THE USE OF WAGES IN COORDINATING HOURS OF WORK

Aloysius Siow*
University of Toronto

* I thank Joe Altonji, Jeff Biddle, Douglas Holtz-Eakin, George Jakubsen, Jacob Mincer, Brenden O'Flaherty and members of seminars at Columbia University, Cornell University, Hoover Institution, Northwestern University, Princeton University, Rice University, SUNY at Stony Brook, University of Chicago and University of Washington for helpful discussions. Joe Altonji also graciously supplied the data for the study. This project was partially funded by the Bradley Foundation. The usual disclaimer applies.
The Use of Wages in Coordinating Hours of Work

This paper estimates models of coordinated labor supply with micro panel data. Macurty's model of labor supply is modified by making the wage relate negatively to the deviation of a worker's hours from the mean of his co-workers. The results suggest that the wage is negatively related to the deviation of a worker's hours from the mean of his co-workers. Moreover, workers choose their hours of work taking this relationship into consideration. Workers with large deviations, either positive or negative, are more likely to leave their occupations and or industries. The movers tend to deviate less in their new environment.

Keywords Coordination of hours of work, estimation of labor supply

JEL subject classification 810

Aloysius Siow
University of Toronto
Department of Economics
Toronto, Ontario M5S 1A1
Canada
(416) 978-4139
and thus his income will suffer. In an earlier draft of this paper (Siow (1988)), I showed that as long as the productivities of agents are affected by each others' choices of hours, they will take the choices of others into account when making their own decisions. The resulting equilibrium due to this implicit coordination is a Nash Equilibrium. Compared with the explicit coordination within a firm, the Nash Equilibrium is inefficient. But there is more bunching in hours of work than would be predicted by the standard model of labor supply.

Thus a robust finding of theoretical models of explicit or implicit coordination is that there will be bunching of hours in equilibrium among heterogeneous workers. This bunching may explain the small labor supply elasticities found in most empirical studies of the standard model of male labor supply. This paper asks whether empirical labor supply models that include the need for coordination will obtain larger labor supply elasticity estimates.

In order to study whether the coordination of hours would affect estimates of labor supply elasticities, I extended Macurdy's dynamic model of labor supply (Macurdy (1981)) to allow for coordination constraints. Even though the coordination problem is a static issue and may be studied in a one

---

2 The possible inefficiency of Nash Equilibria in coordination games are well known. See for e.g. Diamond (1982), Economides and Siow (1988) and Schleifer (1986). Multiple equilibria are also common. It is likely that there are also multiple equilibria in the class of models that I discuss but I am ignoring that issue here.

3 A caveat is necessary in evaluating the necessity of the coordination of hours of work by looking at the tradeoff between wages and hours of work. Even if coordination is important, when there are enough co-workers with the same excess demand for hours of work, they will work together and the observed wage schedules may not reflect the need for coordination.
period framework, I believe that labor supply is best studied in a dynamic context. I chose Macurdy's model as the benchmark because previous work applying it on the same data set that is use in this paper have found small labor supply elasticities (e.g., Abowd and Card (1989), Altonji (1986), Ham (1986), Macurdy). Moreover Ham found that workers were hours constrained. Thus my results can easily be compared with the earlier literature.

To empirically measure the need for coordination, I let the wage relate negatively to the deviation of a worker's hours of work from the mean hours of his co-workers. If co-workers work at the same time so as to coordinate their activities, their hours of work will be the same. If workers work the same hours and or at the same time for their own private reasons, my measure will overestimate the amount of coordination. Thus my measure is a necessary but not sufficient condition for coordination. Another problem is that I lack data on hours of his co-workers. I assume that the mean hours of work of all workers in his occupation, industry and year is a good estimate of the mean hours of his co-workers. With this proxy, I estimate the models using micro panel data from the Michigan Panel Study of Income Dynamics (PSID).

I use annual hours of work in this study. There are two reasons for focusing on annual hours of work. First, variation in annual hours of work is due to both variation in days worked and average hours worked per week. Both margins require coordination. If a worker works only half of the year whereas his co-workers work the whole year, he will be less productive because his firm will not put him on projects that will last longer than a half year. Of course, coordination of hours over a workday also matters. The annual hours
of work measure captures both margins \(^4\) The second reason for focusing on annual hours of work is to make this work directly comparable to the earlier labor supply studies that ignored coordination costs.

The first set of empirical tests of the model is based on the behavior of mean wages by industry, occupation and year. The model predicts that the mean wage in an industry, occupation and year will be negatively related to the variance of hours worked in that category. I am also able to provide tests to distinguish between penalties for hours departures from category specific means versus an economy wide mean \(^5\) The empirical results show that the mean wage in an industry, occupation and year category is negatively related to annual hours variation in that category. It also shows that the penalties are more closely tied to departures from category specific means than to the economy wide mean.

The model also predicts specific mobility patterns that are tied to hours deviations. The empirical results show that workers with large deviations in annual hours, both positive and negative, are more likely to change industries and/or occupations in the next year. These movers deviate less in their new environment.

Finally the structural estimation show that the wage is negatively related to the deviation of a worker's hours from the mean of his co-workers.

\(^4\) There are also results with weekly hours of work in Siow (1988). Those results are largely consistent with ones reported here. The case for using weekly hours of work is to avoid the possible contamination of the annual hours of work results by unemployment behavior.

\(^5\) Motivated by considerations related to that in this paper, Moffitt, Biddle and Zarkin estimated wage schedules that are quadratic functions of hours together with static labor supply responses. I show that their estimates imply wage schedules that penalize workers for choosing hours away from the mean hours of work of the economy.
Moreover, workers choose their hours of work taking this penalty into consideration. My estimates of the wage penalty for a worker, for a one standard deviation departure (20% in annual hours) from the mean hours of his co-workers, ranges from 1 to 7%. My estimates of the intertemporal labor supply elasticity were not much larger than those found by Macurdy, Altonji (1986) and Ham who all used Macurdy's model and the PSID data set. The structural estimation is supportive of the theory.

A qualification in interpreting the empirical work is necessary. Although the results are supportive of my model, I present no direct evidence of coordination. One can argue, as Rosen (1969) and Owen do, that technologies in different industries and occupations require different optimal hours of work. Wage schedules respond to the demands of these technologies. The crucial element does not have to be coordination. However, a simple fixed cost argument will not explain the empirical findings. Different fixed costs in different industries or occupations will only require different minimum hours of work. It will not require bunching of hours around a mean. One may add to the fixed cost a fatigue factor so that too many hours is also suboptimal. Barzel (1973) makes this argument for daily hours of work. In other words, Barzel's theory suggest that the wage is low for a worker, whose hours is below that of his co-workers, because of the fixed cost of employment. The wage is also low for a worker who works above the hours of his co-workers due to the worker's lower productivity at long hours of work. I provide a test between my model and Barzel's. I show that the wage penalty for a worker is more closely tied to his hours deviation from the mean hours of workers in his industry, occupation and year rather than just his industry and occupation. A time varying benchmark for measuring the deviation in hours.
is difficult to reconcile with Barzel's theory

Finally, a branch of macroeconomic research is concerned with models of strategic complementarities which may lead to inefficiencies in the economy. I present empirical evidence of one form of this complementarity. Although I find evidence of induced bunching of hours in the data, I cannot tell if it is the efficient or the Nash outcome.

