focusing on research that emphasizes the possible failure of measured wage rates to separate individual supply and demand decisions. On both the demand and supply sides of the labor market there is evidence that forces from the other side of the market influence employment outcomes through some mechanism other than the wage. On the supply side, this evidence takes the form of correlations between individual labor supply outcomes and market-level measures of employment demand in the individual's local labor market. On the demand side, it takes the form of correlations between firms' employment decisions and measures of their employees' outside opportunities. Both sets of findings are inconsistent with simple supply and demand models, and suggest the need for alternative models of the labor market, which permit an uncoupling of short-run employment decisions from wage rates.

David Card
Department of Economics
Princeton University
Princeton, NJ 08544
609-452-4045
Supply and Demand in the Labor Market

In a conventional model of supply and demand, market participants can obtain all the information they need about the other side of the market from the equilibrium price.¹ This proposition lies behind the widespread practice of analyzing supply and demand behavior separately from each other. In applied studies of the labor market, the notion that wage rates separate the two sides of the market is so pervasive that it has even lead to instances where the same correlations have been interpreted as supply parameters by one author, and as demand parameters by another.²

Recently, however, several strands of empirical research into supply and demand behavior in the labor market have attempted to relax this assumption. On the demand side, research into the implications of efficient wage bargaining stresses that the negotiated wage rate is an incomplete measure of the opportunity cost of labor.³ On the supply side, studies of the effects of hours constraints⁴ and the magnitude of compensating differentials for unemployment risk⁵ stress that knowledge of demand conditions is necessary to fully specify the employment outcomes of individual employees. In a similar vein, research on the covariance structure of individual earnings and hours attempts to test the notion that changes in hours are associated with changes in the marginal wage rate.⁶

This paper presents a partial survey of recent microeconometric studies of the labor market, focusing on those studies that emphasize the possible failure of measured wage rates to separate individual
supply and demand decisions. Although studies on the supply and demand sides have involved somewhat different theoretical models, and very different data sets, I believe that they are largely motivated by the same phenomenon: namely, the absence of large or systematic correlations between measured wage rates and employment outcomes. This absence has lead to some skepticism about the relevance of simple supply and demand models for the interpretation of micro-level employment and wage data. The development of models that relax the separation of the supply and demand sides of the market is one possible route toward a more complete understanding of the functioning of the labor market.

I. Studies on the Supply Side

It is far beyond the scope of this paper to survey the recent literature on labor supply. The study of individual labor supply is one of the most active areas of research in labor economics, and there are several recent and extremely thorough surveys: in particular Pencavel (1986), Killingsworth and Heckman (1986), and the book by Killingsworth (1983). Here I wish mainly to draw attention to research that explicitly incorporates the notion that individual supply decisions may be influenced by demand-side factors via some route other than the marginal wage rate. Before reviewing this research, however, it may be useful to lay the background for this recent work by describing what I believe to be the main obstacle to a straightforward application of conventional supply and demand apparatus to individual labor supply data.

The principle difficulty is this: although the time-series variability in individual labor supply is large, very little of this
variability is correlated with changes in measured wage rates, or any
other variable readily available in conventional longitudinal surveys.
Some evidence on the variability of annual hours changes among adult
males is presented in Table 1. The standard deviation of the percentage
change in reported annual hours in two major panel data sets (the Panel
Study of Income Dynamics or PSID, and the National Longitudinal Survey
of Older Men or NLS) is about 35 percentage points. This number is
somewhat smaller among men with the same employer in consecutive years
(.31), and is considerably smaller among those who report no change in
employer over an extended period (.20). The effect of measured
covariates on this variability is negligible. Altonji and Paxson (1986)
report the cross-section variability of the change in the logarithm of
annual hours, regression-adjusted for age, race, change in health
status, change in marital status, change in family income, and changes
in wage rates. The variability of their regression-adjusted changes
in annual hours, reported in Table 1 as "adjusted" estimates, are very
similar to the unadjusted estimates, both for individuals with the same
employer in consecutive years, and for those who change employers.

Of course, these estimates of the variability in individual hours
may be over-stated by the presence of survey measurement error. In
fact, there is now a variety of evidence that survey responses to
questions on earnings and hours contain substantial measurement
error. One direct estimate of the measurement error in annual hours
changes is presented by Duncan and Hill (1986), who analyze the dif-
ference between employer-recorded earnings and hours measures, and
employee responses recorded by the PSID survey instrument. Assuming
that the employer record is accurate, they estimate that the signal-to-
noise variance ratio in the survey measure of the change in annual hours
is 1.22 (Duncan and Hill (1985), Table 3). Applying this estimate to
the estimated variances in Table 1, the estimated standard deviation of
true hours changes for adult males falls by about one-third, from
.33-.35 to .24-.26.

A less direct but similar estimate of the extent of measurement
error in hours changes can be obtained from a simple model of the
covariance structure of hours. Suppose that the logarithm of reported
hours \((h_t)\) consist of actual hours \((h^*_t)\) and measurement error \((\epsilon_t)\):

\[
\log h_t = \log h^*_t + \epsilon_t .
\]

Suppose further that measurement error is serially uncorrelated with
constant variance \(\sigma^2_\epsilon\). Then the variance of the change in log hours is

\[
\text{var}(\Delta \log h_t) = \text{var}(\Delta \log h^*_t) + 2\sigma^2_\epsilon .
\]

while the covariance of consecutive changes in hours is

\[
\text{cov}(\Delta \log h_t, \Delta \log h_{t-1}) = \text{cov}(\log \Delta h^*_t, \Delta \log h^*_t-1) - \sigma^2_\epsilon .
\]

Abovd and Card (1987, Table 3) report that the average covariances of
consecutive hours changes in their PSID and NLS samples are -.035 and
-.038, respectively. If the first-order autocorrelation of \(\Delta h^*_t\) is
known, these estimates can be combined with the variance estimates in
Table 1 to yield an estimate of the signal-to-noise variance ratio in
the change in hours. Assuming that the first-order serial correlation
coefficient of \(\Delta h^*_t\) is between 0 and .3, the implied estimate of the
standard deviation of the true variability in individual hours is .17-.22.\textsuperscript{13/}

