WORKING PAPER #246
INDUSTRIAL RELATIONS SECTION
PRINCETON UNIVERSITY
JANUARY 1989

Minimum Chi-Square and Three-Stage Least Squares in Fixed Effects Models

by

Joshua D. Angrist and Whitney K. Newey*

January 1989

*Industrial Relations Section and Department of Economics, Princeton University.
Angrist thanks the Industrial Relations Section, the Department of Economics and the Sloan Foundation for financial support. Newey thanks the Sloan Foundation, the National Science Foundation and Bell Communications Research.

Minimum Chi-Square and Three-Stage Least Squares in Fixed Effects Models

Abstract

Two approaches to estimation and testing of fixed effects models are commonly found in the econometrics literature. The first involves variations on instrumental variables. The second, a Minimum Chi-Square (MCS) procedure introduced by Chamberlain, minimizes a quadratic form in the difference between unrestricted regression coefficients and the restrictions implied by the fixed effects model. This paper is concerned with the relationship between Three-Stage Least Squares (3SLS) and MCS. A 3SLS equivalent of the MCS estimator is presented and, in the usual case wherein the time varying error component has a scalar covariance matrix, 3SLS is shown to simplify to the conventional deviations from means estimator. Furthermore, the corresponding over-identification test statistic is the degrees of freedom times the R^2 from a regression of residuals on all leads and lags of right hand side variables. The relationship between MCS and some recently introduced efficient instrumental variables procedures is also considered.

An empirical example from the literature on life-cycle labor supply is used to illustrate properties of 3SLS procedures for panel data under alternative assumptions regarding residual covariance. Estimated labor supply elasticities and standard errors appear to be insensitive to these assumptions. In contrast, the over-identification test statistics are found to be substantially smaller when residuals are allowed to be intertemporally correlated and heteroscedastic. At conventional levels of significance, however, even the smallest of the test statistics leads to rejection of the over-identifying restrictions implicit in the labor supply models.

Joshua D. Angrist
Industrial Relations Section
Firestone Library
Princeton University
Princeton, NJ 08544

Whitney K. Newey
Department of Economics
Princeton University
Princeton, NJ 08544

I. Introduction

In two influential papers, Chamberlain (1982, 1984) developed the theory of multivariate regression models for panel data. An important contribution of this work was to point out that fixed effects specifications impose testable restrictions on coefficients from regressions of all leads and lags of dependent variables on all leads and lags of independent variables. Chamberlain suggested that estimation and testing be carried out by minimizing an optimally weighted quadratic form in the difference between unrestricted parameters and the restrictions implied by the fixed effects model. In the simultaneous equations context, this approach has been referred to as Minimum Chi-Square or MCS (Rothenberg 1973).[2]

In addition to MCS procedures, Chamberlain also considers a generalization of Three-Stage Least Squares (3SLS), demonstrating that there are 3SLS estimators asymptotically equivalent to MCS. The practical implications of this relationship are left somewhat unclear, however, and the purpose of this paper is to explore the relationship between MCS and 3SLS in greater detail. We consider both estimators and test statistics, beginning with an algebraic derivation of the 3SLS equivalent of MCS for fixed effects models. Next, we show that in the usual fixed effects model, wherein the time-varying error component has a scalar covariance matrix, the MCS estimator reduces to the usual deviations from means estimator. Furthermore, the MCS over-identification test statistic is simply the degrees of freedom times the R^2 from a regression of residuals on all leads and lags of right hand side variables.

A substantial literature on instrumental variables (IV) estimation of fixed

[2]The general theory of MCS estimation for parametric models was developed by Ferguson (1958).

effects models has developed over the last few years. In the fourth section of the paper, we briefly discuss the relationship between the 3SLS equivalent of MCS and the efficient instrumental variables estimators of Amemiya and MaCurdy (1986) and Breusch, Mizon and Schmidt (1987).[3] For models without time-invariant regressors, the 3SLS equivalent of Chamberlain's MCS is shown to be the same as the instrumental variables estimator of Breusch, Mizon and Schmidt.

The last section of the paper presents an empirical example taken from the literature on life-cycle models of labor supply. An influential specification popularized by MaCurdy (1981) leads to fixed effects estimation of the relationship between log hours and log wages or log earnings. Re-estimating MaCurdy's equations, we find the restrictions that justify estimation in deviations from means are strongly rejected in two independent samples. Allowing for intertemporal correlation and heteroscedasticity does not appear to affect coefficient estimates. In contrast, the over-identification test statistic becomes substantially more forgiving when the labor supply equations are estimated without restricting the error covariance structure. Even the smallest of the test statistics, however, leads to rejection of the over-identifying restrictions implicit in the labor supply models.

---

[3]Instrumental variables estimators for panel data have also been proposed by Hausman and Taylor (1981), Nickell (1981), Hsaio (1986), Holtz-Eakin, Newey and Rosen (1985) and Holtz-Eakin (1988). In an early paper, Chamberlain (1977) discusses the instrumental variables interpretation of identification in a number of variance components models.

II. Estimation and testing of models with unobserved individual effects

The linear fixed effects model for panel data is

(1) $$y_{it} = x_{it}\beta + \alpha_i + u_{it}$$
$$\alpha_i = x_{i1}\lambda_1 + \ . \ . \ . + x_{iT}\lambda_T + \varepsilon_i$$

where $t = 1, \ldots, T$ indexes time and $i = 1, \ldots, N$ indexes individuals. x_{it} is a $1 \times k$ row vector and β and the λ's are $k \times 1$ column vectors of coefficients. The error components, ε_i and $u_i \equiv [u_{i1}, \ldots, u_{iT}]$, are assumed to be independent, not identically distributed and independent of each other.[4] Correlation between the unobserved individual effect, α_i, and the regressors, x_{it}, is described by the linear combination in the second line of (1). Because of this correlation, the Ordinary Least Squares (OLS) regression of y_{it} on x_{it} will not produce a consistent estimate of the slope coefficient. However, writing $\bar{y}_i$ and $\bar{x}_i$ for the time means of y_{it} and x_{it}, the OLS estimate of β in

(2) $$y_{it} - \bar{y}_i = (x_{it} - \bar{x}_i)\beta + (u_{it} - \bar{u}_i)$$

will be consistent. Least squares estimation of equation (2) is sometimes known as "Analysis of Covariance".