Section II presents three econometric models of coordinated hours of work. The econometric issues associated with estimating the models are presented in Section III and the data is discussed in Section IV. Section V discusses the empirical results. A conclusion is in Section VI.

II. The Econometric Models

Three econometric models of coordinated work hours are derived by modifying Macurdy's model of dynamic labor supply.

On any job, the firm affects a worker's choice of hours by making his wage negatively dependent on the deviation of his hours from the mean hours of his co-workers. The mean hours of his co-workers may differ by industry, occupation and year. In other words, all firms in an industry, occupation and year offer the same wage hours schedule to its workers. Given such a job specific wage schedule, the worker picks his hours of work which determines his earnings on that job as well. The worker's utility from a job is therefore determined by the job specific wage schedule, hours choice, and other characteristics. The worker faces a two steps labor market decision at any point in time. First, by optimally choosing hours and other job specific

---

attributes, the worker calculates his maximum discounted utility from any sequence of jobs. Then he picks the sequence that maximizes his discounted utility. Due to the other job specific attributes, a worker may not pick a job that only minimizes the penalty due to the deviation of his hours choice from those of his co-workers. When the hours preference of a worker changes transitorily, a worker may not change jobs due to the other job attributes that he values. If the hours changes are permanent, due for example to lifecycle changes, workers would be more likely to switch jobs to accommodate these changes in preferences.

With panel data on workers, I can only estimate the wage schedules they face, and their labor supply decisions. From these estimates one cannot tell if the wage schedules they face are efficient in the sense that firms are setting wages as incentive schemes or inefficient as in a non-cooperative coordination game. While there is a presumption of explicit coordination of hours within a firm, coordination across firms are as likely to be based on non-cooperative (Nash) behavior.

Let the intertemporal utility function of worker \( i \) who will be in the labor market for \( T+1 \) periods be

\[
(2.1) \quad V_i = \sum_{t=0}^{T} b^t u(C_{it}, H_{it})
\]

\( b \) discount factor, \( 0 < b < 1 \)
\( C_{it} \) consumption at time \( t \)
\( H_{it} \) labor supply at time \( t \)

The within period utility is
\[ (2.2) \quad U(\cdot,\cdot) = \alpha_{it}C_{it}^\beta - \mu_{it}H_{it}^\sigma \]

\[ 0 < \beta < 1, \quad \sigma > 1 \]

\[ \alpha_{it}, \mu_{it} \text{ non-negative taste shifters at time } t \]

The intertemporal budget constraint faced by the worker, in a world with perfect capital market, is

\[ (2.3) \quad \Sigma_{t=0}^T b^t(W_{it}H_{it} - C_{it}) = A_i \]

\( W_{it} \) refers to the wage on the job at time \( t \) on his optimal career path \( W_{it} = W(H_{it}, M_{it}, X_{it}) \) where \( M_{it} \) is the mean hours of his co-workers and \( X_{it} \) is his skills at time \( t \). \( A_i \) is the initial assets of worker. For simplicity, the market discount factor is assumed to be same as that used by the worker to discount his future utilities.

The worker want to maximize (2.1) subject to (2.3). His optimal labor supply decision at time \( t \) is characterized by

\[ (2.4) \quad U_{H_{it}} = \lambda_i \frac{\delta W_{it}H_{it}}{\delta H_{it}} \]

\( \lambda_i \) is the Lagrange multiplier associated with the intertemporal budget constraint (2.3). Economically, \( \lambda_i \) measures the marginal utility of wealth of that worker given his optimal choices.

Consider three different wage schedules corresponding to models A, B and C respectively. In logs, let the wage schedule for Model A be
\[ \ln W_{it} = X_{it} - 5r(H_{it} - M_{it})^2 \]

In (2.5), the penalty for deviating from \( M_{it} \) is symmetric around \( M_{it} \). While (2.5) looks like a particular hedonic wage function, it should not be interpreted as such. In a standard hedonic wage function, a firm treats \( M_{it} \) and \( r \) as exogenous and it picks \( H_{it} \) to maximize profits. In my framework, a firm uses wages to coordinate hours of work. Thus it determines \( r \) and \( M_{it} \) to affect the wage schedule that it offers its workers.\(^7\) Even if (2.5) is just an empirical description of the data and \( r \) is parametric to the firm, a firm will affect the wage schedule of any worker by its choice of his co-workers. Thus it is inconsistent to treat the observed wage hours locus as a standard hedonic wage equation where a firm cannot influence the wage schedules that workers face. I do not assume that a firm offers each worker a unique earnings and hours of work package, and zero wages for all other hours, because this would require remarkable knowledge on the part of the firm to find the right package for all its workers. Thus I believe that (2.5) is a realistic approximation to what firms actually do.

Substitute (2.5) and (2.2) into (2.4) to get

\[ \mu_{it} e^{H_{it}^d - 1} = \lambda_i W_{it} \{ 1 - r(H_{it} - M_{it})H_{it} \} \]

---

\(^7\) If \( M_{it} \) only consist of the mean hours of co-workers in the firm, then a firm has strong influence over this number by the wage schedules that it offers. However if the worker's co-workers include agents outside the firm, then the firm cannot control \( M_{it} \) as easily. It can set the penalty \( r \) to try to make its own workers be responsive to the coordination problem but the solution is a Nash Equilibrium and achieves only second best efficiency (See Siow (1988))
Take logs of (2.6) and approximate \( \ln(1 - \tau(H_{it} - M_{it})H_{it}) \) as \(-\tau(H_{it} - M_{it})H_{it}\) to get Model A's labor supply function

\[
(2.7) \quad \ln H_{it} = \lambda_i' + (\sigma-1)^{-1} \ln \lambda_i' - \tau(\sigma-1)^{-1}(H_{it} - M_{it})H_{it} - \mu_{it}'
\]

\[
\lambda_i' \quad (\sigma-1)^{-1} \ln \lambda_i
\]

\[
\mu_{it}' \quad (\sigma-1)^{-1} \ln \sigma \mu_{it}
\]

(2.7) is a variant of Macurdy's marginal utility of wealth constant labor supply function. The difference from his version is due to the additional term \(-\tau(\sigma-1)^{-1}(H_{it} - M_{it})H_{it}\). If \(\tau\) is zero, his version applies.