These adjusted estimates imply that there is significant year-to-year variability in individual work hours. For example, if the change in the logarithm of annual hours is normally distributed, a standard deviation of 22 percentage points implies that each year, one-half of adult males experience a change in their annual hours in excess of 15 percent.\textsuperscript{14/} Even among individuals who report the same employer (on their main job) in 11 consecutive years in the PSID sample, the standard deviation of annual hours changes (adjusting for measurement error) may be as high as .15, implying under the normality assumption that each year one half of these individuals experience a change in annual hours in excess of five weeks of full-time employment.\textsuperscript{15/}

a. The Covariance Structure of Earnings and Hours

Most of the analysis of year-to-year variation in individual labor supply has been conducted in the framework of an intertemporal labor supply model (see in particular MacCurdy (1981) and Pencavel (1986), Section 3.5)). Following the derivation in MacCurdy (1981), if preferences are additively separable in annual consumption and leisure, and individuals have perfect foresight, then (to a log-linear approximation) the logarithm of individual labor supply in year $t$ can be written as

$$\log h_t = \psi + \eta \log w_t + \gamma t + \upsilon_t,$$

where $\psi$ is an individual-specific constant reflecting lifetime wealth and any permanent component of tastes for consumption and leisure, $w_t$ is the individual wage rate (assumed to be constant over the year), $\gamma$
is a constant reflecting the ratio of the rate of time preference to the market rate of interest, and $v_t$ represents any transitory taste component in current labor supply. The parameter $\eta$ is the intertemporal substitution elasticity, and represents the responsiveness of annual hours to anticipated changes in hourly wage rates, holding constant the marginal utility of wealth.\textsuperscript{16}

If the logarithm of labor earnings, $\log g_t$, is defined as the sum of the logarithms of annual hours and average hourly wage rates, then the lifecycle labor supply model implies that

$$\log g_t = \psi + (1-\eta) \log w_t + \gamma_t + v_t.$$ 

Appending a pair of survey measurement errors to hours and earnings and taking first differences, the lifecycle model generates a components-of-variance structure for changes in annual hours and earnings of the form:

$$\Delta \log h_t = \gamma + \eta \Delta \log w_t + \Delta v_t + \Delta \epsilon_t,$$

$$\Delta \log g_t = \gamma + (1-\eta) \Delta \log w_t + \Delta v_t + \Delta \nu_t,$$

where $\epsilon_t$ and $v_t$ are measurement errors in the logarithms of observed annual hours and earnings, respectively.\textsuperscript{17} It is interesting to observe that if $\eta$ is in the range of .1 to .3, as estimates in MacCurdy (1981), Altonji (1986), and Ham (1986) suggest, then the lifecycle model implies that cross-sectional variation in wage growth contributes 20-120 times more to the variance of year-to-year changes in earnings than to the variance of year-to-year changes in hours.

In the absence of restrictions on the covariance properties of the other variance components in the model, however, the covariances and cross-covariances of earnings and hours do not supply enough information
to estimate the contribution of wage growth variation to the cross-sectional variances of earnings and hours. Abowd and Card (1986) tackle this problem by assuming that measurement errors are i.i.d. In three different panel data sets (PSID, NLS, and the control group from the Seattle-Denver Income Maintenance Experiment) they find that the remaining common component of earnings and hours variation affects the two variables proportionately. In other words, apart from i.i.d. measurement error, earnings and hours covary at fixed real wage rates. This suggests that the component-of-variance associated with wage growth variation, which according to the lifecycle model affects earnings more than hours, is dominated by random preference shocks.  

An alternative interpretation, however, is that the common component of variation in earnings and hours represents variation in the demand for labor at fixed wage rates. For example, if most of the year-to-year variation in measured labor supply is generated by short-term layoffs at fixed wage rates, then one would expect the common component of variance to affect earnings and hours proportionately. Unfortunately, using only the variances and covariances of earnings and hours, one cannot distinguish between a labor supply model with substantial year-to-year preference variation, and an alternative model in which demand affects individual labor supply directly, with no change in marginal wage rates. The covariance structure of earnings and hours does suggest, however, that behavioral responses to variation in marginal wage rates, which are the focus of the life-cycle theory, are a relatively minor feature of longitudinal labor supply data.
b. Constraints on Labor Supply

If individual labor supply is determined by employer demand at constant wage rates, then individuals may be observed "off" their labor supply functions at a point in time. In situations of low demand, individuals may report that they are unemployed, while in situations of high demand, they may report that they are overemployed. This simple idea suggests that it may be appropriate to interpete the sum of employed and unemployed hours as "labor supply" at a point in time. Ashenfelter (1978) suggested that a conventional labor supply function be augmented with measured unemployment in an equation of the form

\[ \log h_t = h(w_t, x_t) - \theta u_t, \]

where \( h \) represents the labor supply function, \( w_t \) represents the wage rate, \( x_t \) represents a vector of covariates and stochastic terms, \( u_t \) represents measured unemployment and \( \theta \) represents the ratio of constrained or "involuntary" unemployment to measured unemployment.

Evidence that \( \theta \) is nonzero may be interpreted as evidence in favor of the notion that labor supply is "demand determined".

This interpretation of equation (1) has been sharply criticized by Pencavel (1986, pp. 42-44), and Heckman and Macurdy (1987). To see the basis of their criticism, consider the division of time within a year into time allocated to sickness, vacation and time-off, unemployment, and employment. As before, let \( h_t \) and \( u_t \) represent time employed and unemployed, respectively, and let \( l_t \) and \( s_t \) represent leisure time and sickness time. Then, as an accounting identity,

\[ h_t = T - u_t - l_t - s_t. \]
where $T$ is a measure of total time available. Ignoring sickness, cross-sectional or time series variation in employment is completely offset by variation in the sum of unemployment and leisure. In the absence of an hypothesis about how survey respondents divide nonmarket time between leisure and unemployment, the interpretation of the coefficient $\theta$ in equation (1) is clearly ambiguous.

The standard labor supply model makes no distinction between unemployment and leisure. In their highly influential 1969 study, however, Lucas and Rapping suggested one possible interpretation of survey unemployment: reported unemployment is just some fraction of total nonmarket time (net of time lost to illness):

\[(3)\quad u_t = \alpha(T - h_t - s_t) = \alpha T - \alpha h_t - \alpha s_t.\]

According to this interpretation, observed unemployment time is just the mirror image of employment time, and provides no independent sample information on behavior. Evidence that $\theta \neq 0$ in equation (1) can therefore be interpreted as evidence of misspecification of the labor supply function $h$.