The specification that justifies Analysis of Covariance implies many testable restrictions. For example, data from any pair of time periods may be differenced and the differenced equation consistently estimated by OLS. Differenced equations which are more than two periods apart should yield

---

[4] We forgo detailed exposition of statistical regularity conditions. For a discussion of inference with i.n.i.d. samples see Chamberlain (1982) or White (1982).

estimates of β that are not statistically different from estimates based on adjacent differences.[5]

Chamberlain's (1982, 1984) MCS procedure for estimation of β and λ generates an "omnibus test" of all the over-identifying restrictions implied by the fixed effects model. Let $y_i = [y_{i1}, \ldots, y_{iT}]$ and $x_i = [x_{i1}, \ldots, x_{iT}]$ and denote the "reduced form" regression of y_i on x_i by $y_i = x_i\Pi + v_i$. Writing $\Pi(\beta,\lambda)$ for the relationship between reduced form and structural parameters, (1) implies that

$$\Pi(\beta,\lambda) = (I_T \otimes \beta) + \lambda\ell' \tag{3}$$

where $\ell' \equiv [1, \ldots, 1]$ is of dimension $1 \times Tk$ and λ is equal to $\mathrm{vec}[\lambda_1, \ldots, \lambda_T]$.[6] Let $\hat{\Pi}$ be a consistent estimate of Π. Then MCS estimates of β and λ are computed by minimizing

$$h(\beta,\lambda) = \mathrm{vec}[\hat{\Pi} - \Pi(\beta,\lambda)]'\hat{\Omega}^{-1}\mathrm{vec}[\hat{\Pi} - \Pi(\beta,\lambda)], \tag{4}$$

where $\hat{\Omega}$ is a consistent estimate of the variance of $\hat{\Pi}$. Furthermore, when (1) is the true model, the minimized value of $h(\beta,\lambda)$ has a chi-square distribution with $kT^2 - (Tk+k)$ degrees of freedom.[7]

The MCS procedure has a number of 3SLS equivalents, the simplest of which is based on the original model transformed to first differences:

---

[5] This fact has been used by Griliches and Hausman (1986) to test for measurement error in panel data.

[6] For example, in the case where $T = 2$ and $k = 1$, $\Pi = \begin{bmatrix} \beta+\lambda_1 & \lambda_1 \\ \lambda_2 & \beta+\lambda_2 \end{bmatrix}$.

[7] There are Tk λ's and k elements of β to infer from the kT^2 elements of Π.

$$(5) \qquad \begin{aligned} y_{i1}-y_{iT} &= (x_{i1}-x_{iT})\beta + (u_{i1}-u_{iT}) \\ &\ \ \vdots \\ y_{iT-1}-y_{iT} &= (x_{iT-1}-x_{iT})\beta + (u_{iT-1}-u_{iT}) \\ y_{iT} &= x_{iT}(\beta+\lambda_T) + x_{i1}\lambda_1 + \ldots + x_{iT-1}\lambda_{T-1} + \varepsilon_i + u_{iT}, \end{aligned}$$

or,

$$[y_{i1}, \ldots, y_{iT}] \begin{bmatrix} 1 & 0 & . & . & 0 \\ 0 & 1 & . & . & 0 \\ . & . & . & . & . \\ -1 & -1 & . & . & 1 \end{bmatrix} - [x_{i1}, \ldots, x_{iT}] \begin{bmatrix} \beta & 0 & . & . & \lambda_1 \\ 0 & \beta & . & . & \lambda_2 \\ . & . & . & . & . \\ -\beta & -\beta & . & . & \lambda_T+\beta \end{bmatrix}$$
$$= [u_{i1}-u_{iT}, \ldots, u_{iT-1}-u_{iT}, \varepsilon_i+u_{iT}].$$

3SLS estimation of this system using the untransformed x_i as instruments is the same as Chamberlain's MCS. To see this, let $e_i = [u_{i1}+\varepsilon_i, \ldots, u_{iT}+\varepsilon_i]$, so that (5) may be written compactly as

$$(5') \qquad y_i\Gamma - x_iB = e_i\Gamma = v_i.$$

An efficient 3SLS type estimator for (5') is derived from the moment restrictions

$$\text{vec}\{E[x_i'e_i\Gamma]\} = E[\Gamma'e_i' \otimes x_i'] = 0.$$

The variance of $\Gamma'e_i' \otimes x_i'$ is

$$\Psi = E[\Gamma'e_i'e_i\Gamma \otimes x_i'x_i].$$

and so, writing X and Y for the matrices whose rows are x_i and y_i, the 3SLS estimator is derived by minimizing

$$m(\beta,\lambda) = \text{vec}[X'(Y\Gamma + XB)]'\hat{\Psi}^{-1}\text{vec}[X'(Y\Gamma + XB)],$$

where

$$\hat{\Psi} = (1/N)\ \Sigma\ (\hat{v}_i'\hat{v}_i \otimes x_i'x_i)$$

is a consistent estimate of Ψ.

Now, Chamberlain (1982, p. 25-27) has shown that 3SLS minimands such as $m(\beta,\lambda)$ are asymptotically the same as

$$\text{vec}[\hat{\Pi} - B\Gamma^{-1}]'\hat{\Omega}^{-1}\text{vec}[\hat{\Pi} - B\Gamma^{-1}], \tag{6}$$

where $\hat{\Pi} = (X'X)^{-1}X'Y$. But in this case Γ has a simple lower-triangular form and $B\Gamma^{-1}$ is easily verified to be $(I_T \otimes \beta) + \lambda\ell'$. (6) is therefore the MCS minimand for fixed effects models.[8] Also, because MCS and 3SLS estimators minimize asymptotically equivalent quadratic forms with unique minima (if the parameters are identified), the resulting estimators are asymptotically equivalent.[9]

It is useful to confirm that $m(\beta,\lambda)$ and $h(\beta,\lambda)$ have chi-squared distributions with equal degrees of freedom. Note that there are T-1 over-identified equations in (5) and that each over-identified equation has k(T-1) over-identifying orthogonality restrictions (using the columns of X as instruments). In addition, there are k(T-2) equality restrictions on the β's.

[8] Note that $\Omega = (\Gamma^{-1\prime} \otimes \Phi_x^{-1})\Psi(\Gamma^{-1} \otimes \Phi_x^{-1})$, where $\Phi_x = E(x_i'x_i)$. If Φ_x is estimated by X'X/N and the estimate of Ψ used in $m(\beta,\lambda)$ is also used to form an estimate of Ω, then $m(\beta,\lambda)$ and $h(\beta,\lambda)$ will be numerically equal.

[9] The equivalence of MCS and 3SLS for fixed effects models is a special case of Proposition 3 in Newey and West (1987), which establishes the equivalence of MCS and Lagrange Multiplier minimands for restrictions on linear models. In the fixed effects problem, the restrictions in the Lagrange Multiplier version of the omnibus test may be written as $\Pi = B\Gamma^{-1}$, where B and Γ are linear in the parameters of interest. This fact allows conversion of the problem to one of over-identifying orthogonality conditions on a linear system.

$k(T-1)^2 + k(T-2)$ is indeed equal to $kT^2 - (kT+k)$, the degrees of freedom in $h(\beta,\lambda)$.