In logs, the wage schedule for Model B is

\[
(2.8) \quad \ln \bar{W}_{it} = X_{it} - 5\tau(\ln(H_{it}) - \ln M_{it})^2
\]

In (2.8) the penalty for deviating from \(M_{it}\) is based on the percentage deviation. Following the earlier derivation and approximate \(\ln(1 - \tau(\ln H_{it} - \ln M_{it}))\) as \(-\tau(\ln H_{it} - \ln M_{it})\), the marginal utility of wealth constant labor supply function for Model B is

\[
(2.9) \quad \ln H_{it} = \lambda_i'' + (\tau+\sigma-1)^{-1} \ln \bar{W}_{it} + \tau(\tau+\sigma-1)^{-1} \ln M_{it} - \mu_{it}''
\]

\[
\lambda_i'' \quad (\tau+\sigma-1)^{-1} \ln \lambda_i
\]

\[
\mu_{it}'' \quad (\tau+\sigma-1)^{-1} \ln \sigma \mu_{it}
\]
In models A and B, the penalty for the deviation in hours, at rate $\tau$, is the same across all industries, occupations and time. Model C relaxes this assumption by making the penalty proportional to the square of his deviation in hours divided by the standard deviation of hours of his peers. In this case, if his peers show large dispersion in hours, he will not face a large penalty for deviating from the mean hours of his peers. The wage function for Model C is

\begin{equation}
\ln W_{it} = X_{it} - 5\tau((H_{it} - M_{it})/S_{it})^2
\end{equation}

$S_{it}$ Standard deviation of hours of co-workers

Note that (2 10) is similar to (2 7). Following the derivation in Model A, the marginal utility of wealth constant labor supply schedule for Model C is

\begin{equation}
\ln H_{it} = \lambda_i' + (\sigma - 1)^{-1}\ln W_{it} - \tau(\sigma - 1)^{-1}(H_{it} - M_{it})H_{it}/S_{it}^2 - \mu_{it}'
\end{equation}

Associated with the three estimable wage schedules (2 5), (2 8) and (2 10), there are three estimable labor supply schedules (2 7), (2 9) and (2 11) respectively. Note that there is an overidentifying restriction between each pair of wage and labor supply schedules.

III. Econometric Issues

The econometric models are estimated with panel data on individual workers. The data do not contain information on a worker's co-workers.
Therefore the mean hours of the co-workers of worker $i$ at time $t$, $M_{i*t}$, have to be constructed. I construct $M_{i*t}$ as the mean hours of all workers, other than worker $i$, in the data set that are in the same industry, occupation and year as worker $i$. In other words, although there is no information on worker $i$'s co-workers, I assume that a worker interacts the most with his peers. His peers are also likely to face the same constraints that he does. Even if the hours worked of each worker is measured with error, $M_{i*t}$ is measured with less error because it is the mean hours of a group of relevant workers. An alternative to my construction is to use a reduced form hours equation to predict $M_{i*t}$. I do not use this alternative because the hours predicted from a reduced form equation includes the optimizing behavior of the agent.

Therefore the deviation of actual hours from this predicted hours will tend to underestimate the extent of the deviation of a worker's hours from that of his co-workers. Even my construction of $M_{i*t}$ is subject to this criticism. In practice, since I only use occupation at the one digit level to construct $M_{i*t}$, the underestimation is insignificant.

The second issue that have to be addressed is that $\lambda_i'$ and $\lambda_i''$ are both correlated with the characteristics of the worker including his wage and hours choice. Therefore estimating the labor supply schedules directly will result in biased estimates of the structural parameters. Following MaCurdy, I estimate the models in first differences to get rid of any unobserved person specific fixed effect. The price for estimating these models in first
differences of the data is that the estimates may be imprecise.\footnote{An alternative is to follow a suggestion of Chamberlain (1984) and assume that the expectation of the fixed effect is linear in the regressors. Then higher moments of the regressors can be used as instruments in a multivariate regression with data in the levels. If the fixed effect is the marginal utility of wealth, this assumption is likely to be false.} As an example, I estimate Model A in the form

\begin{align}
(3.1) & \quad \Delta \ln W_{it} = \Delta X_{it} - 5t \Delta (H_{it} - M_{it})^2 + \Delta q_{it} \\
(3.2) & \quad \Delta \ln H_{it} = (\sigma - 1)^{-1} \Delta \ln \tilde{W}_{it} - \tau (\sigma - 1)^{-1} \Delta (H_{it} - M_{it})H_{it} - \Delta \mu_{it}' \\
& \quad \Delta Z_{it} = z_{it} - z_{it-1}
\end{align}

In (3.1), I have tacked an error term, $q_{it}$, to the end of the equation to account for factors that affect the measured wage that are unobserved by the econometrician.

The third issue concerns endogeneity and measurement error in observed wages and hours. Given the difficulties of estimating hedonic equation models in general (see e.g., Brown (1983), Epplle (1987)), my strategy is to aim for consistency rather than efficiency. In estimating the wage equations (e.g., (3.1)), I need instruments for the hours deviations that are not in the $X_{it}$'s. Let $SW_{it}$ be a dummy variable that takes the value of 1 if the worker changed industry and or occupation between $t$ and $t-1$. According to the theory, a mismatch in hours is costly. Therefore some workers may change industry or occupation to reduce the mismatch (Evidence of this behavior will be documented in Section VI of this paper). Thus $SW_{it}$ should be negatively
correlated with $\Delta(H_{it}-M_{it})^2$ and be a good instrument for it. In order to control for the fact that job changes increases the wage, I also include a job change dummy variable in the wage regressions. Another instrument for the hours deviation is the number of children. Note that measurement error in $H_{it}$ will not be correlated with the instruments. Other regressors in the wage equations are age, job tenure, education, and some interactions.

In the hours equation (3.2), I need instruments that are correlated with $\Delta(H_{it}-M_{it})H_{it}$ and not with changes in taste for hours of work, $\Delta\mu_{it}$. The prime candidates are $M_{it}$ and its lag. Measurement error in $H_{it}$ will not be correlated with $M_{it}$ because $M_{it}$ is constructed for each worker by excluding the hours worked for that worker. Moreover since $M_{it}$ is the mean of $H_{it}$, the measurement errors in $H_{it}$ are averaged out in $M_{it}$ resulting in a much cleaner measure of $M_{it}$ even if each $H_{it}$ is subject to measurement error. Other instruments in the hours equation are age, job tenure, education, and some interactions. So I identify the structural coefficients in my estimation by both functional form restrictions and choice of instruments.

As a check on the exogeneity assumptions, I test for the exogeneity of the instruments in the regression equations with a Chi-square test (see e.g. Hausman (1984)).