The model behind the test proposed by Ashenfelter (1978, 1980), Ashenfelter and Ham (1979), and Ham (1982, 1986), on the other hand, assumes that reported unemployment includes information on hours constraints faced by workers. A simple model (ignoring hours lost due to sickness) is

\[(4a)\quad h_t = h(w_t, x_t) - c_t,
(4b)\quad u_t = \alpha(T - h(w_t, x_t)) + c_t,
(4c)\quad l_t = (1 - \alpha) (T - H(w_t, x_t)),\]
where \( c_t \) is a measure of the hours constraint facing the individual in period \( t \). According to this model, the finding of a nonzero \( \theta \) in equation (1) can be interpreted as evidence of constraint, maintaining a correct specification for the labor supply function.

Estimates of \( \theta \) alone obviously cannot distinguish the equilibrium model of unemployment (3) from the "disequilibrium" model (4a) - (4c). In my opinion, however, the question of whether measured unemployment belongs on the right-hand side of a labor supply equation misses the main point, which is that demand side information is required to correctly specify observations on the supply side. On this question, the recent paper by Han (1986) sheds some interesting light. Consider an augmented labor supply equation of the form

\[
\log h_t = h(w_t, x_t) + \gamma a_t
\]

where \( a_t \) includes information on aggregated conditions in the individual's occupation, industry, and local labor market in time period \( t \). There is no connection between the measurement of \( h_t \) and the measurement of \( a_t \). If the labor supply model is correctly specified, however, then \( \gamma = 0 \), since market-level information is irrelevant to the individual supply decision, controlling for wages (and individual characteristics). One may object to this test on two grounds. First, individuals may sort themselves into industries, occupations, or local labor markets on the basis of tastes for leisure, so that individual tastes may be correlated with market-level characteristics.\(^{20/}\)

This objection can be overcome by differencing:

\[
\Delta \log h_t = \Delta h(w_t, x_t) + \gamma \Delta a_t
\]
provided that tastes don't change too quickly. Second, recent labor
market changes may signal new information on lifetime opportunities, and
hence induces labor supply changes even holding constant the wage.
This objection can be overcome by using lagged changes in local labor
market conditions to predict $\Delta \alpha_t$:

$$\Delta \log h_t = \Delta h(w_t, x_t) + \gamma^* \Delta \alpha_{t-1},$$

where $\gamma^*$ now represents the product of $\gamma$ and the first-order corre-
lation coefficient of $\Delta \alpha_t$ (which is assumed to be nonzero). Although
he does not report estimates of this equation in his 1986 paper, Ham's
results using $\Delta \alpha_{t-1}$ as instrumental variables for $\Delta \alpha_t$ in a first-
differenced version of equation (1) imply that $\gamma^*$ is far different
from zero. More direct evidence from an earlier unpublished version of
this paper (Ham (1984)) is summarized in Table 2.21/ This table pre-
sents the results of exclusion test for first-differences and lagged
first-differences of four measures of "local" labor market demand con-
ditions in a simple log-linear labor supply equation. The results in
the first row of the table confirm that measures of state, occupational,
and industry unemployment rates help predict changes in individual
hours. The results in the second row, however, suggest that the
influence of local labor market conditions is diminished once individual
unemployment is included on the right-hand side of the labor supply
equation (as might be the case, for instance, if the model of equations
(4a) - (4c) were correct). This evidence confirms that demand-side
information (even based on one-year-ahead forecasts) significantly
improves the predictions of individual labor supply outcomes, and is
clearly at variance with the conventional labor supply model.

c. Compensating Differences for Employer-Determined Hours of Work

Even if hours are determined in the short-run by employer demand, with little or no commensurate wage adjustment, long run mobility of workers between employers (or industries or occupations) will tend to equalize the expected utility of alternative employment packages. This notion of labor market equilibrium is clearly elucidated in The Wealth of Nations, and has motivated the recent study of the relation between wage rates and the distribution of employment outcomes. For the most part, these studies assume that differences across industries, occupations, or employers in the distribution of annual work hours reflect systematic differences in the distribution of demand conditions. Evidence of compensating differentials for these systematic characteristics therefore provides indirect evidence in favor of the hypothesis that ex post employment outcomes involve employer-determined constraints on individual supply behavior.

It is worth pointing out that there is, in fact, considerable diversity across industries, occupations, and employers in the distribution of annual hours. Some illustrative evidence of the diversity across industries is presented by Murphy and Topel (1987), and is summarized in Table 3. This table shows average annual hours over an 8-year period for operatives employed in nine major industries, and in all industries. The range of mean hours over the nine industries is 356, or 17 percent of average annual hours for all operatives. In contrast to this variability in expected hours of work, the sum of hours
of work and reported hours of unemployment is remarkably stable across industries.

Differences across industries or occupations in mean hours, however, do not necessarily reflect employer-determined constraints. Individuals may sort themselves into different industries on the basis of preferences for leisure.\textsuperscript{24} Evidence in Murphy and Topel (1987) based on the comparison between cross-sectional and longitudinal estimates of the industry-effects in conventional earnings equations suggests that there is considerable sorting across industries in the dimension of unobserved "ability" or earnings capacity. It seems reasonable to expect similar sorting in the dimension of preferences for leisure.\textsuperscript{25}

Nevertheless, Altonji and Paxson (1986) present strong evidence of heterogeneity in hours of work across different jobs, controlling for individual characteristics. Their findings, some of which were reported in Table 1, suggest that changes in annual hours associated with a change in jobs have approximately three times larger cross-sectional variability than those associated with no change in jobs. This is true for weeks per year and hours per week, and it is also true for changes in hours over three and five year intervals. Furthermore, comparing the variation of hours changes for the same individual during intervals with a change in employer, and during intervals with no change in employer, Altonji and Paxson (1986) continue to find that movements across jobs are associated with roughly three times greater variability in hours changes. Their results suggest that a substantial component of individual labor supply is job-specific, and that further analysis of the
relation between hours characteristics and job choices is justified.