3SLS and MCS for a fixed effect model with lagged dependent variable

The relationship between MCS and 3SLS applies equally well to models with lagged dependent variables. To see this, consider a dynamic model with one regressor and one lagged dependent variable, assumed to be stationary:[10]

$$
\begin{aligned}
y_{it} &= y_{it-1}\gamma + x_{it}\beta_0 + \alpha_i + u_{it} \qquad t = 2, \ldots, T \\
y_{i1}(1-\gamma) &= x_{i1}\beta_0 + \alpha_i + u_{i1} \\
\alpha_i &= x_{i1}\lambda_1 + \ldots + x_{iT}\lambda_T + \varepsilon_i .
\end{aligned}
$$

Estimates of β_0, β_1 and γ may be recovered from the same reduced form regression coefficients as in model (1), although in this case MCS estimation is a forbidding task because the restrictions on the reduced form are nonlinear. An (asymptotically) equivalent procedure is to estimate the first differenced system by Three-Stage Least Squares, using the columns of X as instruments:

$$
\begin{aligned}
y_{iT}-y_{iT-1} &= (y_{iT-1}-y_{iT-2})\gamma + (x_{iT}-x_{iT-1})\beta_0 + (u_{iT}-u_{iT-1}) \\
&\vdots \\
y_{i3}-y_{i2} &= (y_{i2}-y_{i1})\gamma + (x_{i3}-x_{i2})\beta_0 + (u_{i3}-u_{i2}) \\
y_{i2}-y_{i1} &= (x_{i2}-x_{i1})\beta_0 + (u_{i2}-u_{i1}) \\
y_{i1}(1-\gamma) &= x_{i1}(\beta_0+\lambda_1) + x_{i2}\lambda_2 + \ldots + x_{iT}\lambda_{iT} + \varepsilon_i + u_{i1} .
\end{aligned}
$$

---

[10] As an alternative to the assumption of stationarity, the initial condition (y_{i0}) may be treated as a fixed effect with time-varying coefficients. For this version of an autoregressive model see Chamberlain (1984, p. 1267).

The MCS equivalent of 3SLS on the transformed system is found by defining Γ as the matrix of coefficients on dependent variables, lagged dependent variables and the error term in the transformed model, while B is defined as the matrix of coefficients on the independent variables in the transformed model. The matrix of reduced for regression coefficients is then fit to $B\Gamma^{-1}$ exactly as in the static model.

It is easy to verify that the MCS and 3SLS minimands have chi-squared distributions with the same degrees of freedom. In the MCS procedure for this problem, T^2 reduced form parameters are fit to T+2 structural parameters (T λ's, 1 γ, and 1 β_0). The first differenced system to be estimated by 3SLS contains T-2 equations with T-2 over-identifying restrictions, 1 equation with T-1 over-identifying restrictions, T-3 equality restrictions on γ and T-2 equality restrictions on β_0. The total number of restrictions is indeed equal to $T^2-(T+2)$.

To conclude this section we note that a related procedure for dynamic models with fixed effects has been described by Holtz-Eakin, Newey and Rosen (1985) and by Holtz-Eakin (1988). These authors are concerned with testing causality and estimating vector autoregressions (VAR) in short panels. In contrast, the models discussed here are likely to have some sort of structural interpretation. Yet, the practical difference between the 3SLS procedure for VAR's and the one described here is only that the VAR procedure requires estimation of differenced equations using lagged dependent as well as lagged independent variables in the instrument list. The 3SLS equivalent of Chamberlain's MCS uses all leads and lags of independent variables as instruments but excludes lagged dependent variables.

III. Analysis of Covariance revisited

The Three-Stage Least Squares procedures considered in the previous section are based on systems transformed to first differences. Deviations from means transformations may also be used to construct a 3SLS estimator equivalent to Chamberlain's MCS. To see this, consider model (1) in deviations from means form:

$$
\begin{aligned}
y_{i1}-\bar{y}_i &= (x_{i1}-\bar{x}_i)\beta + (u_{i1}-\bar{u}_i) \\
&\ \ \vdots \\
y_{iT-1}-\bar{y}_i &= (x_{iT-1}-\bar{x}_i)\beta + (u_{iT-1}-\bar{u}_i) \\
y_{iT} &= x_{iT}(\beta + \lambda_T) + x_{i1}\lambda_1 + \cdots + x_{iT-1}\lambda_{T-1} + \varepsilon_i + u_{iT}.
\end{aligned}
\tag{7}
$$

Systems (7) and (5) are related by a non-singular transformation and so they have the same reduced form (Schmidt 1976).

Let I_T denote the T-dimensional identity matrix. Mundlak (1978) shows that when α_i is correlated with x_i and $Eu_i'u_i = \sigma^2 I_T$ for all i, the Analysis of Covariance estimator is Best Linear Unbiased for β. Therefore, any 3SLS equivalent of the fully efficient MCS estimator should reduce to Analysis of Covariance in the scalar covariance case. We verify this conclusion, below, and show how the MCS test statistic may be computed by regressing the Analysis of Covariance residuals on x_i.

The 3SLS estimator when $Eu_i'u_i = \sigma^2 I_T$

The Tth equation in system (7) is just-identified and so 3SLS estimates of the first T-1 equations are unaffected by dropping the Tth equation.[11] Note that

[11] See, for example, Theil (1971, 513-514). In the appendix to Chamberlain (1984) this result is established for the heteroscedastic case.

the residuals in (7) may be written as $u_i[F\ e_T]$ where F is the $T \times T\text{-}1$ matrix that transforms the first T-1 elements of a $1 \times T$ vector into deviations from means and $e_T = [0\ .\ .\ .\ 0\ 1]'$. Also, note that post-multiplying by $F \otimes I_k$ transforms the $N \times Tk$ matrix X into an $N \times (T\text{-}1)k$ matrix of deviations from means. Stack the submatrices of $X[F \otimes I_k]$ for each period into an $N(T\text{-}1) \times k$ matrix, W, and let vec(YF) = q. When $Eu_i'u_i = \sigma^2 I_T$, the covariance matrix of u_iF is $(F'F)\sigma^2$. Using this notation, the 3SLS estimator for the first T-1 equations in (7) may be written

$$(8) \qquad \hat{\beta} = (W'[(F'F)^{-1} \otimes X(X'X)^{-1}X']W)^{-1}W'[(F'F)^{-1} \otimes X(X'X)^{-1}X']q.$$

Now,

$$F'F = [I_{T-1} - (1/T)\ell_{T-1}\ell_{T-1}'] \text{ and } (F'F)^{-1} = [I_{T-1} + \ell_{T-1}\ell_{T-1}'].$$

Substituting for $(F'F)^{-1}$ in the cross-product matrix and dropping the T-1 subscript, we have

$$(9) \quad W'[(F'F)^{-1} \otimes X(X'X)^{-1}X']W = W'(I \otimes X(X'X)^{-1}X')W + W'(\ell\ell' \otimes X(X'X)^{-1}X')W.$$

Let W_t and q_t be the deviations from means for all individuals in period t; W_t is N by k and q_t is N by 1. Let $P_x = X(X'X)^{-1}X'$. The two terms on the right hand side of (9) may now be written

$$\sum_{t=1}^{T-1} W_t'P_xW_t + \sum_{t=1}^{T-1}\sum_{s=1}^{T-1} W_t'P_xW_s = \sum_{t=1}^{T-1} W_t'P_xW_t + (\sum_{t=1}^{T-1} W_t)'P_x(\sum_{s=1}^{T-1} W_s).$$

Because the sum of T-1 deviations from means is equal to minus the Tth deviation from mean, we also have

$$\sum_{t=1}^{T-1} W_t = \sum_{s=1}^{T-1} W_s = - W_T.$$

Therefore,

$$\sum_{t=1}^{T-1} W_t'P_xW_t + \sum_{t=1}^{T-1} \sum_{s=1}^{T-1} W_t'P_xW_s = \sum_{t=1}^{T-1} W_t'P_xW_t + W_T'P_xW_T = \sum_{t=1}^{T} W_t'P_xW_t.$$

Finally, note that because each W_t is a linear combination of the columns of X,

$$\sum_{t=1}^{T} W_t'P_xW_t = \sum_{t=1}^{T} W_t'W_t,$$

which is the cross product matrix for Analysis of Covariance on the stacked observations. Likewise, the term $W'[(F'F)^{-1} \otimes X(X'X)^{-1}X']q$ in the formula for $\hat{\beta}$ may be shown to be $\sum_{t=1}^{T} W_t'q_t$. Therefore, the 3SLS estimator is indeed Analysis of Covariance.

