Title

An information based sample-selection estimation model of agricultural workers' choice between piece-rate and hourly work

Authors

Golan, Amos
Morttie, Enrico
Perloff, Jeffrey M

Publication Date

1998-12-01

DEPARTMENT OF AGRICULTURAL AND RESOURCE ECONOMICS AND POLICY
DIVISION OF AGRICULTURE AND NATURAL RESOURCES
UNIVERSITY OF CALIFORNIA AT BERKELEY

WORKING PAPER NO. 861

AN INFORMATION BASED SAMPLE-SELECTION ESTIMATION MODEL OF AGRICULTURAL WORKERS' CHOICE BETWEEN PIECE-RATE AND HOURLY WORK

by

Amos Golan, Enrico Moretti and Jeffrey M. Perloff

California Agricultural Experiment Station
Giannini Foundation of Agricultural Economics
December 1998

KW
KEY WORDS
maximum entropy, information, labor, piece rate, wage differentials

AB
ABSTRACT

This paper presents a new generalized maximum entropy (GME) approach to estimation of sample-selection models with small data sets, such as are found in many empirical agricultural economic analyses. For small samples, the GME approach produces more stable estimates and has smaller mean square error measures than other well-known estimators such as ordinary least squares, Heckman's two-step method, full-information maximum likelihood, and Ahn and Powell's method. The technique is used to analyze whether hired agricultural workers will work in piece-rate or time-rate jobs and to compare female-male wage differentials for both types of jobs.

An Informational Based Sample-Selection Estimation Model of Agricultural Workers' Choice Between Piece-Rate and Hourly Work

Amos Golan, Enrico Moretti, and Jeffrey M. Perloff*

This paper presents a new generalized maximum entropy (GME) approach to estimation of sample-selection models with small data sets, such as are found in many empirical agricultural economic analyses. For small samples, the GME approach produces more stable estimates and has smaller mean square error measures than other well-known estimators such as ordinary least squares, Heckman's two-step method, full-information maximum likelihood, and Ahn and Powell's method. The technique is used to analyze whether hired agricultural worker will work in a piece-rate or time-rate jobs and to compare female-male wage differentials for both types of jobs.

Key Words: maximum entropy, information, labor, piece rate, wage differentials

The problem of sample selection arises frequently in agricultural economics, such as in studies of individuals' wages or labor supply. With large data sets of "well-behaved" data, the traditional (parametric and semi-parametric) approaches perform well. These models include the two-step and maximum likelihood sample-selection approaches (e.g., Heckman) as well as

---

* The authors are a visiting professor of economics at American University; a graduate student in the Economics Department at the University of California, Berkeley; and a professor in Department of Agricultural & Resource Economics at the University of California, Berkeley and a member of the Giannini Foundation. We thank Richard Mines of the Office of the Assistant Secretary for Policy, U.S. Department of Labor for permission to use the NAWS data. Perloff gratefully acknowledges support from the Giannini Foundation and the Institute of Industrial Relations at Berkeley.

the semi-parametric class of estimators (e.g., Ahn and Powell). However, when sample sizes are small, data are non-experimental and somewhat contaminated, perhaps due to multicollinearity, and the researcher is not sure what data-generation process underlies the data, the traditional models may have difficulties and may produce unstable results. Unfortunately, many if not most data sets have these limitations and therefore traditional methods may not be fully satisfactory.

Our objective is to summarize a new, semi-parametric approach for estimating small data set sample-selection problems and use it to examine an important problem in agricultural labor economics. The approach we take grew out of information theory and is based on the classical maximum entropy approach (Jaynes, Levine) and the generalized maximum entropy (GME) work of Golan, Judge and Miller. Our main goal is to estimate the set of unknown parameters, incorporating all the possible information (data points as well as economic theoretical requirements) in the estimation procedure without making a priori assumptions regarding the underlying distribution.

We use our method to study how agricultural employees choose to work in piece-rate or time-rate sectors, how the wage equations differ across these sectors, and how the female-male wage differential varies across regions. Because we are interested in regions, the sample sizes are relatively small and traditional approaches may not perform well. We compare our estimates to those of four other methods.