A simple framework for measuring compensating differentials for employer-determined hours constraints and relating these differentials to conventional labor supply parameters was proposed by Abowd and Ashenfelter (1979). They noted that the percentage differential between a wage offer that permits an unconstrained hours choice of \( h^0 \), and the wage that is required to induce the same individual to accept a constrained hours choice of \( \bar{h} \) is approximately

\[
\Delta_1 = \frac{1}{2\eta} \left( \frac{h^0 - \bar{h}}{h^0 \bar{h}} \right)^2
\]

where \( \eta \) is the compensated labor supply elasticity. They also showed that a second order approximation for the wage differential required to compensate for stochastic variation in \( \bar{h} \) at fixed hourly wage rates is

\[
\Delta_2 = \frac{1}{2} r c^2
\]

where \( c \) is the coefficient of variation of \( \bar{h} \), and \( r \) is interpreted as a generalized index of relative risk aversion. In fact, it can be easily shown that \( r = 1/\eta \), so that the rates of compensation for these two aspects of hours constrains are equal, at least to a second order approximation.26/

Empirical evidence on the magnitude of compensating differentials for employer-determined hours, however, is mixed. In their 1979 paper, Abowd and Ashenfelter constructed individual-specific estimates of the probability \((\pi)\) and duration \((\delta)\) of layoffs for white males in the PSID, using industry, occupation, and aggregate information as well as previous layoff histories for each individual. They used the product
as a measure of the expected proportional constraint

\( (h^0 - E(\tilde{h}))/h^0 \), and an estimate of the variance of time laid off, based

on \( \pi, \delta \), and the variance of their layoff duration equation, as a

measure of the variation in \( \tilde{h} \). They then included the differentials

\( \Delta_1 \) and \( \Delta_2 \) on the right-hand side of a wage equation, treating the

inverse labor-supply elasticity \( (1/\eta) \) and the risk coefficient \( (r) \) as

parameters. Their estimates are fairly sensitive to the definition of

the wage rate and also to the treatment of individual effects, although

in no case is their estimate of the inverse labor supply elasticity

significantly different from zero by conventional standards. The risk

coefficient \( r \) (associated with the regression coefficient of the

variance of layoff durations) is more precisely estimated. Taken

literally, Abowd and Ashenfelter's (1979) results suggest an estimate of

\( \eta \) in the range of .10 - .20, which is not very different from the range

of estimates based on conventional labor supply models (either static or

intertemporal). It is important to keep in mind, however, that if Abowd

and Ashenfelter's model of the labor market is correct, and hours vary

at fixed real wage rates, then the apparent labor supply elasticity as

measured by conventional methods is (essentially) infinite, at least for

employees observed in the same job over time.

Later research by Abowd and Ashenfelter (1984) and Murphy and Topel

(1987) uses estimates of the means and variances of hours within

industry and occupation cells to estimate the compensating differentials

earned by individuals in those industries and occupations. Following

this approach, Abowd and Ashenfelter (1984) again report large but
imprecise estimates of the inverse labor supply elasticity, suggesting that the relation between wage rates and mean hours by industry/occupation is weak. The simple averages by industry for male operatives in Table 3 give some idea of the nature of the problem: in these data the correlation between industry average wage rates and industry average hours or unemployment rates is approximately zero.

Although Murphy and Topel’s (1987) specifications are not directly comparable to Abowd and Ashenfelter’s, they too report difficulty in obtaining an interpretable relation between individual earnings and the squared deviations of industry and occupation mean earnings and hours from their overall means. The hypothesis that compensating differentials for employer-determined hours can adequately describe the observed inter-industry and occupation differences in mean earnings is overwhelmingly rejected. Murphy and Topel (1987) suggest that this is due, in part, to sorting of individuals into different industries by unobserved ability, and present some evidence of this phenomenon.

The weak evidence of compensating differentials for hours constraints in cross-sectional data is consistent with the finding of rather small wage premiums for other job amenities, including safety and work effort. 27/ One explanation for the absence of significant differentials is unobserved individual heterogeneity. The statistical framework for estimating compensating differentials relies heavily on the absence of unmeasured heterogeneity in individual tastes and opportunities. In cross-section, however, workers of higher unmeasured ability may earn higher wages and choose less constrained jobs, so that
the observed differentials are small, or even wrong-signed. Alternatively, taste variation may accommodate the observed variation in hours distribution across industries, occupations, and employers with no associated wage premiums.

A simple solution to the problem of heterogeneity is to use repeated observations on the same individual over time. Abowd and Ashenfelter's (1979) attempt to include individual effects in their compensating differentials equations resulted in less precise and sometimes wrong-signed estimates of the differential for expected hours constraints. A slightly different approach is taken by Altonji and Paxson (1987), who estimate first-differenced wage equations for job changers in the PSID sample that include measures of reported hours constraints on the old and new jobs. They find that movements between jobs that result in the relaxation of previous constraints are associated with smaller wage growth, while those that result in the imposition of new constraints are associated with higher wage growth. Their estimates of the wage premiums for reported constraints are relatively small but are roughly consistent with those reported by Abowd and Ashenfelter (1979, 1984).

Overall, there is only weak empirical support for the notion of compensating differentials for employer-determined hours. The magnitude of the estimated differentials involved is uniformly small, and goes only part of the way toward explaining observed earnings differences across industries, occupations, and employers. There are also a number of recurrent problems in the literature, including findings by Abowd and
Ashenfelter (1979) and Topel (1984) that the effect of unemployment insurance programs is to more than offset the earnings losses associated with reduced hours, and the basic fact that average industry and occupation hours or unemployment rates are only weakly related to wage rates. The results in the literature so far do not lend overwhelming credence to a model of the labor market in which employers respond to variations in demand for labor at fixed hourly wage rates in return for equalizing wage premiums based on the expected distribution of hours.

II. Studies on the Demand Side

In comparison to the study of labor supply, the study of labor demand using individual (i.e., firm-specific) data is in its infancy. As Stafford (1986) has pointed out, this is mainly due to an absence of data at the firm level. One of the most interesting developments in research on labor demand in the past decade is the emergence of time-series-cross-sectional data sets with information on the employment decisions of individual firms. Much of this research has focused on the question of whether employment determination in a unionized setting involves setting the marginal product of labor equal to the observed wage rate, or some other measure of the opportunity cost or alternative value of employees' time.