Finally, in order to maximize the number of observations for estimation, I include individuals who may not be present for all years in the sample. The price for this inclusion is that fully efficient estimation via the optimum minimum distance estimators (see e.g. Chamberlain) is too expensive to

---

9 Permanent differences across people are differenced out in the regressions. It is unlikely that transitory changes in individual tastes for hours are correlated with $M_{it}$ or its lag.
compute I use the two stage least squares estimator and correct the standard errors of the estimates for correlated errors across equations

IV. Data

The data is from the 1971-1981 panels of the Michigan Panel Study of Income Dynamics (PSID) This data set was used by MaCurdy, Altonji, and Ham to estimate marginal utility of wealth constant labor supply functions So the results from this study can be compared with the earlier work

I construct two different data sets for this study Let the main data set be D1 For each year Only married male heads of households are considered Observations were excluded if the worker is below 22 or above 60 years old, he had less than 15 or more than 5000 hours of work on the main job, he was retired, permanently disabled or a student

\( M_{it} \) for worker i is the mean annual hours of work of all workers, excluding worker i, in the data set in his occupation, industry and year If the mean is constructed from less that five workers, that observation is excluded from the analysis Occupations are at the one digit level

\[ \text{Note that (3.2) is linear in the variables and only non-linear in the parameters It is also just identified So the point estimates of the parameters and their standard errors from two stage least squares are identical to that from non-linear two stage least squares The savings in computation time from two stage least squares is overwhelming} \]
Industries are at the two digits level. 11

\(H_{it}\), the annual hours of work, is constructed by multiplying weeks worked on the main job and average hours per week.

\(W_{it}\) is the reported hourly wage divided by the implicit price deflator for consumption expenditures. There is a problem of timing with the data because \(W_{it}\) is measured in March whereas \(H_{it}\) refers to that in the preceding year. The wage measure is available before 1976 only for non-salaried workers, and from 1976 on also for salaried workers.

Since I am worried about workers' misreporting their industry, I construct a smaller data set D2. A worker is included in D2 at \(t\) and \(t-1\) if his industry at \(t\) is equal to \(t+1\), and also \(t-1\) is equal to \(t-2\). The advantage of D2 is that misreporting is likely to be minimized. The disadvantage is that I may be throwing out true mobility, thereby reducing variation in the data. I am less worried about occupational misreporting because it is at the one digit level.

I also used average hours worked per week as a measure of \(H_{it}\). The rationale for using weekly hours of work is to make sure that the results using annual hours of work were not primarily driven by unemployment behavior since most of the variation in annual hours of work consist of variation in days worked. The results with weekly hours are contained in an earlier draft.

---

11 The occupations are professionals, managers, clerical and sales, craftsmen and foremen, operatives, laborers, and service workers. The industries are mining, manufacturing durables (30,31,32,33), manufacturing non-durables (40,41,42,43,44,45,49), construction, transportation, communication, other public utilities, retail trade, wholesale trade, other trade, finance insurance and real estate, repair services, business services, personal services, recreation services, publishing, health services, education services and other professional services. These categories are those defined in the PSID codebook.
of this paper, Siow (1988) They are consistent with the results presented here I also used annual earnings per hour as an alternative wage measure The results were similar to that of using wages per hour and are not reported here

Summary statistics of the data set is in the Appendix

V. Empirical Results

In order to relate the findings of Moffitt, Biddle and Zarkin to my work, let workers be penalized only for choosing hours that deviated from that of the mean of the population, $M$. The wage schedule in (2.5) for worker $i$ becomes

$$
\ln W_{it} = X_{it} - 5\tau (H_{it} - M)^2 + q_{it}
$$

(5.1) can be rewritten as

$$
\ln W_{it} = X_{it}' + a_1 H_{it} - a_2 H_{it}^2 + q_{it}
$$

(5.2) 

In (5.2) the log wage is a quadratic function of log hours of work Estimates of $\tau$ and $M$ can be obtained from estimates of $a_1$ and $a_2$. Biddle and Zarkin use log wage as a quadratic function of log annual hours, that is (5.2) where they use $\ln H_{it}$ rather than $H_{it}$, to derive a cross-section annual
earnings regression \(^{12}\) Their preferred estimate of \(M\) is 2617 hours. Their implied estimate of \(r\) says that a 10\% deviation in hours from \(M\) reduces the wage by 6.5\%. Moffitt estimates a cross-section wage function which is quadratic in weekly hours jointly with labor supply. He finds the peak of the wage schedule at 34 hours. The results in both papers clearly suggest that the wage schedule is quadratic in hours. Moreover, their estimates, using different data and estimation techniques, are consistent with my interpretation of their models \(^{13}\).

One question that arises from the above discussion is whether worker \(i\) is penalized for hours deviations around an industry, occupation and year mean, \(M_{it}\). Perhaps workers are only penalized for deviations from a grand mean, \(M\), as in (5.1) In order to shed light on this issue, note that (5.1) may also be rewritten as

\[
(5.3) \quad \ln W_{it} = X_{it} - 5r(H_{it} - M_{it})^2 + r(H_{it} - M_{it})(M_{it} - M) - 5r(M_{it} - M)^2 + q_{it}
\]

Take the expectation of equation (5.3) by industry, occupation and year to get

\[
(5.4) \quad \ln \underline{W}_{it} = \underline{X}_{it} - 5r \text{Cov}(H_{it}) - 5r(\underline{M}_{it} - M)^2
\]

\(\underline{Z}_{it}\) Mean of \(Z_{it}\) by industry, occupation and year

---

\(^{12}\) This is equivalent to my equation (2.8) where \(\ln M\) is substituted for \(\ln M_{it}\).

\(^{13}\) In conversation, Zarkin interprets their results by appealing to Duncan and Stafford's theory of coordination. Moffitt uses Barzel's theory.
\( \text{Cov}(H_{it}) \)  Variance of \( H_{it} \) by industry, occupation and year

Now consider the following regression where an observation is an industry, occupation and year

\[
\ln W_{it} = X_{it}' + a_3 \text{Cov}(H_{it}) + a_4 (M_{it} - M)^2 + q_{it}'
\]

\( X_{it}' \)  Part of \( X_{it} \) that are observed by the econometrician

\( q_{it}' \)  Part of \( X_{it} \) that are not observed by the econometrician

If the grand mean theory is correct, that is equation (5.1), then \( a_3 = a_4 < 0 \)  If my theory is correct, then equation (2.5) holds and it will imply that \( a_4 = 0 \) in equation (5.5)  If one believes that the penalty is for hours deviations around an industry and occupation mean, \( M_i \), one can derive an equivalent equation to (5.5)

\[
\ln W_{it} = X_{it}' + a_5 \text{Cov}(H_{it}) + a_6 (M_{it} - M_i)^2 + q_{it}'
\]

If the penalty is for deviations around an industry and occupation mean, \( a_5 = a_6 < 0 \)  If the penalty is for deviations around an industry, occupation and year mean, \( a_6 = 0 \)  Note that Barzel's theory suggest that the penalty is for deviations around an industry and occupation mean  A worker's wage is low if his hours is below the industry and occupation mean because of the fixed cost of employment  A worker's wage is also low if his hours is above the industry and occupation mean due to fatigue from long hours  There is no role
for a time varying mean in Barzel's theory. So his theory suggest that \( a_5 = a_6 \).