The first section specifies the sample-selection model. Section 2 develops the background and discusses the GME estimation model. Section 3 lists the relevant inference and diagnostic measures. Section 4 discusses the data and the main empirical results.

Section 5 contains conclusions.

The Model

Suppose each Person i decides whether he or she prefers to work in the time-rate sector or in the piece-rate sector of the agricultural labor market. Let the logarithm of the hourly earnings (henceforth just called the "hourly earnings" for short) that an individual would receive in the time-rate and piece-rate sectors be w_t^* and w_p^* respectively. An individual's hourly earnings in Sector j is $w_j^* = X\underline{\beta}_j + \underline{\varepsilon}_j$, where $j = t$ or p, X is a $N \times K$ matrix demographic and other socioeconomic covariates, where the first column consists of ones, and $\underline{\beta}_t$ and $\underline{\beta}_p$ are K-dimensional vectors of unknown parameters.

We only observe earnings for the sector in which the individual works. If the i^{th} individual works in the time-rate sector, the observed wage is $w_{ti} = w_{ti}^*$; otherwise, we observe $w_{pi} = w_{pi}^*$. An individual works in the time-rate sector if

$$(1) \qquad \begin{cases} w_{ti} = w_{ti}^* = X\underline{\beta}_t + \varepsilon_{ti} \\ w_{ti}^* - w_{pi}^* > C_i\underline{\beta}_c, \end{cases}$$

where ε_{ti} is $N \times 1$, $w_{ti}^* - w_{pi}^*$ is the relative benefit from working in the time-rate sector rather than the piece-rate sector, $C\underline{\beta}_c$ is the relative cost of being in the time-rate sector, C is a $N \times L$ matrix of covariates, and $\underline{\beta}_c$ is a L-dimensional vector of unknown parameters. Otherwise, that individual works in the piece-rate sector:

$$(2) \qquad \begin{cases} w_{pi} = w_{pi}^* = X\underline{\beta}_p + \varepsilon_{pi} \\ w_{ti}^* - w_{pi}^* \le C_i\underline{\beta}_c. \end{cases}$$

A common approach to estimating this model is to use Heckman's two-step procedure, where the maintained assumption is that of normality. The first step is to estimate a probit of the sector choice conditional on the union of the variables in X and C. The next step is to estimate, for the appropriate subsamples, the wage equations including an extra variable, the inverse Mills ratio from the probit, to avoid a sample-selection bias. Because this method requires a normality assumption and does not perform well in small samples (see Golan, Moretti, and Perloff), we propose an alternative approach.

GME Estimation Approach

Given the estimation problem in Equations (1) and (2), our objective is to estimate $\underline{\beta}_t$, $\underline{\beta}_p$, and $\underline{\beta}_c$ with minimum distributional assumptions, while incorporating all the information in the data. To do so, we follow Golan, Judge, and Miller and employ the GME estimation method.

Within the GME framework, the estimation procedure is converted into a constrained optimization problem, where the objective function consists of the joint entropy of the signal $\underline{\beta} \equiv (\underline{\beta}_t', \underline{\beta}_p', \underline{\beta}_c')'$ and the noise $\underline{\varepsilon} \equiv (\underline{\varepsilon}_t', \underline{\varepsilon}_p')'$. The constraints to this optimization problem are the system of N data points specified in Equations (1) and (2) as well as another set of normalization requirements discussed below.

Specifically, let $\beta_{jk} = \Sigma_m \, p_{jkm} z_{jkm}$, where $k = 1, 2, \ldots, K$, for $j = t$ or p and $k = 1, 2, \ldots, L$ for $j = c$, and $\underline{p}_{jk}$ is a M-dimensional proper probability vector corresponding to a M-dimensional vector of weights $\underline{z}_{jk} = (z_{j1}, z_{j2}, \ldots, z_{jm})'$. These support spaces are $M \geq 2$ dimensional vectors that serve as the discrete support spaces for each one of the K unknowns. Thus, each vector $\underline{\beta}_j$ is converted from the real line into a well-behaved set of proper probabilities, defined over the supports. If we have no prior knowledge as to the possible values of $\underline{\beta}$, we specify $\underline{z}$ to be equally spaced and symmetric around zero, with large negative and positive boundaries. If we have some knowledge about the possible range of values for $\underline{\beta}$, we can use this information to specify $\underline{z}$. For further discussion and sampling experiments see, for example, Golan, Judge, and Perloff.