The motivation for this research is two-fold. On one hand, there is a long theoretical literature which distinguishes two alternative models of the determination of quantities and prices under bilateral monopoly: a monopoly model (usually attributed to Dunlop (1944)) in
which the trade union sets wages and the firm determines employment from its conventional labor demand curve; and a co-operative model (first applied to the employment determination problem by Leontief (1946)) in which wages and employment are jointly determined on the contract curve formed by the locus of tangencies between the parties' indifference curves in wage-employment space. Except in the case where the union places no marginal value on increased employment, the co-operative approach implies that the marginal product of labor is different from the observed wage rate, while the monopoly model implies that the two are equal.  

A second and related motivation is the literature on implicit contracts. A key prediction of this literature is that employment is related to a shadow-value of labor that falls below the observed wage rate. In fact, except for the interpretation of the employees' preference function, the cooperative or efficient bargaining model is equivalent to a simple implicit contracting model. The basic idea of both models is that, in the absence of employee mobility, the employment allocation decision can be separated from the wage determination decision. The monopoly bargaining or inefficient contracting model, by comparison, assumes that employment determination is related solely to observed wage rates, as in the traditional demand and supply model.

Just as on the supply side, interest in the more flexible models of employment determination offered by the cooperative bargaining or implicit contracting models has been stimulated by an absence of systematic correlations between wage rates and short-run employment.
determination. In contrast to the supply side, however, most of the available evidence on employment demand is based on aggregated data. The careful study of quarterly two-digit manufacturing industry data by Nadiri and Rosen (1973), for example, shows no evidence of systematic wage effects in a partial adjustment model. Remarkably similar conclusions emerge from an application of the same model to quarterly firm-specific data for mechanics in the airline industry in Card (1986). Efficient contracting models (or implicit contracting models) offer the potential, at least, of explaining the lack of success of short-run employment demand models by focusing on an alternative measure of the shadow of labor.

Before reviewing those recent studies that focus on the efficient-contract interpretation of firm-specific employment and wage outcomes, however, I would like to point out that these same firm-level data may be very useful for studying the supply side! For example, if variation in employment demand is mainly a result of exogenous demand shocks, then a regression of wage rates on employment or hours per employee using output as an instrumental variable identifies the inverse labor supply elasticity faced by the firm. If short-term variations in labor demand are met by changes in weekly hours or layoff-rates among a fixed pool of workers, then this labor supply elasticity is in principle the same one that should be estimated using wage rates and hours for the firm's employees (assuming, of course, that the simple demand and supply framework is appropriate.)

The assumption of hours variations among a fixed pool of workers is
more or less appropriate to describe fluctuations in man-hours brought about by predictable seasonal demand fluctuations. In this case, suitable instrumental variables are the seasonal dummy variables themselves, and one can obtain a rough idea of the inverse supply elasticity (defined over the period of a month or a quarter) by comparing the amplitude of the seasonal cycle in wages to the amplitude of the seasonal cycle in employment.\textsuperscript{36} It is my impression that the seasonality in wage rates is very small, even in seasonal industries like durable manufacturing.\textsuperscript{37} This suggests that the firm-level labor supply elasticity is relatively large, or alternatively, that seasonal variation in output demand is met by variation in employment at fixed wage rates.

A similar approach could be applied to year-to-year fluctuations in employment arising because of the aggregate business cycle. Over the longer term, however, a larger and larger fraction of employment adjustment is presumably met by variation in the number of workers attached to the firm. The inverse labor supply elasticity estimated from such data therefore incorporates both the labor supply responses of permanently attached workers and the effect of movements of workers between jobs. The latter movements are presumably very wage-elastic, and provide no information on the labor supply elasticity of a given individual.

A basic framework for the demand-side interpretation of firm-specific employment and wage data, which allows for the possibility of both the monopoly or "inefficient" outcome, as well as a variety of alternative "efficient" outcomes, is presented by Brown and Ashenfelter (1986, Section III). They suggest a log-linear approximation to the
structural equation determining employment \((N)\) in terms of the observed wage \((w)\), some measure of the opportunity wage available to workers \((a)\), and a vector \((X)\) of control variables (other factor prices, time trends, etc.) of the form

\[
\log N = \alpha_0 + \alpha_1 \log w + \alpha_2 \log a + X\beta.
\]

This specification includes as special cases the conventional labor demand model \((\alpha_2 = 0, \alpha_1 = -\phi)\), where \(\phi\) is the elasticity of labor demand; a strong form of the co-operative bargaining model in which the marginal product of labor is equated to the alternative value of workers' time \((\alpha_1 = 0, \alpha_2 = -\phi)\); and intermediate cases associated with alternative parameterizations of workers' preferences. The employment demand equation associated with the simple implicit contract model of Azarjadis (1975), for example, implies \(\alpha_1 = m \phi\) and \(\alpha_2 = -(1 + m\delta)\phi\), where \(m\) is the average percentage markup of \(w\) over the alternative wage and \(\delta\) is the coefficient of relative risk aversion associated with individual employees' preferences over income risks.

This simple equation makes clear the sense in which efficient wage bargaining can help to explain the absence of strong correlations between employment and wages. Suppose that bargaining is efficient in the Coasian sense that the net income of the firm and workers as a whole is maximized. Then observed wages represent a pure transfer payment between the firm and its employees, and one could not expect to measure the elasticity of labor demand from the covariation of \(N\) and \(w\). In this case, knowledge of supply-side information, in particular the value
of a, is necessary to correctly specify observed employment outcomes on the demand side.

Brown and Ashenfelter (1986) present a wide variety of estimates of equation (6) on a sample of 10 locals of the International Typographic Union (ITU) over the period 1948-65. Their estimates of the coefficient \( \alpha_2 \) are generally imprecise, and distressingly sensitive to the definitions of employment and alternative wages. On the other hand, when they define the opportunity wage as \( a(1-U) \), where \( a \) is some measure of the alternative wage rate and \( U \) is the measure of the unemployment rate in the local labor market, they find consistent evidence that employment is positively related to \( U \). This finding confirms the basic idea of efficient contracting models, which is that employment responds to supply side opportunities even holding constant the wage rate. Finally, Brown and Ashenfelter's empirical results suggest that contract wage rates have a generally negative effect on employment. This rules out the "strongly efficient contract" hypothesis \( \alpha_1 = 0 \), but does not rule out the possibility that employment determination is based on a measure of the opportunity cost of labor that depends on both the contract wage for printers, and some measure of their extra- contractual opportunities.