Table 1 contains OLS regressions of equations (5.5) and (5.6). In column 1 to 4, there are no \( x_{it} \)'s except for \( M_{it} \) in some of the regressions. This is not a problem if the regressors that we are interested in are not correlated with the left out variables. If there is a correlation, we will have bias estimates of the structural coefficients. In order to deal with the possible problem of left out variables, I estimate the equations also in first differences in columns 5 to 8. The first differencing also reduces measurement error in \( \text{Cov}(H_{it}) \).

First note that the coefficient for \( \text{Cov}(H_{it})/10^6 \) is negative and significantly different from zero in all 8 regressions. The estimates of \( \tau \) are different for the levels regressions (Col 1-4) versus the first differenced regressions (Col 5-8) suggesting bias for the levels regressions. If we use my theory as the null hypothesis, then we cannot reject the null that \( a_4 \) and \( a_6 \) are equal to zero in columns 2, 3, 4, 6, 7 and 8. If Barzel's theory is the alternative hypothesis, then it is rejected. For estimates of \( \tau \) of 8 and 2, an increase of 402 hours (1 standard deviation or 20%) away

---

14 A time varying mean can be obtained from the need to coordinate hours of work across workers to reduce the fixed cost of employment (e.g., switching on the air-condition in a building). Then his theory merges into mine. As such, I do not believe that his theory and mine are mutually exclusive.

15 \( H_{it} \) and \( M_{it} \) are divided by 1000 for estimation.

16 Let measured hours, \( H_{it} \), be equal to true hours, \( H_{it}^* \) and an additive error term, \( U_{it} \), that is uncorrelated with \( H_{it}^* \). Then \( \text{Cov}(H_{it}) = \text{Cov}(H_{it}^*) + \text{Var}(U_{it}) \). If \( \text{Var}(U_{it}) \) is equal to \( \text{Var}(U_i) \), that is the variance of the measurement error is constant across years, \( \text{Cov}(H_{it}) = \text{Cov}(H_{it}^*) \).

17 For convenience, I use the 90% confidence interval for testing hypotheses in this paper.
From the mean hours of his co-workers will reduce his wage by 6.5% and 1.6% respectively. These estimates seem reasonable in the sense that total earnings are still increasing for this worker although at a declining rate. However, due to the large standard errors in estimating \( a_4 \) and \( a_6 \), if we use the grand mean or only industry and occupational mean hypotheses, they also cannot be rejected in columns 4, 6, 7, and 8. In summary, I conclude from Table 1 that there is evidence to support the hypothesis that workers are penalized for hours deviations around an industry, occupation and year mean.

The results in Table 1 also shed light on two other issues. First, Model C in my paper (equation 2.10) is probably too crude. If Model C is correct, it implies that the coefficients for \( \text{Cov}(H_{it})/10^6 \) are zeros in the regressions in Table 1. Second, the coefficients for \( \text{Cov}(H_{it})/10^6 \) are systematically negative in Table 1. These negative coefficients make it difficult to fit models that explain wage differentials across industries as a compensating differential for employment variations.\(^{18}\)

There are two implications from my theory that relates to the mobility of workers. If workers are penalized for choosing hours of work that deviated from those of their peers, they may change industry or occupation to find a more compatible set of peers and reduce the penalty associated with their initial choice of hours. Of course, if they receive substantial rent from their initial industry or occupation, or if they perceive the deviations to be

\(^{18}\) If employment variations are undesired by workers, industries that have substantial employment variations will pay wage premiums to attract workers. These models are investigated in Abowd and Ashenfelter (1981), Murphy and Topel (1987). Their mixed results are consistent with the evidence in Table 1.
transitory, they will not move ¹⁹ Let SWᵢᵣ be equal to 1 if a worker switches
industry or occupation between t-1 and t, and zero otherwise. In Table 2,
Column 7 presents a logit regression of SWᵢᵣ on 1000|Hᵢᵣ₋₋₁-Mᵢᵣ₋₋₁|, the
absolute value of 1000x(Hᵢᵣ-Mᵢᵣ). The estimated coefficient for 1000|Hᵢᵣ₋₋₁-
Mᵢᵣ₋₋₁| is 460 and significant. The estimate implies that a one standard
deviation reduction in measured hours away from mean hours will increase
the probability of the worker switching industry and/or occupation by 0.41 (12%)
While the estimate is supportive of the theory, an alternative hypothesis is
that the result is driven by workers who have hours less than those of their
peers. ²⁰ In column 8, I decompose the deviations into negative and positive
deviations 1000|Hᵢᵣ-Mᵢᵣ|(+)=1000x(Hᵢᵣ-Mᵢᵣ) if Hᵢᵣ-Mᵢᵣ >0, = 0 otherwise
100|Hᵢᵣ-Mᵢᵣ|(-) is defined likewise. The coefficient on the positive
development is 311 and significantly different from zero. The coefficient on
the negative deviations is 601 and also significantly different from zero.
While the workers with negative deviations are more likely to move, workers
with positive deviations may also move. Another issue that arise is
measurement error in the reported industry of the workers. Workers may report
a change in industry even when they did not change or vice versa. Therefore I
reran the regressions with data set D2 where a worker is included in the data
set at t-1 and t if his industry is the same for t and t+1, and also the same
for t-1 and t-2. In the D1 data set (columns 7 and 8), the mean of SWᵢᵣ was

¹⁹ An earlier draft of this paper found evidence of substantial
transitory hours deviations

²⁰ Robert Topel suggested that the unemployed, those with hours less
than the mean hours in an industry, occupation and year, were driving some of
the results in an earlier draft of this paper. E.g. the result in column 7
may be due to only the unemployed leaving their industry or occupation.
In D2, it was 239 which is less than that of D1. The results with this smaller data set is in columns 9 and 10. The point estimates are slightly smaller and the standard errors of the estimates are larger than that with the larger data set. Still, all the estimates still have the correct sign and they are also significantly different from zero. Moreover, the distance between the estimated coefficients for the positive and negative deviations is smaller in column 10. For the estimate in column 9, a one standard deviation departure in hours from the mean of his co-workers will increase the probability of a worker leaving his industry or occupation by 03 (13%). A caveat is necessary in interpreting the quantitative magnitudes of the estimates. Since $H_{it-1}$ is likely to be measured with error, the point estimates are probably biased towards zero. It is all the more reassuring that the estimates are statistically different from zero.