Similarly, we transform the errors $\underline{\varepsilon}_j$, $j = t$ or p, into two sets of N proper probability distributions defined over some support space $\underline{v}$. Let $\underline{v}$ be a discrete support space of dimension $G \geq 2$, equally spaced and *symmetric around zero*. Now, define a set of proper probabilities (weights) $\underline{q} = (\underline{q}_{ti}', \underline{q}_{pi}')'$ such that $\varepsilon_{ji} = \Sigma_g \underline{v}_{ig}\underline{q}_{jig}$, $j = t$ or p. The boundary points for $\underline{v}_j$, say $\underline{v}_{j1}$ and $\underline{v}_{jG}$, are just $-3\sigma_j$ and $3\sigma_j$ where σ_j is the *empirical standard deviation* of q_j calculated from the data $\underline{w}_j$.

In the following applications, $M = 5$ and $G = 3$. The supports are $\underline{z}_{jk}$ = (-100, -50, 0, 50, 100)' and $v_j = (-3\sigma_j, 0, 3\sigma_j)'$ for $j = t$ or p. Thus, the supports are symmetric around zero for all the coefficients and for the error terms.

Having reparameterized both the signal and the noise parts of the system as two sets of proper probabilities, the GME sample selection model is

$$
\begin{aligned}
(3) \qquad \max_{\underline{p},\underline{q}} H(\underline{p}_t, \underline{p}_p, \underline{p}_c, \underline{q}_t, \underline{q}_p) = & -\sum_k \sum_m p_{tkm} \ln p_{tkm} - \sum_k \sum_m p_{pkm} \ln p_{pkm} \\
& - \sum_l \sum_m p_{clm} \ln p_{clm} - \sum_i \sum_g q_{tig} \ln q_{tig} - \sum_i \sum_g q_{pig} \ln q_{pig}
\end{aligned}
$$

subject to the data, where, if the individual works in the time-rate sector

$$
(4) \qquad \sum_k \sum_m X_{tik} z_{tkm} p_{tkm} + \sum_g v_{tg} q_{tig} = w_{ti}
$$

$$\sum_k \sum_m X_{tik} z_{tkm} p_{tkm} + \sum_g v_{ti} q_{tig} - \left[\sum_k \sum_m X_{pik} z_{pkm} p_{pkm} + \sum_g v_{pi} q_{pig}\right] > \sum_l \sum_m C_{il} z_{clm} p_{clm}, \tag{5}$$

if the individual works in the piece-rate sector

$$\sum_k \sum_m X_{pik} z_{pkm} p_{pkm} + \sum_g v_{pg} q_{pig} = w_{pi} \tag{6}$$

$$\sum_k \sum_m X_{tik} z_{tkm} p_{tkm} + \sum_g v_{tl} q_{tlg} - \left[\sum_k \sum_m X_{pik} z_{pkm} p_{pkm} + \sum_g v_{pi} q_{pig}\right] \leq \sum_l \sum_m C_{il} z_{clm} p_{clm}, \tag{7}$$

and the requirements of proper distributions that $\Sigma_m p_{jkm} = 1, j = t, p,$ or c, and $\Sigma_g q_{jlg} = 1, j = t$ or p. The optimization yields $\underline{\hat{p}} = (\underline{\hat{p}}_t', \underline{\hat{p}}_p', \underline{\hat{p}}_c')'$ and $\underline{\hat{q}} = (\underline{\hat{q}}_t', \underline{\hat{q}}_p')'$, which in turn yield the estimates $\underline{\beta} = (\underline{\beta}_t', \underline{\beta}_p', \underline{\beta}_c')'$ and $\underline{\hat{e}} = (\underline{\hat{e}}_t', \underline{\hat{e}}_p')'$ respectively.