Similar data for a different set of ITU locals is analyzed in a more highly structured framework by Macurdy and Pencavel (1986). They too conclude that there is evidence of some intermediate model of efficient contracting in which both the measured contract wage and extra-contract opportunity wages determine the shadow value of employment.
They do not, however, provide direct evidence on whether measures of the alternative wage may be excluded from the employment demand equation, arguing that such a test provides no information on the issue of whether or not contracting is "efficient". Specifically, Macurdy and Pencavel point out that certain assumptions on the form of union preferences imply that $\alpha_2 = 0$ in equation (6), even though the marginal product of labor is set to a value different than the contract wage. Of course, in this case, apart from functional form restrictions, the efficient contracting model is observationally equivalent to a conventional labor demand model, and fluctuations in employment are related directly to fluctuations in contract wages, so that much of the appeal of the efficient contracting hypothesis is lost.

The simple model of equation (6) is extended to a dynamic setting in Card (1986) and applied to quarterly data on employment, output, and wage rates for airline mechanics at 7 U.S. airlines during the period 1969-76. Again, there is evidence that employment responds to variations in both the contract wage and a measure of the alternative wage available to airline mechanics, although the restrictions implied by the model on the joint representation of employment, wages, and output are strongly rejected. As Brown and Ashenfelter (1986) report for many of their specifications that use ITU tax receipts to measure employment, there is evidence for the airline mechanics that employment responds positively to alternative wage rates. This finding consistent with an efficient contracting model in which optimal combinations of employment and wages lie on a positively sloped contract
curve. It is also roughly consistent, however, with a more conventional supply and demand model in which supply depends positively on contract wages and negatively on alternative wages. 40/

Both Brown and Ashenfelter (1986) and Card (1986) also report that short-run price changes, while significantly negatively related to real contract wage rates and measured real alternative wage rates, have no correspondingly positive effect on employment. This finding suggests that employment does not respond to changes in real wage rates induced by nominal price increases, and is consistent with an efficient contracting model in which the alternative value of employees' time is invariant to aggregate price movements. Such might be the case if the primary alternative to contractual employment is "home production" or leisure.

A very different test between alternative models of employment determination in unionized firms is proposed by Abowd (1987), who uses the stock market reaction to news of a wage settlement to measure the tradeoff between increases in union member "wealth" (income flows over the term of the contract) and the market value of the firm. The basic idea of Abowd's test can be illustrated very simply in a one-period model where the value of the firm is equal to its profits during the period. Let \( w \) represent the negotiated contract wage and let \( \pi(w) \) represent the conventional profit function of the firm, assuming that the firm can select employment \( (N) \) after \( w \) is known. Then, using a second-order expansion of the profit function around the expected wage settlement \( w^* \), the change in the value of the firm in response to the announcement of a wage settlement \( w \) is
\[ \Delta V_1 = \pi(w) - \pi(w^*) = -N^*(w - w^*) + \frac{1}{2} \left( \frac{w - w^*}{w^*} \right)^2 \phi w^* N^* , \]

where \( N^* \) is the level of employment associated with the expected wage \( w^* \) and \( \phi \) is the absolute value of the elasticity of demand for labor.

On the other hand, the change in the value of the firm associated with the same wage announcement, assuming that the firm sets employment so that the marginal product of labor equals the alternative wage available to workers, is

\[ \Delta V_2 = -N^E(w - w^*) , \]

where \( N^E \) is the "efficient" level of employment.\(^{41}\)/ A test of the efficient contracting model in this framework amounts to a test of the strict linearity of the profit function under the null hypothesis of efficiency, versus the convexity of the profit function under the alternative.\(^{42}\)/

Aboud (1987) uses a Phillips-curve type equation to forecast the nominal wage settlement for each of some 2200 collective bargaining agreements negotiated during the period 1976-82 for which associated stock market information is available. The residual from this equation is an estimate of \((w - w^*)\). Multiplying by the number of employees on hand at the settlement date, he performs a regression of the actual change in the stock market value of the firm during a three-month window around the contract settlement date on the unexpected change in union income \((N(w - w^*))\) associated with the settlement. Aboud finds no evidence against the strict linearity hypothesis (dividing settlements into those with \( w - w^* > 0 \) and those with \( w - w^* < 0 \)), although the standard errors are somewhat imprecise. His point estimates of the rate
of trade-off between the value of the firm and the the income stream of union members average something less than unity, although this could be expected on the basis of measurement error in $N$ and $w^*$, and imprecision in the appropriate time horizon of the parties.

In summary, evidence for the efficient contracting hypothesis based on augmented employment demand equations of the form estimated by Brown and Ashenfelter (1986), Macurdy and Pencavel (1986), and Card (1986) is perhaps mildly supportive of the theory, although significant problems remain. First, exclusion tests for the presence of the alternative wage variable are not always significant. Second, the sign pattern of the effects of contract wages and alternative wages is not always explicable within the confines of the theory. On the other hand, the findings of Brown and Ashenfelter (1986) with respect to the effect of local unemployment rates on contract employment, and the findings of Brown and Ashenfelter (1986) and Card (1986) with respect to the absence of aggregate price effects on employment may be considered as evidence in favor of some notion of efficiency, and against the simple demand and supply model of short-run employment determination.

Evidence from the event study methodology devised by Abowd (1987) is also difficult to interpret. Abowd's test relies on significant convexity of the conventional profit function in order to detect departures from the strongly efficient model that forms his null hypothesis. If the actual curvature of the profit function (i.e., the elasticity of labor demand) is small, departures from linearity will be difficult to detect. In any case, Abowd's results suggest that the stock market
reacts to wage announcements as if there will be no adjustment to employment in response to unanticipated changes in wages. This finding is really just another statement of the difficulty in isolating the responses to real wage changes that are predicted by simple supply and demand models of the labor market.