Another test of the hypothesis on mobility is to look at the difference in hours choice when workers move. The model argues that part of the reason why workers move is to reduce their deviations in hours. We first use the data set D1. In Table 2, column 1, the first difference of the absolute deviation in hours, $\Delta|H_{it} - M_{it}|$, is regressed on $SW_{it}$. The coefficient for $SW_{it}$ is negative as predicted and significantly different from zero. It says that when a worker switches industry or occupation, he reduces his deviations on average by 24 hours. I then broke up the sample into two groups, those whose deviations in hours in $t-1$ were negative versus those that were positive. Column 2 contains the result for those workers who had positive deviations in hours in $t-1$. The coefficient for $SW_{it}$ is -15.5 and almost significantly different from zero. For the other group of workers, the coefficient for $SW_{it}$ in column 3 is -31.4 and significantly different from zero.
The estimated mean reduction in the post move deviation is twice as large for workers with initial negative deviations that for workers with initial positive deviations. Again to check against measurement error in industry reporting, I reran the regressions using data set D2. The estimated coefficients on $SW_{it}$ in columns 4, 5 and 6 are all negative and significant. There is no systematic difference in these regressions versus the ones using data set D1. Measurement error in the hours variables do not affect the consistency of the estimates in these regressions because the hours variables are the dependent variables. To summarize, the results on mobility are supportive of the theory: Workers that initially choose hours of work that deviate from that of their co-workers have a tendency to move to reduce those deviations. Workers with negative deviations are more likely to move and reduce their post move deviations by a larger margin than workers with positive deviations.21

Table 3 contains structural estimates of Models A, B, and C.22 Consider the estimates in column 1 (equations (2.5) and (2.7)). From the wage equation (i.e. (2.5)), I estimated a $\tau$ of 4209 with a standard error of 278. This estimate of $\tau$ would imply a reduction of 2.4% in the wage if a worker reduces his hours of work one standard deviation from that of his co-workers (20%). From the hours equation (i.e. (2.7)), I get an estimate of $\tau$ of 07883 with a standard error of 0488. Both estimates of $\tau$ have the right signs. Using the Wald test, I cannot reject the hypothesis that they are the same.

---

21 It is unclear if the compensating wage differential for employment variation argument to explain interindustry wage variation is consistent with these mobility results.

22 $H_{it}$ and $M_{it}$ are divided by 1000 each for estimation.
However due to the wide standard errors, it is also difficult to reject the hypothesis that both coefficients are zeros. The estimated wage coefficient from the hours equation, \((\sigma-1)^{-1}\), was 3772. This point estimate is similar to that obtained by Altonji, Ham and McCurdy.

In column 2, I present estimates for Model B (equations (2.8) and (2.9)). From the wage equation, I estimated \(\tau\) of 3856 with a standard error of 3719. The hours equation gave an estimate of \(\tau\) of 3403 with a standard error of 2004. If \(\tau\) is 35, a reduction in hours of one standard deviation away from the mean (20%) would reduce the wage by 7%. The estimate of the wage coefficient is similar to that in column 1.\(^{23}\) Given the similarity of the point estimates, the Wald test cannot reject the null hypothesis of equality of the estimates from the wage and hours equations.

In column 3, I present estimates for model C (equations (2.10) and (2.11)). The estimate of \(\tau\) from the wage equation is 1422 which implies a 7% reduction in wage for a 20% increase in hours from mean hours. From the hours equation, we get an estimate of \(\tau\) of 0.179 with a large standard error. The wage coefficient in the hours equation is about the same as obtained earlier.

The tests for exogeneity showed that the instruments were exogenous in the wage equations except for model B. It also showed that exogeneity is rejected in the hours equations. Fortunately, the exogeneity assumption in

\(^{23}\) I was surprised not to find larger estimates of the intertemporal labor supply elasticity relative to those obtained by Altonji and McCurdy. One explanation is that they have a constant and sometimes year dummies in the labor supply equations. If the penalty for hours deviations is well approximated by deviations to the grand mean of the population, \(M\), or a time varying grand mean, \(M_t\), then the right hand side of equation (2.9) in the text reduces to a function of the wage, intercept and year effects. Note that there is no justification for year effects in the labor supply equation in their formal framework.
the hours equations is not rejected for the other data sets

My conclusion from Table 3 is that there is evidence in the data that supports the theory. The estimated penalties for deviations in hours worked are sensible economically since we do observe workers at 1 standard deviation away from mean hours of work.

Table 4 consists of the same regressions as Table 3 except that it uses the data set D2. The estimates of $r$ from the wage equations are larger than that of Table 2. For example, from the wage equation in column 1, the estimate of $r$ implies that the wage falls by 5.7% for a one standard deviation departure of a worker's hours from mean hours. The standard errors of the estimates are about the same as that of Table 2. The tests of the exogeneity of the instruments for the wage equations are not rejected. The estimates of $r$ and its standard errors from the hours equations are about the same as in the earlier Table. The Wald tests suggest that the null hypothesis of equality cannot be rejected in model B, and is just rejected in models A and C. The estimate of the wage elasticity, $(\sigma-1)^{-1}$, is about 41 in the three models. This estimate is larger than that in the earlier table and is marginally larger than those found by Altonji, Ham and McCurdy. The tests of exogeneity of the instruments in the hours equations were not rejected in all three equations.

One restriction that I imposed on the wage equations for analytical convenience was that the penalty for hours deviations was symmetric around $M_{it}$. In order to test this restriction, I decomposed $(H_{it}-M_{it})^2/10^6$ into two

---

24 These rejections show that there is power in tests of the overidentifying restriction even with the large standard errors of the estimates.
components, $H^+$ and $H^-$. $H^+$ is equal to $(H_{it} - M_{it})^2/10^6$ when it is positive and zero otherwise. $H^-$ is equal to $(H_{it} - M_{it})^2/10^6$ when it is negative and zero otherwise. I then estimated the wage equations in column 1 of Table 3 and 4 again with $H^+$ and $H^-$ in the equations instead of $(H_{it} - M_{it})^2/10^6$. For data set D1, the estimates of the coefficients for $H^+$ and $H^-$ were -2582 and -1758 with standard errors of 2555 and 2094 respectively. For data set D2, the coefficient estimates of $H^+$ and $H^-$ were -3630 and -3422 with standard errors of 2353 and 2559 respectively. From the above estimates, there is little evidence that the estimates of $\tau$ in the wage regressions in Table 3 and 4 are driven primarily by workers whose hours are below that of the mean of their co-workers.