The 3SLS minimand when $Eu_i'u_i = \sigma^2 I_T$

The over-identification test statistic associated with the MCS estimator is the MCS minimand, which has been shown here to be the 3SLS minimand for system (7). Study of the test statistic is simplified by the fact that the minimized 3SLS minimand for a system which includes just-identified equations is equal to the minimand for the over-identified equations only.[12]

Using x_i as instruments, the minimand for the T-1 over-identified equations

[12] This result is proved in Appendix B. It is analogous to the result (noted above) that the over-identified equations in a system with just-identified equations may be estimated separately.

in system (7) is

$$(10) \qquad (q - W\beta)'[(F'F)^{-1} \otimes P_x](q - W\beta)/\sigma^2 = \hat{v}'[(F'F)^{-1} \otimes P_x]\hat{v}/\sigma^2,$$

where $\hat{v}$ is the stacked analysis of covariance residual.[13] By working through the same algebra as was done for the estimator, it is straightforward to show when u_i has a scalar covariance matrix, (10) simplifies to

$$(1/\sigma^2) \sum_{t=1}^{T} \hat{v}_t' P_x \hat{v}_t.$$

But each term in this sum is simply the degrees of freedom times the R^2 from a regression of the analysis of covariance residuals for a particular period on all leads and lags of the right-hand side variables. The R^2 from a regression of residuals on instruments is also the basis of orthogonality test statistics in conventional homoscedastic linear simultaneous equations models (Newey 1985, Hausman 1984).

IV. Efficient instrumental variables estimators

Write $\nu_i = u_i + \alpha_i$ for the $1 \times T$ vector of compound error terms in the original, untransformed model. The moment restrictions that lead to an estimator based on using the deviations from means of x_i as instruments for an undeviated model may be written $E[\nu_i' x_i (F \otimes I_k)] = 0$. The moment restrictions that lead to an estimator based on using x_i as instruments in the deviated model may be written $E[(\nu_i F)' x_i] = 0$. In Appendix A, the following lemma is proved:

[13] Let $\hat{v}_t$ be the vector of analysis of covariance residuals in period t; each element of $\hat{v}_t$ is $\hat{v}_{it} = (y_{it} - \bar{y}_i) - (x_{it} - \bar{x}_i)\beta$.

Lemma 1. The optimal Generalized Method of Moments (GMM) estimators and test statistics based on the moment restrictions $E[(\nu_i F)'x_i] = 0$ and $E[\nu_i' x_i(F \otimes I_k)] = 0$ are identical.[14]

This lemma relates instrumental variables estimators of the type considered by Amemiya and MaCurdy (1986) and Breusch, Mizon and Schmidt (1987) to the 3SLS estimator described in the previous section.

Amemiya and MaCurdy's (AM) "efficient" instrument list contains all leads and lags of a set of variables assumed uncorrelated with the fixed effect plus contemporaneous deviations from means for correlated right hand side variables. Invoking somewhat stronger assumptions, Breusch, Mizon and Schmidt (BMS) propose an instrument list that includes all leads and lags of uncorrelated variables plus all leads and lags of correlated variables in deviations from means form.[15] But Lemma 1 says that using leads and lags of deviated regressors as instruments is equivalent to using leads and lags of undeviated instruments in the deviated model. Thus, the BMS procedure for models where all regressors are both time-varying and correlated with the fixed effect is equivalent to the 3SLS estimator introduced in the previous section. This, in turn, has been shown to be the same as the MCS estimator.

The AM and BMS instrumental variables procedures may also be used to estimate fixed effects models with time-invariant regressors if some of the regressors are assumed uncorrelated with the fixed effect. In this case, however, neither the AM or BMS procedure corresponds to an MCS estimator that

---

[14] This is not simply a restatement of the well-known instrumental variables interpretation of Analysis of Covariance (Hausman and Taylor 1981). The standard IV interpretation of Analysis of Covariance uses only contemporaneous deviations from means as instruments, while the moment restrictions considered here involve all leads and lags.

[15] For an empirical example using alternative instrumental variables estimators in panel data models see Cornwell and Rupert (1988).

imposes restrictions on the reduced form considered by Chamberlain. This is because the Chamberlain reduced form does not contain enough information to identify all the parameters in models with time-invariant regressors. In particular, the reduced from regression on all independent variables cannot be used to distinguish between coefficients on time-invariant regressors in the equation for α_i and coefficients on time-invariant regressors in the equations for y_{it}.

V. A labor supply example

In life-cycle models with intertemporally additive preferences in a certain environment, current period labor supply depends solely on the contemporaneous wage rate and the time-invariant marginal utility of wealth. The marginal utility of wealth is an unobserved variable that is correlated with wage rates. An attractive feature of life-cycle specifications, however, is that this source of omitted variables bias may be treated as a fixed effect in panel data.

One well-known variant of the life-cycle model is derived from a utility function discussed by Heckman and MaCurdy (1980) and MaCurdy (1981). This specification gives rise to a linear relationship between the log of annual hours worked and the log of hourly wages:

$$\ln h_{it} = \alpha + \rho_t + \theta \ln w_{it} + \lambda_i + u_{it}, \tag{11}$$

where $\ln h_{it}$ and $\ln w_{it}$ are the logs of hours and wages, ρ_t is a function of time-varying discount and interest rates, λ_i is a transformation of the marginal utility of wealth and θ is the intertemporal substitution elasticity. Inference usually focuses on θ, which must be positive for labor supply theory to provide a plausible explanation of cyclical fluctuations in employment.

Equation (11) appears to be a natural candidate for estimation using fixed effects techniques, however applied researchers generally find that fixed effects estimates of θ are negative.[16] An explanation commonly invoked to account for this unsatisfactory result is the presence of measurement error in the average hourly earnings data used to proxy for hourly wage rates. If annual hours are poorly measured, the ratio of earnings to hours will be negatively correlated with the error term in the hours equation. Differencing or deviations from means transformations are likely to aggravate this measurement error bias (Griliches and Hausman 1986).