Conclusion

Recent microeconometric research on individual- and firm-specific data suggests that simple supply and demand models of the labor market are incomplete. On the supply side, the variation over time in individual employment outcomes is enormous. Yet, only a trivial fraction of this variation is associated with variation in observed wage rates. At the same time, however, individual employment outcomes are highly correlated with market-level measures of employment demand in the individual's industry and local labor market. Taken together, these facts are evidence against the strict separation of the two sides of the labor market predicted by simple supply and demand models. They suggest that economists' attention might fruitfully shift toward alternative models of the labor market, which permit an uncoupling of short-run employment and hours choices from wage rates.

Similar conclusions emerge from studies of firm-level employment demand. The recent literature on efficient contracting suggests that, in the unionized sector at least, the link between employment and wage rates is more complex than predicted by the simplest supply and demand models. This is not to say, however, that the alternative models that
have been proposed to date have been unambiguously successful. Research on compensating differentials for employer-determined hours has generally failed to uncover large or systematic premiums predicted by models in which labor supply responds directly to labor demand with little or no commensurate wage adjustment. Similarly, models of employment determination at the firm level which relate employment outcomes to a shadow wage rate that varies with the alternative value of workers' time have only been moderately successful.

A major obstacle to both line of research is the paucity of micro-level data sets which contain information on both sides of the labor market. While such data are unnecessary for a strict supply and demand model, they are obviously essential in any model that attempts to relax the separation of the two sides of the labor market. The collection and analysis of such data should be a high priority for future research.
Footnotes

1. The market price can also serve as a sufficient statistic for information that is privately held by other individuals on the same side of the market: see Grossman (1981).

2. Examples are mentioned by Kennan (1985, p. 2) and Pencavel (1986, footnote 31, p. 34).


4. See Ham (1986) and the review article by Ham (1987).

5. See Abowd and Ashenfelter (1979, 1984), and Murphy and Topel (1987).


7. Some of the same material is summarized in the recent survey by Kniesner and Goldsmith (1986).


9. These standard deviations are based on cross-sectional variances around the average annual change in hours. The average change in hours in the PSID and NLS data sets exhibits some variation from year-to-year: see Abowd and Card (1986, Table 1).

10. Altonji and Paxson estimate the coefficients of their regression adjustments by instrumental variables, to account for measurement error in reported wage coefficients are between .03 and .08, with standard errors of about .3 to .4.

Duncan and Hill's (1986) estimates are for hourly-rated workers who remained with the same firm over the two years of their validation study. Using a sample of male heads from the PSID who report earnings and hours data for 1976-80, I estimated the ratio of the variance of the change in annual hours on the main job for hourly-rated workers with the same employer in consecutive years to the variance of the change in annual hours on the main job for all workers to be 1.09 for changes between 1978 and 1979, and 1.03 for changes between 1979 and 1980. Thus Duncan and Hill's sample of hourly rated workers with the same employer in each year may be fairly representative of all workers in terms of variability in individual hours measures.

Of course the autocorrelation coefficient of $\Delta \log h_t$ is not identified in this model, and could be outside the range of zero to .3. I base my estimate on the fact that most aggregate annual labor market series (real wage rates, employment, unemployment) display zero or slight positive autocorrelation in first-differences.

There is some evidence that changes in the logarithm of annual earnings and hours are not normally distributed: see Macurdy (1982) and Abowd and Card (1986).

Some additional evidence of the variability of employment outcomes is provided by Feenberg (1987), who uses individual tax-return data to estimate the year-to-year variation in earnings. Feenberg's estimate of the standard deviation of the change in the logarithm of taxable earnings for males under 65 years of age is .32.
This simple model can be extended to the case where individuals lack perfect foresight: see Macurdy (1982) and Pencavel (1986, pp. 88-89).

If the model is amended to include uncertainty, then both $\Delta \log h_t$ and $\Delta \log g_t$ include an additional stochastic component reflecting the updating of the individual's marginal utility of wealth.

A similar conclusion emerges when a component of variance is added to reflect unanticipated changes in the marginal utility of wealth.

Here I am following the convention of the PSID survey, which divides annual weeks into these four categories.

This objection is raised by Pencavel (1986, p. 43).

I am grateful to John Ham for supplying these tabulations.


Murphy and Topel (1987) provide similar data for a variety of other broad occupations. The variation across industries in annual work hours is greater for laborers and transportation operatives, and lesser for professional and clerical workers.

This observation was made by Lewis (1969).

It would be very useful to repeat Murphy and Topel's (1987) analysis of industry effects using individual hours data. For individuals who change employer between annual censuses, however, reported hours in the second survey contain a mixture of hours in the old and new jobs. Thus an analysis of industry effects for job changes requires at least three years of data on each individual.
Abowd and Ashenfelter's approximations are exact in the limit as \( \eta \to \infty \), but may be quite misleading for relatively low values of \( \eta \). For example, if \( \eta = 0 \) the proportional compensating differential for underemployment is \( (h^0 - \bar{h})/h^0 \), while the differential for overemployment is arbitrarily large. This suggests that researchers should use an alternative approximation if the implied estimates of the labor supply elasticity are relatively small.

See Brown (1980), who reports mixed results for the estimation of wage premiums for a variety of job amenities in both cross-sectional and longitudinal specifications.

So far, union contracts and/or records have been the major source of these data. It is to be hoped that similar data from a nonunion environment can eventually be assembled.


Brown and Ashenfelter (1986) provide some interesting data on controlled bargaining experiments designed to test between these two competing models of quantity and price determination under bilateral monopoly.

Useful surveys of the implicit-contracting literature are provided by Hart (1983) and Rosen (1985).

Brown and Ashenfelter (1986, pp. 556-557) make this connection explicit.
For example, McDonald and Solow (1981) motivate their theoretical analysis of efficient wage bargaining by asking why it is that fluctuations in the demand for labor "... so often lead to large changes in employment and small, unsystematic changes in real wage rates?" (p. 896).

See also Sargent (1978) and Geary and Kennan (1982). In contrast to the generally negative results reported in studies of short-run employment demand in aggregate U.S. data, results for the U.K. have been more successful. See Nickell (1984), for example.

This idea is used by Sullivan (1986) to test for the presence of monopsony power in the market for nurses.

Consider a variable equal to unity in the season of highest demand, -1 in the season of lowest demand, and zero otherwise. The instrumental variables estimate of the inverse labor supply elasticity using this variable as an instrument for employment is just the relative amplitude in the seasonal changes in wages and employment.