The objective function (3) is a dual objective function that places equal weights on both prediction and precision of estimates, while shrinking all the estimates to the center of their supports. That is, the GME method shrinks the errors toward zero but does not force them to be identically zero, while the $\underline{\beta}$'s are "pushed" toward the center of their supports. Out of all known estimators that restrict the parameter space (such as restricted maximum likelihood) or the number of the moments of the distribution (any maximum likelihood model) or both, the GME uses the least amount of information: See Golan.

In Golan, Moretti, and Perloff, we conducted Monte Carlo experiments with a similar sample selection problem (wage equation if the individual is in the labor market). We find that, with small samples, the OLS, Ahn and Powell (AP, a semi-nonparametric method), and GME methods always provide estimates, but that Heckman's two-step and full-information maximum likelihood methods frequently fail to converge or provide estimates of the correlation coefficient that do not lie within the plausible range of [-1, 1]. Under all scenarios, the GME proved to be the most stable estimator (had the lowest variance and mean square errors) and dominated the other estimators where sample sizes were very small.

Inference

Let $\underline{\lambda} \equiv (\underline{\lambda}'_{te}, \underline{\lambda}'_{ti}, \underline{\lambda}'_{pe}, \underline{\lambda}'_{pi})'$ be the vector of Lagrange multipliers associated with Equations 4-7, where "*e*" and "*i*" stand for the "equality" (Equations 4 and 6) and "inequality" restrictions (Equations 5 and 7) for each sector. Define $H^*(\hat{\underline{\lambda}})$ as the maximum value of the objective function where $\hat{\underline{\lambda}} \neq \underline{0}$, or, equivalently $\hat{\underline{\beta}} \neq \underline{0}$. That is, $H^*(\hat{\underline{\lambda}})$ is the optimal value of (3) when all the constraints (data) are employed. Next, let $H_u(\hat{\underline{\lambda}})$ be the maximum possible value of the objective function when <u>no data constraints</u> are imposed, $\hat{\underline{\lambda}} = \underline{0}$. That is, $H_u(\hat{\underline{\lambda}})$ is the maximum possible value of the objective function where the only constraints imposed are the proper probability requirements. Thus, $H_u(\hat{\underline{\lambda}})$ is just the entropy value of the four sets of discrete uniform distributions, so $H_u(\hat{\underline{\lambda}}) = (2K + L) \ln M + 2N \ln G$.

Assuming (i) the errors' support v is symmetric around zero, (ii) $\underline{z}_k$ and $\underline{z}_l$ span the true values of each one of the unknown parameters $\underline{\beta}$ coefficients, (iii) the errors are iid, and (iv) the matrix X is of full rank, then the *entropy-ratio statistic* for testing the null hypothesis H_0: $\underline{\beta} = \underline{0}$ is $\mathscr{E}(\hat{\underline{\lambda}}) = 2[H_u(\hat{\underline{\lambda}}) - H^*(\hat{\underline{\lambda}})]$ and $\mathscr{E}(\hat{\underline{\lambda}}) \rightarrow \chi^2_{(2K+L-3)}$ as $N \rightarrow \infty$ when H_0 is true. The approximate α-level confidence intervals for the β are obtained by setting $\mathcal{E}(\underline{\hat{\lambda}}) \leq C_\alpha$ where C_α is chosen such that Prob $(\chi^2_{(2K+L-3)} < C_\alpha) = \alpha$. Similarly, any hypothesis of the form H_0: $\underline{\beta} = \underline{\beta}_0$ for all, or any subset, of the $2K + L$ parameters, can be tested.

A "goodness of fit" measure for the whole system is

$$R^{*2} \equiv 1 - S(\underline{\hat{p}}) = 1 - \frac{H^*(\underline{\hat{p}}_t, \underline{\hat{p}}_p, \underline{\hat{p}}_c)}{2K \ln M + L \ln M},$$

where $S(\underline{\hat{p}})$ is the normalized entropy measure. This normalized entropy measure is a continuous measure between zero and one. A measure of zero reflects perfect knowledge (no uncertainty), and a measure equal to one indicates a state of full ignorance where all the $\underline{\beta}$'s are zero, or at the center of their supports. Similarly, the normalized entropy measure for each part of the system $S(\underline{\hat{p}}_j) = H^*(\underline{\hat{p}}_j)$, $j = t, p$, or c, and the corresponding goodness of fit measure is $R^{*2}_j = 1 - S(\underline{\hat{p}}_j)$.