One solution to the measurement error problem, originally proposed by MaCurdy (1981), is to add $\theta \ln h_{it}$ to both sides of equation (11). This gives

$$(12) \qquad \ln h_{it} = \alpha^* + \rho_t^* + [\theta/(1+\theta)]\ln y_{it} + \lambda_i^* + u_{it}^*,$$

where * denotes the original parameters divided by $(1+\theta)$ and $\ln y_{it}$ is the log of annual earnings. Measurement error in average hourly earnings need not bias coefficient estimates based on equation (12). On the other hand, a different sort of bias is introduced as a consequence of the correlation between $\ln y_{it}$ and u_{it}.[17] It is straightforward to show that the Analysis of Covariance estimate of $\theta/(1+\theta)$ in (12) actually converges to

$$(13) \qquad \eta = \theta/(1+\theta) + \sigma^2/[\tau^2(1+\theta)],$$

[16] See Ashenfelter (1984) for a pessimistic summary of applied work on labor supply.

[17] Note that $\ln y_{it} = \alpha + \rho_t + (1+\theta)\ln w_{it} + \lambda_i + u_{it}$. MaCurdy (1981) avoids bias from correlation between $\ln y_{it}$ and u_{it} by using instrumental variables other than earnings to estimate equation (12).

where σ^2 is the variance of u_{it} and τ^2 is the variance of $y_{it}-\rho_t^*-\lambda_i^*$. Solving for θ in (13) gives $\theta = (\eta - \gamma)/(1 - \eta)$, where $\gamma = \sigma^2/\tau^2$.

Ignoring the endogeneity of earnings in equation (12) biases estimates of θ upwards. In the current context, however, it is still of interest to see whether the over-identification test statistics actually reflect this bias. In Angrist (1988), evidence is presented which suggests that Analysis of Covariance estimates of $\theta/(1+\theta)$ do not differ significantly from a set of instrumental variables estimates known to be consistent. One reason for this may be that σ^2/τ^2 is small relative to θ.[18] The empirical example therefore includes estimates of (12) as well as (11).

To compute fixed effects estimates, two independent samples were drawn from the 1969-1979 waves of the Panel Survey of Income Dynamics. The samples include male heads of household with positive earnings and hours in every year from 1969 to 1979. Individuals with extreme values for wages, hours or earnings were removed, as were members of the Survey of Economic Opportunity low income subsample.[19] The removal of the SEO sample and outliers reduced the original sample of continuously employed male heads from a size of 1437 to 803. Although we are agnostic regarding the use of such sample selection criteria, the sample is restricted so as to be consistent with earlier work (i.e., MaCurdy 1981, Abowd and Card 1986 and Altonji 1986).

The 803 observations in the selected sample were randomly divided into two half-samples, henceforth referred to as sample 1 (with 387 observations) and

---

[18] In Angrist (1988), there is assumed to be no time-varying error other than a parametric trend.

[19] Excluded are men with average hourly earnings in 1967 dollars above 100 dollars in any year, men with annual hours worked above 4680 in any year, men with change in hours greater than 3000 or change in log hours greater than 1 in any year and men with change in log earnings in 1967 dollars greater than 1 in any year. Detailed PSID variable definitions and sample means are reported in Appendix C.

sample 2 (with 416 observations). Estimates of θ in equation (11) and $\theta/(1+\theta)$ in equation (12) are presented in Table 1 for both samples. In equation (11), w_{it} is taken to be average hourly earnings, computed by dividing annual earnings by annual hours worked. In equation (12), y_{it} is annual wage and salary earnings and in both equations the dependent variable is the log of annual hours worked. The equations also include a full set of time dummies to represent ρ_t.

Three estimation techniques are used in the empirical work. First, conventional Analysis of Covariance estimates were constructed, along with the corresponding over-identification test statistic discussed in Section III. The test statistic was computed by regressing the Analysis of Covariance residuals for each period on all the instruments.[20]

The second set of estimates was computed by Three-Stage least Squares allowing for unrestricted intertemporal correlation in the residuals of the first T-1 equations in system (7). In this case, the test statistic is simply the 3SLS minimand.[21] The last set of estimates allows for intertemporal correlation and heteroscedasticity using the generalized Three-Stage Least squares procedure discussed in Chamberlain (1982, 1984). In this case, optimal heteroscedasticity corrected estimates and covariance matrices were computed by replacing $\hat{v}_i$ with

---

[20] We used SAS PROC SYSLIN with the SDIAG option to estimate a diagonal 3SLS system for the residual regressions. Multiplying the SAS "System Weighted R^2" by the degrees of freedom in the stacked system gives the test statistic. The degrees of freedom should be adjusted to reflect the deviations from means transformation.

[21] SAS PROC SYSNLIN was used to perform the 3SLS estimation. SYSNLIN reports the value of the 3SLS minimand as "N times OBJECTIVE". Kiefer (1980) shows that the sample size is an asymptotically correct denominator for the estimate of residual covariance in panel data models estimated in deviations from means. We therefore set SAS option VARDEF = N. Note that SYSNLIN writes the residuals to a data set where they may be saved for use in subsequent heteroscedasticity-consistent estimation.

Two-Stage Least Squares residuals in $\hat{\Phi} = (1/N) \Sigma (\hat{v}_i'\hat{v}_i \otimes x_i'x_i)$.[22]

The top half of Table 1 shows estimates of equation (11) and the bottom half shows estimates of equation (12). Analysis of Covariance estimation of equation (11) gives an intertemporal substitution elasticity of -.277 . It is apparent that regardless of the technique used, estimates of θ from regressions on wages are always negative. Estimates of $\theta/(1+\theta)$ from regressions on earnings, shown in the lower half of the table, are all positive. For example, the Analysis of Covariance estimate of $\theta/(1+\theta)$ implies an intertemporal substitution elasticity around .30 . This is quite close to MaCurdy's (1981, Table 1) instrumental variables estimates.

The one percent critical value for a chi-square statistic with 109 degrees of freedom is 147; all chi-square statistics in the table therefore lead to rejection at the one percent level. The fact that over-identification test statistics for model (12) also reject suggests that estimates of $\theta/(1+\theta)$ reflect bias induced by correlation between earnings and the error term in the hours equation. The restrictions tested in 3SLS estimation of equation (12), however, appear to fit the data somewhat better than the restrictions tested in equation (11). Thus, it may be that measurement error in hourly wages is of greater practical importance than simultaneity bias.

In both equations, test statistics that allow for correlation and heteroscedasticity are substantially smaller than the test statistics from analysis of covariance estimation. Results from other samples, not reported here, also show this pattern. Test statistics computed from Sample 1 and Sample

---

[22] Estimates were constructed using SAS PROC MATRIX to form $\hat{\Phi}$ from residuals. We then computed X'W and X'q. These are the stacks of second moments of the instruments (X) and the dependent and independent variables in deviations from means. Finally, the moments were premultiplied by the square root of the inverse of $\hat{\Phi}$, and heteroscedasticity consistent 3SLS estimates computed by Ordinary Least Squares (OLS) on the transformed moments. The test statistic is the sum of squares from this regression.

2 combined (803 observations) and from the larger, unselected sample (1437 observations) are always ordered in the same manner as the test statistics in Table 1. Of course, test statistics from the larger samples are larger than those reported in the table.