A major source of seasonal variation in marginal wage rates is changes in the fraction of hours paid at overtime rates. Assuming that the overtime premium is fixed at 50 percent, the range of monthly variation in the fraction of hours paid at overtime rates for production workers in durable manufacturing is tiny: based on an average of 1984-86 data, the fraction of workers paid at overtime rates varies between .080 and .095.

One could not expect to recover elasticities of supply from the covariation of employment and wages among the employees in such a contract either.
If the union utility function is of the form

$$U(N, w, a) = g(N)(w/a)^e$$

where $g$ is a constant elasticity function, then the gap between the contract wage and the shadow value of labor in the efficient contract is simply a constant. See Macurdy and Pencavel (1986, p. S13).

A similar finding is reported by Neelin (1987), who studies employment and wage outcomes for teachers in 200 bargaining units in Ontario (Canada) over the period 1976-1983.

A maintained assumption is that the wage settlement does not reveal any new information about the alternative wage or future demand conditions, so that the level of $N^E$ is independent of wage announcements.

Note that the expressions in (7) and (8) are equivalent if the elasticity of demand is negligible.
References


Pencavel, John. "The Tradeoff Between Wages and Employment in Trade


<table>
<thead>
<tr>
<th>Authors</th>
<th>Survey &amp; Hours Measure</th>
<th>Sample</th>
<th>Std Dev of Chg in Ann Hrs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abowd &amp; Card (1987), Table 3</td>
<td>PSID, 1969-79</td>
<td>Males 21-64, continuously heads, nonzero hrs and earnings in all years</td>
<td>.34</td>
</tr>
<tr>
<td></td>
<td>All Hours</td>
<td>Subsample with nochg in employer during sample</td>
<td>.20</td>
</tr>
<tr>
<td>Abowd &amp; Card (1987), Table 3</td>
<td>NLS, 1969-75</td>
<td>Males 45-64, nonzero hrs and earnings in all years</td>
<td>.33</td>
</tr>
<tr>
<td></td>
<td>Hours on Current or Last Job</td>
<td>Subsample with nochg in employer during sample</td>
<td>.20</td>
</tr>
<tr>
<td>Altonji &amp; Paxson (1986), Tables 1, 2</td>
<td>PSID, 1968-81</td>
<td>Males 18-60, nonzero hrs in consecutive years</td>
<td>.35</td>
</tr>
<tr>
<td></td>
<td>Hours on Main Job</td>
<td>Subsample with same employer in consecutive years</td>
<td>.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adjusted</td>
<td>.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Subsample with diff employer in consecutive years</td>
<td>.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adjusted</td>
<td>.52</td>
</tr>
</tbody>
</table>

Note: The standard deviation of the percent change in hours is estimated by the cross-sectional standard deviation of the change in the logarithm of annual hours. Adjusted estimates refer to the cross-sectional standard deviation of the regression-adjusted change in the logarithm of annual hours.
<table>
<thead>
<tr>
<th>Other Explanatory Variables:</th>
<th>Current</th>
<th>Lagged</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Year Effects</td>
<td>Year Effects</td>
</tr>
<tr>
<td>1. Change in log average hourly earnings (%)</td>
<td>19.1 (.001)</td>
<td>26.2 (.000)</td>
</tr>
<tr>
<td>2. Change in log average hourly earnings, change in hours of unemployment (%)</td>
<td>6.6 (.159)</td>
<td>3.7 (.448)</td>
</tr>
</tbody>
</table>

Note: Reproduced from Table A-6 of Ham (1984). The statistics are exclusion test statistics for the first-differences of four labor demand variables (occupation, industry, and local unemployment rates, a dummy variable for whether or not the individual reported that his job had folded) in a first-differenced log-hours equation fit to male household heads in the PSID.
Table 3
Hours Worked, Hours Unemployed, and Wage Rates for Male Operatives in 9 Industries

<table>
<thead>
<tr>
<th>Industry</th>
<th>Hourly Wage</th>
<th>Unempl. Hours</th>
<th>Hours Worked</th>
<th>Total Unempl &amp; Work Hrs</th>
<th>Unempl. Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construction</td>
<td>3.87</td>
<td>290</td>
<td>1834</td>
<td>2124</td>
<td>13.7</td>
</tr>
<tr>
<td>Lumber</td>
<td>2.99</td>
<td>158</td>
<td>1979</td>
<td>2136</td>
<td>7.4</td>
</tr>
<tr>
<td>Primary Metals</td>
<td>3.89</td>
<td>128</td>
<td>2009</td>
<td>2137</td>
<td>6.0</td>
</tr>
<tr>
<td>Fabric. Metals</td>
<td>3.10</td>
<td>126</td>
<td>2007</td>
<td>2133</td>
<td>5.9</td>
</tr>
<tr>
<td>Machinery</td>
<td>3.37</td>
<td>88</td>
<td>2099</td>
<td>2187</td>
<td>4.0</td>
</tr>
<tr>
<td>Automobiles</td>
<td>4.04</td>
<td>167</td>
<td>1983</td>
<td>2150</td>
<td>7.8</td>
</tr>
<tr>
<td>Paper</td>
<td>3.69</td>
<td>60</td>
<td>2159</td>
<td>2219</td>
<td>2.7</td>
</tr>
<tr>
<td>Petroleum Ref.</td>
<td>4.65</td>
<td>59</td>
<td>2095</td>
<td>2154</td>
<td>2.7</td>
</tr>
<tr>
<td>Retail Trade</td>
<td>2.76</td>
<td>88</td>
<td>2190</td>
<td>2278</td>
<td>3.9</td>
</tr>
<tr>
<td>All (Weighted)</td>
<td>3.21</td>
<td>116</td>
<td>2050</td>
<td>2167</td>
<td>5.4</td>
</tr>
<tr>
<td>(Standard Dev.)</td>
<td>(.61)</td>
<td>(52)</td>
<td>(80)</td>
<td>--</td>
<td>(2.4)</td>
</tr>
</tbody>
</table>

Notes: Taken from Murphy and Topel (1987), Table 5.1. Data are from March CPS Annual Demographic Supplement 1977-84. Unemployment rate is ratio of unemployed hours to sum of hours unemployed and hours worked. The number of observations is 8779.