The asymptotic variance matrix is found by calculating $\hat{\sigma}^2_j = (1/N)\, \Sigma_i\, \hat{\varepsilon}^2_{ji}$ and the covariance term $\hat{\sigma}_{tp} = (1/N)\, \Sigma_i\, \hat{\varepsilon}_{ti}\, \hat{\varepsilon}_{pi}$, and observing that, given our four conditions,

$$\sqrt{N}\left(\underline{\hat{\beta}} - \underline{\beta}\right) \xrightarrow{d} N\left(\underline{0}, \left[A'\left(\hat{\Sigma}^{-1} \otimes \mathrm{I}\right) A\right]^{-1}\right),$$

where $\hat{\Sigma}$ is a 2 x 2 matrix of $\hat{\sigma}^2_t$, $\hat{\sigma}^2_p$, and $\hat{\sigma}_{tp}$ terms and A is a N x $2K$ matrix, where two X matrixes are on the main diagonal.

Empirical Results

The data used in this study are from the National Agricultural Workers Survey (NAWS), which is an annual survey of U. S. seasonal agricultural service workers (SAS). SAS workers, as defined by the U. S. Department of Agricultural, are most field workers in perishable crop agriculture. We use data from random sample interviews conducted in April and May of 1995. See Mines, Gabbard and Boccalandro for details on how the survey is conducted.

We want to examine how individuals decide whether to work in the piece-rate or time-rate sectors of the agricultural labor market, whether women are paid less then men in these sectors, and whether these earnings differentials vary geographically. Consequently, we estimate the same model for various regions of the country. In these models, wage depends on the X matrix which includes age and age squared; farm work experience and its square; and dummies for white (due to a lack of variation in most samples, we do not include dummies for black and Hispanic), females, and legal status (citizen, permanent resident [green card], amnesty recipient under the Immigration Reform and Control Act of 1986, or non-authorized worker). The C matrix includes these variables and whether the individual can speak English. For the Western Plains region, we drop the amnesty dummy due to lack of variation and include a dummy for Texas. We do not estimate the model for the North West region due to the lack of variation in many variables.

We estimated models of piece-rate and time-rate wage equations and (where relevant) selection equations for each region using the GME and four other models: ordinary least squares (which ignores selection), Heckman's two-step estimator, Heckman's full-information maximum likelihood estimator, and Ahn-Powell's (AP) method. The consistency of both of Heckman's estimators depend on the assumption of joint normality of the residuals, which may be violated in our samples. Neither Heckman model produces fully acceptable estimates for any region. In the following tables, we do not report estimates for Heckman's maximum likelihood estimator because it either fails to converge or its estimated correlation coefficient lies outside the [-1, 1] range for every region. We do list estimates for the Heckman two-step procedure even though the correlation between residuals of the selection equation and the wage in at least one sector lies outside [-1, 1], for each sector. Where such a violation occurs, we report a "constrained" correlation coefficient of -1.

The AP model uses a two-step estimator where both the joint distribution of the error term and the functional form of the selection equation is unknown. Because the AP estimator is robust to misspecification of the distribution of residuals and the form of the selection equation, we expect the AP estimator to perform better than Heckman's parametric two-step estimator for large samples. Whether the AP method has an advantage in small samples is not clear. (See Golan, Moretti, and Perloff for a detailed comparison of AP and GME models for small data set, sample-selection models.)

Table 1 reports estimates of the wage coefficients (and their associated asymptotic standard deviations) for the Mid West. Though the general sign patterns are similar across the models, the GME coefficients tend to have much smaller asymptotic standard errors than the other estimates — especially in the piece-rate sector, which has few observations. The coefficient patterns are generally similar to those found in the literature (e.g., Rubin and Perloff), but less precisely measured by the Heckman estimators, presumably because the earlier studies used larger samples than here.