The results from Tables 1, 2, 3 and 4, reflecting different tests of the theory, different estimation techniques and different selection criterion for creating data sets, are supportive of the theory. Note that whenever possible, I have tested whether the results were driven by the restriction that the wage penalty is symmetric around the mean hours of the worker’s co-workers. In all the tests where symmetry was relaxed, I found that workers tried to reduce their absolute deviations and workers were penalized also for absolute deviations as predicted by the theory. The exogeneity tests in the structural estimation were rejected less often in the more restrictive sample D2, suggesting that the rejections in the larger sample is due more to measurement error rather than endogeneity of the instruments. Due to the lack of valid instruments, the estimates of $\tau$ were not precise. Since I was primarily concerned with consistency of the estimates, I had little freedom in choosing the instruments. Finally, there is little evidence to suggest that
Model C better describes the data than the other two models Perhaps my attempt to allow for variation in the penalty rate by industry and occupation was too crude

VII Conclusions

The coordination of hours of work is an important phenomenon in most economies What this paper shows is that wages are used to at least partially achieve that coordination Given the measurement error in measured industry cells, the results from the smaller data sets are more reliable A worker’s wage falls between 1 to 7% when his hours of work differ from that of his co-workers by one standard deviation (20% in annual hours) Although the point estimates of the penalty for the lack of coordination are imprecise, there is little support for the hypothesis that different industries and occupations have different penalty rates Estimates of the labor supply elasticities were not much larger than those found in the earlier literature which ignored the coordination issue So even though wages are used to coordinate hours of work, the small labor supply elasticities are not explained by my investigation

25 Other factors include the government that defines the standard work week
REFERENCES


Cooper, Russell and Andrew John (1985), "Coordinating Coordination Failures in Keynesian Models", Cowles Foundation Discussion Paper 745, Yale University


 Economides, Nicholas and Aloysius Siow (1988), "The Division of Markets is Limited by the Extent of Liquidity", *American Economic Review*, Vol 78, No 2, pp 108-121


Heller, Walter P (1985), "Coordination Failure Under Complete Markets with Application to Effective Demand", University of California at San Diego Working Paper


Basil Blackwell


Owen, John D (1979), *Working Hours*, Lexington Books


Siew, Aloysius (1988), "The Use of Wages in Coordinating Hours of Work", manuscript
### Table 1

Mean Log Wage Regressions using Annual Hours

<table>
<thead>
<tr>
<th>Dep Var</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>lnWit</td>
<td>lnWit</td>
<td>lnWit</td>
<td>lnWit</td>
<td>lnWit</td>
</tr>
<tr>
<td>Cov(Hit)/10^6</td>
<td>-4284</td>
<td>-4659</td>
<td>-4184</td>
<td>-3879</td>
</tr>
<tr>
<td></td>
<td>(0793)</td>
<td>(0829)</td>
<td>(0801)</td>
<td>(0815)</td>
</tr>
<tr>
<td>(Mit-M)^2/10^6</td>
<td>3640</td>
<td>0402</td>
<td>2308</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2354)</td>
<td>(2308)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Mit-Mi)^2/10^6</td>
<td></td>
<td></td>
<td>-1152</td>
<td>(5547)</td>
</tr>
<tr>
<td>Mit/10^3</td>
<td></td>
<td></td>
<td>4546</td>
<td>(0592)</td>
</tr>
</tbody>
</table>

| r       | 85    | 93    | 83    | 77    |

| Equality test^2 | 99 | 91 | 81 |
| R^2             | 0371 | 0402 | 1097 | 0426 |
| DF              | 756  | 755  | 754  | 755  |

<table>
<thead>
<tr>
<th>Dep Var</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>lnWit</td>
<td>lnWit</td>
<td>lnWit</td>
<td>lnWit</td>
<td>lnWit</td>
</tr>
<tr>
<td>Cov(Hit)/10^6</td>
<td>-1162</td>
<td>-1037</td>
<td>-1131</td>
<td>-1096</td>
</tr>
<tr>
<td></td>
<td>(0294)</td>
<td>(0302)</td>
<td>(0306)</td>
<td>(0309)</td>
</tr>
<tr>
<td>(Mit-M)^2/10^6</td>
<td>1820</td>
<td>1839</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1014)</td>
<td>(1013)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Mit-Mi)^2/10^6</td>
<td></td>
<td></td>
<td>-1320</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1888)</td>
<td></td>
</tr>
<tr>
<td>Mit/10^3</td>
<td></td>
<td></td>
<td>-06243</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(03464)</td>
<td></td>
</tr>
</tbody>
</table>

| r       | 23    | 21    | 23    | 22    |

| Equality test | 52 | 47 | 9 | 09 |
| R^2           | 0255 | 0308 | 0360 | 0263 |
| DF            | 595  | 594  | 593  | 594  |

---

1 Inconcepts included in all regressions  Standard errors are in parenthesis

2 Marginal confidence level for test of either (Mit-M)^2 = Cov(Hit) or (Mit-Mi)^2 = Cov (Hit)
### Table 2

Deviations in Annual Hours and Mobility

<table>
<thead>
<tr>
<th>Method</th>
<th>Dep Var</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Δ</td>
<td>(H_{it}-M_{it})</td>
<td></td>
<td>Δ</td>
</tr>
<tr>
<td>SWit</td>
<td></td>
<td>-23.59</td>
<td>-15.50</td>
<td>-31.43</td>
<td>-30.49</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(5960)</td>
<td>(241)</td>
<td>(603)</td>
<td>(646)</td>
</tr>
<tr>
<td>R(^2)</td>
<td></td>
<td>0.0009</td>
<td>0.004</td>
<td>0.015</td>
<td>0.016</td>
</tr>
<tr>
<td>DF</td>
<td></td>
<td>17503</td>
<td>8662</td>
<td>8839</td>
<td>9854</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td>OLS</td>
<td>OLS</td>
<td>LOGIT</td>
</tr>
<tr>
<td>Dep Var</td>
<td>Δ</td>
<td>(H_{it-1}-M_{it-1})</td>
<td></td>
</tr>
<tr>
<td>SWit</td>
<td>-21.38</td>
<td>-39.97</td>
<td>4597</td>
</tr>
<tr>
<td></td>
<td>(10.68)</td>
<td>(9.68)</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>(H_{it-1}-M_{it-1})</td>
<td></td>
<td>3111</td>
</tr>
<tr>
<td>1000</td>
<td>(H_{it-1}-M_{it-1})(^+)</td>
<td></td>
<td>6008</td>
</tr>
<tr>
<td>1000</td>
<td>(H_{it-1}-M_{it-1})(^-)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R(^2)</td>
<td>0.008</td>
<td>0.027</td>
<td>104.6</td>
</tr>
<tr>
<td>DF</td>
<td>4940</td>
<td>4912</td>
<td>1</td>
</tr>
<tr>
<td>NO OF OBS</td>
<td>18161</td>
<td>18161</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(9)</th>
<th>(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td>LOGIT</td>
</tr>
<tr>
<td>Dep Var</td>
<td>SWit</td>
</tr>
<tr>
<td>1000</td>
<td>(H_{it-1}-M_{it-1})</td>
</tr>
<tr>
<td>1000</td>
<td>(H_{it-1}-M_{it-1})(^+)</td>
</tr>
<tr>
<td>1000</td>
<td>(H_{it-1}-M_{it-1})(^-)</td>
</tr>
<tr>
<td>CHI-SQUARE</td>
<td>31.52</td>
</tr>
<tr>
<td>DF</td>
<td>1</td>
</tr>
<tr>
<td>NO OF OBS</td>
<td>10142</td>
</tr>
</tbody>
</table>