Table 2 reports the estimated residual covariance matrices for the residuals used to form $\hat{\Psi}$ in Sample 1. Correlation matrices for sample 2, not shown here, are similar. The matrices for both earnings and wages show correlation that is positive close to the diagonal, but decreasing and negative at longer lags. This is the pattern to be expected if the residuals of untransformed variables exhibit positive first order serial correlation. To see this, suppose that earnings have first order serial correlation coefficient ϕ. The t,k element of the correlation matrix for earnings deviations from means is

$$\phi^{|t-k|} - (1/T) \sum_{j=1}^{T} \phi^{|j-k|} .$$

For fixed k, it is apparent that this is a decreasing function of t that eventually becomes negative.

The fact that the correlation matrix of deviations from means has a large number of negative elements is important because it suggests an explanation for the difference between Analysis of Covariance test statistics and 3SLS test statistics. To see this, consider a t-statistic for the difference between two negatively correlated means, $\bar{x}_1$ and $\bar{x}_2$. The denominator of the test statistic is the standard error of the difference. Ignoring negative correlation between the means, the standard error will be too small (and the statistic too large) because $V(\bar{x}_1-\bar{x}_2) = V(\bar{x}_1) + V(\bar{x}_2) -2Cov(\bar{x}_1,\bar{x}_2)$.

VI. Conclusions

The relationship between Minimum Chi-Square and Three-Stage Least Squares procedures for fixed effects models has been considered here in some detail. For the practitioner, the most useful result is that in the special case of a static fixed effects model with scalar covariance matrix and no time-invariant regressors, MCS is equivalent to Analysis of Covariance. Furthermore, the MCS over-identification test statistic in this case is simply the degrees of freedom times the sum of R^2's from regressions of Analysis of Covariance residuals on regressors. This is the same method used to calculate orthogonality test statistics in homoscedastic linear simultaneous equations models. We have also considered the relationship between the 3SLS equivalent of MCS and efficient instrumental variables estimators previously found in the literature. The 3SLS equivalent of MCS has been shown to be the same as an estimator proposed by Breusch, Mizon and Schmidt (1987).

In the last section of the paper, an empirical example taken from the literature on life-cycle labor supply is presented and the 3SLS over-identification test statistic computed under alternative assumptions regarding the residual covariance matrix. Coefficient estimates and standard errors were found to be insensitive to these assumptions. The over-identification test statistic, however, is much smaller when allowance is made for intertemporal correlation and heteroscedasticity. Nevertheless, the assumptions justifying fixed effects estimation are found to be inappropriate for both of the labor supply equations estimated here.

Table 1

Indep. Var.	Technique	Sample 1 θ	Sample 1 $\chi^2(109)$	Sample 2 θ	Sample 2 $\chi^2(109)$
Equation (11)					
$Ln(w_{it})$	ANACOVA	-.277 (.010)	341.0	-.272 (.010)	313.1
	3SLS	-.333 (.010)	318.4	-.336 (.010)	238.0
	Heteroscedastic 3SLS	-.295 (.009)	189.7	-.288 (.010)	162.2
		$\theta/(1+\theta)$	$\chi^2(109)$	$\theta/(1+\theta)$	$\chi^2(109)$
Equation (12)					
$Ln(y_{it})$	ANACOVA	.230 (.011)	309.8	.232 (.011)	230.3
	3SLS	.216 (.012)	283.4	.238 (.011)	231.6
	Heteroscedastic 3SLS	.231 (.011)	161.6	.220 (.014)	148.6

Standard errors in parentheses.

Estimates are for regressions of log hours on the independent variables indicated.

All equations include a full set of year effects.

Table 2

Correlations of Deviations from Means (logs), 1969-79 in 1967 Dollars

Sample 1 Earnings

1.0										
.3675	1.0									
.1456	.2829	1.0								
.0184	.0428	.1044	1.0							
-.1808	-.2239	-.0997	.1503	1.0						
-.2219	-.1435	-.2172	-.0946	.0470	1.0					
-.1723	-.2522	-.3613	-.2843	-.0635	.0421	1.0				
-.3491	-.2877	-.2693	-.2755	-.1666	-.0407	.2564	1.0			
-.3737	-.4009	-.3665	-.2265	-.1021	.0497	.0863	.2307	1.0		
-.3266	-.3968	-.2100	-.2181	-.0856	-.1232	-.0384	.0467	.1413	1.0	
-.2067	-.2603	-.2240	-.2239	-.1457	-.1260	-.1332	-.0311	.1371	.3485	1.0

Sample 1 Wages

1.0										
.5030	1.0									
.3243	.3680	1.0								
.0856	.0796	.2492	1.0							
-.2087	-.1446	-.1003	.1472	1.0						
-.3203	-.1841	-.2771	-.1746	.1158	1.0					
-.3243	-.2430	-.4564	-.3854	-.0491	.1946	1.0				
-.3728	-.3176	-.3980	-.3303	-.2076	.0495	.3579	1.0			
-.4261	-.4381	-.4332	-.2904	-.0839	.0374	.1409	.2441	1.0		
-.3889	-.4966	-.3329	-.2444	-.1151	-.0784	.0114	.1326	.3056	1.0	
-.3508	-.4721	-.2674	-.1537	-.1192	-.0472	-.0879	.0499	.1935	.3870	1.0

Appendix A

Proof of Lemma 1.

The two GMM estimators referred in lemma 1 choose the parameter vector (δ) to minimize an optimally weighted quadratic form in the sample analog of $g_0(\delta) = \mathrm{vec}\{E[\nu_i{}'x_i(F\otimes I_k)]\}$ or $g_1(\delta) = \mathrm{vec}\{E[F'\nu_i{}'x_i]\}$. Note that

$$\mathrm{vec}[\nu_i{}'x_i(F\otimes I_k)] = (F'\otimes I_k\otimes I_T)\mathrm{vec}(\nu_i{}'x_i) = (F'\otimes I)\mathrm{vec}(\nu_i{}'x_i)$$

and that

$$\mathrm{vec}[F'\nu_i{}'x_i] = (I\otimes F')\mathrm{vec}(\nu_i{}'X),$$

where I is of dimension Tk.

Denote the asymptotic covariance matrix of $\mathrm{vec}(\nu_i{}'x_i)$ by Ξ and stack ν_i and x_i into $N \times T$ and $N \times kT$ matrices, ν and X. Optimal GMM minimands for $g_0(\delta)$ and $g_1(\delta)$ are given by

$$m_0 = [\mathrm{vec}(\nu'X)]'(F\otimes I)[(F'\otimes I)\Xi(F\otimes I)]^{-1}(F'\otimes I)[\mathrm{vec}(\nu'X)]$$

and

$$m_1 = [\mathrm{vec}(\nu'X)]'(I\otimes F)[(I\otimes F')\Xi(I\otimes F)]^{-1}(I\otimes F')[\mathrm{vec}(\nu'X)].$$

Both m_0 and m_1 have chi-square distributions with kT^2-(kT+k) degrees of freedom, as may be verified by counting orthogonality and equality restrictions in the original moment vectors.