For all models that we can, we calculate the R^2 goodness of fit measure for both wage equations using the same method as for ordinary least squares. The AP model does not have a goodness of fit statistic as it does not estimate constants.

The following outcome tables demonstrate how well the Heckman two-step model and the GME model predict the sector in which individuals choose to work:

	Predicted by Two-Step		*Predicted by GME*	
Actual	*time*	*piece*	*time*	*piece*
time	86	3	82	7
piece	11	6	1	16

The Heckman does slightly better at predicting the time-rate (the larger category) sector, but the GME does better in predicting the piece-rate sector. The GME does better overall, correctly predicting 92.5 percent compared to 86.8 percent for the two-step method. Results are similar in other regions. For example, in the Western Plains region (23 piece-rate and 54 time-rate observations), the Heckman model predicts 79.2 percent of the observations accurately, while GME predicts 98.7 percent correctly. The corresponding percentages are 69.5 percent and 93.4 percent for the South East (27 piece rate, 65 hourly) and 93.5 percent and 100 percent for California (221 piece rate, 37 hourly).

For ease in comparing the various models, the Heckman sample-selection probit equation contains the same variables as in the C matrix, which we use in the GME model to estimate the relative cost of being in the time-rate sector in each of the inequality restrictions (Equations 5 and 7). However, one might argue that only the constant term and the "extra" variable — the ability to speak English — belongs in the C matrix. The entropy-ratio test that the other nine coefficients are zero is 0.02, which is smaller than the critical value of χ^2_9 using a 0.05 criterion. Thus, we conclude that these other nine variables do not contain statistically significant information.

We also examined whether the female-male wage differential varies across the country. We expect these differentials to vary regionally because agricultural labor markets are regional (average wages differ substantially geographically), cover different crops, have different lengths of employment, and employ workers with different demographic characteristics. Table 2 shows the estimates of the coefficient on the female dummy for each estimated region. Because the left-hand variable is the logarithm of hourly earnings, these values are approximately the percentage difference between women's wages and men's. We find large differentials (unlike those in most existing studies) that vary substantially across regions. The GME estimates are closer to zero in most cases and have much smaller asymptotic standard errors than do the two-step estimates. The sign patterns for the two estimators are the same except for piece-rate workers in the Western Plains. The GME estimates indicate that women are paid substantially less than men except in the piece-rate sector in the Western Plains and the time-rate sector in California and that these differentials are statistically significant using a 0.05 criterion.

Conclusions

We present a practical alternative method for estimating sample selection models with small samples. Monte Carlo experiments in Golan, Moretti, and Perloff indicate that this method dominates traditional methods with small samples in the work-do not work sample-selection problem. Here, we show how to use the method to estimate the sector choice sample-selection problem. We apply this method to examine the choice between working in agricultural piece rate or in time rate jobs and the corresponding wage equations. The GME approach tends to have smaller estimated asymptotic standard errors and better explains sector choice than do ordinary least squares, Heckman's methods, and Ahn and Powell's approach in our small sample application.

References

Ahn, H. and J. L. Powell. "Semiparametric estimation of censored selection models with a nonparametric selection mechanism." *J. Econometrics* 58(July 1993):3-29.

Golan, A., G. Judge, and D. Miller. *Maximum Entropy Econometrics: Robust Estimation With Limited Data.* New York: John Wiley & Sons, 1996.

Golan, A., G. Judge, J. M. Perloff. "Recovering Information from Censored and Ordered Multinomial Response Data." *J. Econometrics* 79(July 1997):23-51.

Golan, A., E. Moretti, and J. M. Perloff. "A Small-Sample Estimator for the Sample-Selection Model." U. C. Berkeley Working Paper, 1998.

Golan, A. "Maximum Entropy, Likelihood and Uncertainty: A Comparison." G. J. Erickson, ed., *Maximum Entropy and Bayesian Methods*, Kluwer Academic Pub. 1998.Heckman, J. "Shadow Prices, Market Wages, and Labor Supply." *Econometrica* 42(1974):679-94.

Jaynes, E. T., "Information Theory and Statistical Mechanics" *Physics Rev.* 106(1957):620-630.