---

\(^3\) Intercepts included in all regressions. Standard errors are in parenthesis.
### Table 3

Structural Estimation with Data Set D1

<table>
<thead>
<tr>
<th>Model</th>
<th>A&lt;sup&gt;5&lt;/sup&gt;</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔlnW&lt;sub&gt;it&lt;/sub&gt; Regression&lt;sup&gt;6&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;i&gt;τ&lt;/i&gt;</td>
<td>4209</td>
<td>3856</td>
<td>1422</td>
</tr>
<tr>
<td></td>
<td>(2780)</td>
<td>(3719)</td>
<td>(09345)</td>
</tr>
<tr>
<td>MSE</td>
<td>04497</td>
<td>04187</td>
<td>05602</td>
</tr>
<tr>
<td>EX TEST&lt;sup&gt;7&lt;/sup&gt;</td>
<td>8969</td>
<td>9528</td>
<td>7697</td>
</tr>
<tr>
<td>ΔlnH&lt;sub&gt;it&lt;/sub&gt; Regression&lt;sup&gt;8&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;i&gt;τ&lt;/i&gt;</td>
<td>07883</td>
<td>3403</td>
<td>01787</td>
</tr>
<tr>
<td></td>
<td>(04885)</td>
<td>(2004)</td>
<td>(01091)</td>
</tr>
<tr>
<td>(&lt;i&gt;σ-1&lt;/i&gt;)&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>3773</td>
<td>3539</td>
<td>3876</td>
</tr>
<tr>
<td></td>
<td>(1827)</td>
<td>(1711)</td>
<td>(1860)</td>
</tr>
<tr>
<td>MSE</td>
<td>08447</td>
<td>07393</td>
<td>08746</td>
</tr>
<tr>
<td>EX TEST</td>
<td>9889</td>
<td>9899</td>
<td>9849</td>
</tr>
<tr>
<td>WALD TEST&lt;sup&gt;9&lt;/sup&gt;</td>
<td>7733</td>
<td>08421</td>
<td>8131</td>
</tr>
</tbody>
</table>

---

<sup>4</sup> There were 7823 observations in all regressions. Standard errors are in parentheses. H<sub>it</sub> and M<sub>it</sub> divided by 1000 each.

<sup>5</sup> H<sub>it</sub> and M<sub>it</sub> divided by 100 for estimation of model A.

<sup>6</sup> Other regressors included intercept, job change dummy, education, age, age squared, and tenure. Instruments were no of kids, no of kids under 6, education squared, age*education, age*education squared, SW<sub>it</sub>.

<sup>7</sup> Chi square test of exogeneity of instruments. Marginal confidence interval is reported.

<sup>8</sup> Other regressor was an intercept. Instruments were lnH<sub>it</sub>, lnH<sub>it-1</sub>, education, age, age squared, tenure, education squared, age*education, age*education squared.

<sup>9</sup> Marginal confidence level of Wald Test of equality of <i>τ</i> from the hours and wage equations (Chi-square with 1 df).
Table 4

Structural Estimation with Data Set D2\(^{10}\)

<table>
<thead>
<tr>
<th>Model</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A(^{11})</td>
<td>B</td>
<td>C</td>
</tr>
</tbody>
</table>

\(\Delta \ln W_{it}\) Regression\(^{12}\)

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(\tau)</td>
<td>7073</td>
<td>6140</td>
<td>1736</td>
</tr>
<tr>
<td></td>
<td>(3722)</td>
<td>(4059)</td>
<td>(08969)</td>
</tr>
<tr>
<td>MSE</td>
<td>04311</td>
<td>03365</td>
<td>04915</td>
</tr>
<tr>
<td>EX TEST(^{13})</td>
<td>6265</td>
<td>8988</td>
<td>4678</td>
</tr>
</tbody>
</table>

\(\Delta \ln H_{it}\) Regression\(^{14}\)

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(\tau)</td>
<td>08749</td>
<td>3751</td>
<td>02272</td>
</tr>
<tr>
<td></td>
<td>(06842)</td>
<td>(2819)</td>
<td>(01953)</td>
</tr>
<tr>
<td>((\sigma-1)^{-1})</td>
<td>4350</td>
<td>4043</td>
<td>4040</td>
</tr>
<tr>
<td></td>
<td>(2894)</td>
<td>(2650)</td>
<td>(2991)</td>
</tr>
<tr>
<td>MSE</td>
<td>06676</td>
<td>05644</td>
<td>07186</td>
</tr>
<tr>
<td>EX TEST</td>
<td>6297</td>
<td>6237</td>
<td>5337</td>
</tr>
<tr>
<td>WALD TEST(^{15})</td>
<td>9000</td>
<td>3687</td>
<td>9023</td>
</tr>
</tbody>
</table>

\(^{10}\) There were 4462 observations in all regressions. Standard errors are in parentheses. \(H_{it}\) and \(M_{it}\) divided by 1000 each.

\(^{11}\) \(H_{it}\) and \(M_{it}\) divided by 100 for estimation of model A.

\(^{12}\) Other regressors included intercept, job change dummy, education, age, age squared and tenure. Instruments were no of kids, no of kids under 6, education squared, age*education, age*education squared, \(SW_{it}\).

\(^{13}\) Chi square test of exogeneity of instruments. Marginal confidence interval is reported.

\(^{14}\) Other regressor was an intercept. Instruments were \(\ln H_{it}\), \(\ln H_{it-1}\), education, age, age squared, tenure, education squared, age*education, age*education squared.

\(^{15}\) Marginal confidence level of Wald Test of equality of \(\tau\) from hours and wage equations (Chi-square with 1 df).
### Appendix

**Summary Statistics of Data Set D1**

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W_{it}$</td>
<td>4.233</td>
<td>1.927</td>
</tr>
<tr>
<td>$H_{it}$</td>
<td>2022</td>
<td>402.3</td>
</tr>
<tr>
<td>$M_{it}$</td>
<td>1999</td>
<td>169.7</td>
</tr>
<tr>
<td>$SW_{it}$ (D1)</td>
<td>3459</td>
<td></td>
</tr>
<tr>
<td>$SW_{it}$ (D2)</td>
<td>2246</td>
<td></td>
</tr>
<tr>
<td>$\Delta \ln W_{it}$</td>
<td>0.1653</td>
<td>1.964</td>
</tr>
<tr>
<td>$\Delta \ln H_{it}$</td>
<td>0.0000</td>
<td>2.646</td>
</tr>
<tr>
<td>$\Delta \ln M_{it}$</td>
<td>-0.01043</td>
<td>1.004</td>
</tr>
</tbody>
</table>