Denote the kronecker product (tensor) commutator matrix by ∇. ∇ is a matrix with the properties that $(F'\otimes I) = \nabla(I\otimes F')$, $F\otimes I = (I\otimes F)\nabla'$ and $\nabla' = \nabla^{-1}$ (Pollock 1979). Equality of m_1 and m_0 follows by manipulation of these relationships:

$$
\begin{aligned}
m_0 &= [\mathrm{vec}(\nu'X)]'(F\otimes I)[(F'\otimes I)\Xi(F\otimes I)]^{-1}(F'\otimes I)[\mathrm{vec}(\nu'X)] \\
&= [\mathrm{vec}(\nu'X)]'(I\otimes F)\nabla'[\nabla(I\otimes F')\Xi(I\otimes F)\nabla']^{-1}\nabla(I\otimes F')[\mathrm{vec}(\nu'X)] \\
&= [\mathrm{vec}(\nu'X)]'(I\otimes F)\nabla'\nabla[(I\otimes F')\Xi(I\otimes F)']^{-1}\nabla'\nabla(I\otimes F)[\mathrm{vec}(\nu'X)] \\
&= m_1.
\end{aligned}
$$

The optimal GMM estimators that minimize m_0 and m_1 are identical because, provided the model is identified, they minimize identical quadratic forms with unique minima.

Appendix B

The 3SLS Minimand for a System with Just-identified Equations

The 3SLS estimator for over-identified equations in a system which contains just-identified equations is well-known to be 3SLS on the over-identified equations alone. An analogous result for the 3SLS minimand does not seem to be in the literature, although Abowd and Card (1986) present a related result for Optimal Minimum Distance (OMD) over-identification test statistics. They show that the OMD test statistic for a set of restrictions that leaves some elements of the parameter vector unrestricted is the OMD quadratic form for the restricted elements only.

Let "subscript a" denote the over-identified equations, let "subscript b" denote the just-identified equations and let δ_a and δ_b be the corresponding vectors of parameters. Then, we may write

$$\begin{aligned} y_a &= W_a\delta_a + v_a \\ y_b &= W_b\delta_b + v_b \end{aligned} \tag{b1}$$

for the stacked system of over-identified and just-identified equations. Let I_a and I_b be identity matrices with dimensions equal to the number of equations in systems a and b and let $(I_a \otimes X')W_a = H_a$, $(I_a \otimes X')y_a = h_a$, $(I_b \otimes X')y_b = h_b$ and $(I_b \otimes X')W_b = H_b$, where X denotes the matrix of exogenous variables. Note that, because the equations in b are just-identified, H_b is non-singular. Denote the covariance matrix of $v = [v_a{}' \ v_b{}']'$ by Σ and partition Σ and it's inverse by writing

$$\Sigma^{-1} = \begin{pmatrix} P_1 & R_1 \\ R_1{}' & Q_1 \end{pmatrix}^{-1} = \begin{pmatrix} P_2 & R_2 \\ R_2{}' & Q_2 \end{pmatrix},$$

where P_1 is the covariance matrix of v_a and Q_1 is the covariance matrix of v_b. Finally, let P_x denote the projection matrix, $X(X'X)^{-1}X'$.

We consider here only the case where u is homoscedastic. Algebraic development of the heteroscedastic case is similar. The conventional 3SLS estimator is motivated by the moment restrictions $E[(I \otimes X')v] = 0$. When v is homoscedastic, the variance of $(I \otimes X')v$ is $\Sigma \otimes (X'X)$ and so the optimal GMM estimator based on these restrictions minimizes

$$v'(I \otimes X)[\Sigma^{-1} \otimes (X'X)^{-1}](I \otimes X')v = v'[\Sigma^{-1} \otimes P_x]v, \tag{b2}$$

which is also the 3SLS minimand.

Theil (1971, p. 513-514) shows that the 3SLS estimator of parameters in the just-identified equations, δ_b, is

$$\begin{aligned} \hat{\delta}b &= \hat{d}_b - H_b^{-1}[R_1'P_1^{-1} \otimes I](h_a - W_a\hat{\delta}_a) \\ &= \hat{d}_b - H_b^{-1}[R_1'P_1^{-1} \otimes I][I \otimes X']\hat{v}_a, \end{aligned} \tag{b3}$$

where $\hat{d}_b$ is the Two-Stage Least Squares (TSLS) estimator of δ_b and $\hat{v}_a$ is the vector of 3SLS residuals from system a. Therefore, the 3SLS residuals from system b are

$$\hat{v}_b = \hat{z}_b + W_bH_b^{-1}[R_1'P_1^{-1} \otimes X']\hat{v}_a \tag{b4}$$

where $\hat{z}_b$ is the vector of TSLS residuals. Using the fact that TSLS residuals from the just-identified system are orthogonal to X and that $(I \otimes X')W_b = H_b$, we have

$$(I \otimes X')\hat{v}_b = [R_1'P_1^{-1} \otimes X']\hat{v}_a = [I \otimes X'][R_1'P_1^{-1} \otimes I]\hat{v}_a. \qquad \text{(b5)}$$

Now, applying the notation introduced above, the 3SLS minimand, (b2), may be written as

$$v_a'[P_2 \otimes P_x]v_a + v_b'[Q_2 \otimes P_x]v_b + 2v_a'[R_2 \otimes P_x]v_b. \qquad \text{(b6)}$$

Replacing v_b with $\hat{v}_b$, the minimized second term in (b6) is

$$\hat{v}_b'[Q_2 \otimes P_x]\hat{v}_b = \hat{v}_b'[I \otimes X][Q_2 \otimes (X'X)^{-1}][I \otimes X']\hat{v}_b,$$

so that substituting for $[I \otimes X']\hat{v}_b$ using (b5) gives

$$\begin{aligned}\hat{v}_b'[Q_2 \otimes P_x]\hat{v}_b &= \hat{v}_a'[P_1^{-1}R_1 \otimes I][I \otimes X][Q_2 \otimes (X'X)^{-1}][I \otimes X'][R_1'P_1^{-1} \otimes I]\hat{v}_a \\ &= \hat{v}_a'[P_1^{-1}R_1 \otimes P_x][Q_2 \otimes P_x][R_1'P_1^{-1} \otimes P_x]\hat{v}_a.\end{aligned}$$

Likewise, the minimized third term in (b6) may be shown to be

$$2\hat{v}_a'[R_2 \otimes P_x]\hat{v}_b = 2\hat{v}_a[R_2 \otimes P_x][R_1'P_1^{-1} \otimes P_x]\hat{v}_a.$$

The minimized 3SLS minimand for the entire system is therefore

$$\begin{aligned}&\hat{v}_a'[P_2 \otimes P_x]\hat{v}_a + \hat{v}_a[P_1^{-1}R_1 \otimes P_x]'[Q_2 \otimes P_x][R_1'P_1^{-1} \otimes P_x]\hat{v}_a \\ &\qquad + 2\hat{v}_a'[R_2 \otimes P_x][R_1'P_1^{-1} \otimes P_x]\hat{v}_a \\ &= \hat{v}_a'[(P_2+2R_2R_1'P_1^{-1}+P_1^{-1}R_1Q_2R_1'P_1^{-1}) \otimes P_x]\hat{v}_a. \qquad \text{(7)}\end{aligned}$$

From partitioned inversion formulas (e.g., Theil 1971, Chapter 1), it immediately follows that

$$R_2R_1' = I - P_2P_1 \text{ and } Q_2R_1' = -R_2'P_1.$$

Using these formulas to substitute for R_2R_1' and Q_2R_1' in (7), the matrix in the quadratic form may be shown to be $P_1^{-1} \otimes P_x$. Thus, the minimized 3SLS minimand for the entire system is $\hat{v}_a'[P_1^{-1} \otimes P_x]\hat{v}_a$, which is also the minimand for the over-identified equations only.