Levine, R. D. "An Information Theoretical Approach to Inversion Problems," *J. Physics* A 13(1980):91-108.

Mines, R., S. Gabbard, and B. Boccalandro. "Findings from the National Agricultural Workers Survey (NAWS) 1990." Office of Program Economics, Research Report No.

1. Office of the Assistant Secretary for Policy, U. S. Department of Labor. July 1991.

Rubin, D. K., and J. M. Perloff. "Who Works for Piece Rates and Why." *Amer. J. Agr. Econ.* 75(November 1993):1036-43.

Table 1. Wage Equations for the Mid West

	Piece Rate (N = 17)				*Time Rate* (N = 89)			
	OLS	*2-step*	*AP*	*GME*	*OLS*	*2-step*	*AP*	*GME*
Constant	0.766	0.993	-	0.884	1.164	1.135	-	1.189
	(0.900)	(0.808)	-	(0.127)	(0.161)	(0.162)	-	(0.135)
Age/100	5.575	7.278	6.545	4.699	2.466	2.338	2.467	2.465
	(6.513)	(5.640)	(9.259)	(0.815)	(1.005)	(0.999)	(0.696)	(0.866)
Age^2/10,000	-5.002	-5.501	-6.977	-3.635	-3.062	-2.819	-3.071	-3.247
	(9.831)	(8.753)	(13.558)	(1.102)	(1.357)	(1.370)	(0.932)	(1.171)
Farm Exp./100	-4.197	-6.901	-4.103	-3.263	0.969	1.155	0.984	0.315
	(6.378)	(4.969)	(5.505)	(0.789)	(1.286)	(1.268)	(0.960)	(0.839)
Farm $Exp.^2$/10,000	6.447	8.169	6.517	3.298	1.222	0.097	1.136	3.570
	(21.012)	(17.015)	(16.886)	(2.588)	(5.120)	(5.152)	(3.608)	(2.751)
White	-0.322	-0.410	-0.329	-0.352	0.184	0.209	0.196	0.208
	(0.260)	(0.229)	(0.172)	(0.060)	(0.127)	(0.126)	(0.068)	(0.064)
Female	-0.124	-0.356	-0.141	-0.136	-0.160	-0.164	-0.160	-0.092
	(0.185)	(0.205)	(0.184)	(0.044)	(0.057)	(0.056)	(0.049)	(0.047)
Citizen	0.230	0.290	0.214	0.232	-0.086	-0.076	-0.098	-0.112
	(0.200)	(0.225)	(0.263)	(0.066)	(0.140)	(0.134)	(0.072)	(0.070)
Amnesty	0.464	0.645	0.458	0.448	-0.146	-0.126	-0.144	-0.108
	(0.344)	(0.308)	(0.158)	(0.070)	(0.087)	(0.088)	(0.055)	(0.074)
Green Card	-0.062	-0.067	0.059	-0.103	-0.369	-0.419	-0.408	-0.370
	(0.357)	(0.299)	(0.291)	(0.083)	(0.135)	(0.143)	(0.070)	(0.089)
σ	0.26	0.318	-	0.17	0.19	0.19	-	0.18
R^2	0.61	0.75	-	0.62	0.42	0.43	-	0.86

Notes: The standard errors or asymptotic standard errors are in the parentheses. The covariance for the GME model is σ_{tp} = -0.0000007. For the 2-step estimator, the estimated inverse Mills ratio is -0.353 (0.180) for piece rate and -0.117 (0.156) for hourly, ρ_p = -1 (constrained), and ρ_t = -0.629.

Table 2. Female Wage Differential by Region and Sector

	Piece Rate		*Time Rate*	
	Two Stage	*GME*	*Two Stage*	*GME*
California	-0.551	-0.433	0.156	0.114
	(0.316)	(0.046)	(0.043)	(0.020)
Mid West	-0.356	-0.136	-0.164	-0.092
	(0.205)	(0.044)	(0.056)	(0.046)
Western Plains	-0.0001	0.183	-0.057	-0.041
	(0.283)	(0.041)	(0.034)	(0.022)
South East	-0.529	-0.327	0.121	-0.165
	(2.605)	(0.080)	(0.411)	(0.045)