Appendix C

Data Sources, Sample Selection and Sample Means

1. Data Sources

Earnings, average hourly earnings (wages) and hours are taken from the Michigan Panel Survey of Income Dynamics, Wave 13 Individual Family Tape.

Annual Earnings variables are V1196, V1897, V2498, V3051, V3463, V3863, V5031, V5627, V6174, V6767, V7413.

Annual Hours of work variables are V1138, V1839, V2439, V3027, V3423, V3823, V4332, V5232, V5731, V6336, V6934.

Average hourly earnings equals Annual Earnings divided by Annual Hours.

Nominal data were deflated by the CPI (all items), taken from the Economic Report of the President 1987, p. 307.

2. Sample Selection

Male heads of household aged 21-64 over the period 1969-79. Only household heads in households that experienced no change in head were included.

Sample restricted to heads with positive hours and earnings in each year from 1969-79.

Sample restricted to heads with average hourly earnings (in 1967$) less than 100$ in each year and with annual hours less than 4680 in each year.

Sample excludes men with change in hours greater than 3000 or change in log hours greater than 1 in any year.

Sample excludes men with change in log earnings (in 1967$) greater than 1 in any year.

There are 387 men in Sample 1 and 416 men in Sample 2.

3. Sample means

<u>Sample 1</u>

Year	Earnings	Hours	Wages
69	9145.5	2340.22	4.01780
70	9238.6	2290.33	4.14080
71	9606.2	2333.13	4.21585
72	10056.5	2346.59	4.40760
73	10400.1	2334.73	4.57161
74	9930.9	2294.33	4.44448
75	9695.6	2253.66	4.42116
76	10128.1	2272.27	4.54970
77	10346.8	2256.32	4.68656
78	10327.7	2272.50	4.68094
79	9918.0	2232.48	4.56083

<u>Sample 2</u>

69	9442.0	2319.22	4.11528
70	9532.2	2305.98	4.24354
71	9902.6	2324.60	4.39923
72	10683.0	2322.12	4.74795
73	10916.9	2290.23	4.88179
74	10400.9	2255.02	4.69688
75	10435.3	2236.97	4.70685
76	10741.3	2263.54	4.83063
77	11143.9	2301.40	4.93921
78	11145.5	2262.82	5.01614
79	10970.1	2226.50	4.97757

References

Abowd, J.A. and D. Card (1986). "On the Covariance of Structure of Earnings and Hours Changes," National Bureau of Economic Research Working Paper N. 1832, February 1986.

Altonji, J. G. (1986). "Intertemporal Substitution in Labor Supply: Evidence from Micro Data," Journal of Political Economy 94 (Vol. 3, Part 2): S176-S215.

Amemiya, T. and T.E. MaCurdy (1986). "Instrumental Variable Estimation of an Error-Components Model," Econometrica 54: 869-880.

Angrist, J.D. (1988). "Grouped Data Estimation and Testing in Simple Labor Supply Models," Princeton University, ndustrial Relations Section Working Paper No. 234, July 1988.

Ashenfelter, O.C. (1984). "Macroeconomic and Microeconomic Analyses of Labor Supply," Carnegie-Rochester Conference on Public Policy 21: 117-156.

Breusch, T.S., G.E. Mizon and P. Schmidt (1987). "Efficient Estimation Using Panel Data," Econometrics and Economic Theory Paper No. 8608, Michigan State University, East Lansing, Michigan, 1987.

Chamberlain, G. (1982). "Multivariate Regression Models for Panel Data," Journal of Econometrics 18: 5-46.

Chamberlain, G. (1984). "Panel Data," Chapter 22 in Handbook of Econometrics Volume II, Edited by Z. Griliches and M. Intriligator, Elsevier Science Publishers BV, 1984.

Chamberlain, G. (1977). "An Instrumental Variables Interpretation of Identification in Variance-Components and MIMIC Models," in Kinometrics: The Determinants of Socioeconomic Success Within and Between Families, edited by Paul Taubman, Amsterdam: North Holland, 1977.

Cornwell, C. and P. Rupert (1988). "Efficient Estimation with Panel Data: An Empirical Comparison of Instrumental Variables Estimators," Journal of Applied Econometrics 3:149-155.

Ferguson, T.S. (1958). "A Method of Generating Best Asymptotically Normal Estimates with An Application to the Estimation of Bacterial Densities," Annals of Mathematical Statistics 29: 1046-1062.

Griliches, Z. and Hausman, J.A. (1986). "Errors in Variables in Panel Data," Journal of Econometrics 31: 93-118.

Hausman, J.A. (1984). "Specification and Estimation of Simultaneous Equation Models," Handbook of Econometrics Chapter 7, Z. Griliches and M.D. Intriligator, eds., Amsterdam: North-Holland.

Hausman, J.A. and W. Taylor (1981). "Panel Data and Unobservable Individual Effects," Econometrica 49: 1377-1398.

Heckman, J.J. and T.E. MaCurdy (1980). "A Life-Cycle Model of Female Labor Supply," The Review of Economic Studies 47: 47-74.

Holtz-Eakin, D. (1988). "Testing for Individual Effects in Autoregressive Models," Journal of Econometrics 39: 297-307.

Holtz-Eakin, D., W. Newey and H. Rosen (1985). "Implementing Causality Tests with Panel Data, with an Example from Local Public Finance," NBER Technical Working Paper No. 48, June 1985.

Hsaio (1986). Panel Data, Cambridge: Cambridge University Press, 1986.

Kiefer, N.M (1980). "Estimation of Fixed Effect Models for Time Series of Cross-Sections with Arbitrary Intertemporal Covariance," Journal of Econometrics 14: 195-202.

MaCurdy (1981). "An Empirical Model of Labor Supply in a Life-Cycle Setting," Journal of Political Economy 89: 1059-1086.

Mundlak, Y. (1978). "On the Pooling of Times Series and Cross Section Data," Econometrica 46: 69-85.

Newey, W.K. (1985). "Generalized Method of Moments Estimation and Testing," Journal of Econometrics 29: 229-256.

Newey, W.K. and K. West (1987). "Hypothesis Testing with Efficient Method of Moments Estimation," International Economic Review 3: 777-787.

Nickell, S. (1981). "Biases in Dynamic Models with Fixed Effects," Econometrica 49: 1417-1426.

Pollock, D.S.G. (1979). The Algebra of Econometrics, New York: John Wiley and Sons, 1979.

Rothenberg, T.J (1973). Efficient Estimation with **A Priori** Information, New Haven: Yale University Press, 1973.

Schmidt, P. (1976). Econometrics, New York: Marcel Dekker, Inc., 1976.

Theil, H. (1971). Principles of Econometrics, New York: John Wiley and Sons, Inc., 1971.

White, H. (1982). "Instrumental Variables Regression with Independent Observations," Econometrica 50: 483-501.