DOCUMENT RESUME

ED 117 332

CE 005 900

AUTHOR Boskin, Michael J.; Nold, Frederick C.
TITLE Welfare Dependency and Low Income Labor Markets [and] A Markov Model of Turnover in Aid-to-Families with Dependent Children.
INSTITUTION Stanford Univ., Calif. Dept. of Economics.; Stanford Univ., Calif. Inst. for Mathematical Studies in Social Science.
SPONS AGENCY Manpower Administration (DOL), Washington, D.C. Office of Research and Development.; National Science Foundation, Washington, D.C.
PUB DATE Jul 75
NOTE 60p.

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage
DESCRIPTORS *Labor Market; *Low Income Groups; Minimum Wage; Models; *Statistical Analysis; Tables (Data); Unemployment; *Welfare; *Welfare Recipients

ABSTRACT

Two models of the duration of stay on welfare are developed and estimated using panel data from the California Aid to Families with Dependent Children AFDC panel survey. The first model characterizes the distribution of length of stay on welfare as drawn from the lognormal distribution with a truncation at the duration of the experiment (sixty months). The second model analyzes the movements on and off welfare, and duration of stay on welfare as a Markov process. Methodological findings indicate that statistical procedures fail to account for the special characteristics of the limited duration of observation and can be quite misleading. Substantive findings indicate that the welfare population as a whole involves an enormous turnover and modest length of stay. Those earning below the minimum wage are less likely to leave welfare and are more likely to return, stay off welfare for shorter periods and stay on for longer periods, and are more likely to be on welfare in a steady state than those earning above the minimum wage. Persons with high expected unemployment or low non-wage incomes respond similarly to those with below minimum wage earnings. Seven tables supplement the discussions. (Author/LH)

Welfare Dependency and Low Income Labor Markets

by

Michael J. Boskin
Department of Economics
Stanford University

The material in this project was prepared under Grant No. 51-06-73-06 from the Manpower Administration, U. S. Department of Labor, under the authority of title I of the Manpower Development and Training Act of 1962, as amended. Researchers undertaking such projects are encouraged to express freely their professional judgment. Therefore, points of view or opinions stated in this document do not necessarily represent the official position or policy of the Department of Labor.

STANDARD TITLE PAGE FOR TECHNICAL REPORTS

1. Report No.
DLMA

2. Govt Accession No.

3. Recipient's Catalog No.

4. Title and Subtitle

Welfare Dependency and Low Income Labor Markets

5. Report Date
July 1975

6. Performing Organization Code

7. Author(s)
Michael J. Boskin

8. Performing Organization Rept. No.

9. Performing Organization Name and Address

Department of Economics
Stanford University
Stanford, CA 94305

10. Project/Task/Work Unit No.

11. Contract/Grant No.
DL 51-06-73-06

12. Sponsoring Agency Name and Address
U.S. Department of Labor
Manpower Administration
Office of Research and Development
1111 20th St., N.W. Washington, D.C. 20210

13. Type of Report & Period Covered
Final

14. Sponsoring Agency Code

15. Supplementary Notes

Findings to be published in Journal of Human Resources, Fall 1975.

16. Abstracts

Two models of the duration of stay on welfare are developed and estimated using panel data from the California AFDC panel survey. The first model characterizes the distribution of length of stay on welfare as drawn from the lognormal distribution with a truncation at the duration of the experiment (sixty months). The second model analyzes the movements on and off welfare, and duration of stay on welfare as a Markov process.

The most important methodological finding of the study is that statistical procedures which fail to account for the special characteristics of the limited duration of observation and the piling up of the density at the limit can be quite misleading. Indeed, several frequently used statistical procedures impart a substantial downward bias to the coefficient estimates. (continued...)

17. Key Words and Document Analysis. 17a. Descriptors

economic analysis
statistical analysis
supply
unemployment

17b. Identifiers/Open-Ended Terms

welfare dependency
minimum wage

17c. COSATI Field/Group 5C

18. Distribution Statement
Distribution is unlimited. Available from National Technical Information Service, Springfield Va., 22151.

19. Security Class (This Report)
UNCLASSIFIED

20. Security Class (This Page)
UNCLASSIFIED

21. No. of Pages

22. Price

16. Abstracts continued.

The major substantive findings are that the popular notion of the welfare population as more or less permanently entrenched on welfare is erroneous (an enormous turnover accompanied by a modest average length of stay on welfare is more accurate); that persons facing a wage below the minimum wage are much less likely to leave welfare, much more likely to return, stay off welfare for shorter periods, stay on welfare for longer periods, and are much more likely to be on welfare in the steady state than those facing wages above the minimum. Persons with low non-wage income or high expected unemployment respond similarly to those with low wages.

WELFARE DEPENDENCY AND LOW INCOME LABOR MARKETS: ABSTRACT

by

Michael J. Boskin

Two models of the duration of stay on welfare are developed and estimated using panel data from the California AFDC panel survey. The first model characterizes the distribution of length of stay on welfare as drawn from the lognormal distribution with a truncation at the duration of the experiment (sixty months). The second model analyzes the movements on and off welfare, and duration of stay on welfare as a Markov process.

The most important methodological finding of the study is that statistical procedures which fail to account for the special characteristics of the limited duration of observation and the piling up of the density at the limit can be quite misleading. Indeed, several frequently used statistical procedures impart a substantial downward bias to the coefficient estimates.

The major substantive findings are that the popular notion of the welfare population as more or less permanently entrenched on welfare is erroneous (an enormous turnover accompanied by a modest average length of stay on welfare is more accurate); that persons facing a wage below the minimum wage are much less likely to leave welfare, much more likely to return, stay off welfare for shorter periods, stay on welfare for longer periods, and are much more likely to be on welfare in the steady state than those facing wages above the minimum. Persons with low non-wage income or high expected unemployment respond similarly to those with low wages.

Table of Contents

List of Tables

REGRESSION ANALYSIS WHEN THE DEPENDENT VARIABLE IS TRUNCATED LOGNORMAL, WITH AN APPLICATION TO THE DETERMINANTS OF THE DURATION OF WELFARE DEPENDENCY

by

Takeshi Amemiya and Michael Boskin[1]

1. Introduction

The purpose of the present paper is two-fold: to present an estimation technique which may be useful in a variety of econometric studies and to apply it to a new body of data in order to gain some insight into the determinants of the duration of welfare dependency.

The former is motivated by two considerations with respect to the distribution of the dependent variable. First, many variables must be non-negative to have any economic meaning. For example, theoretical considerations preclude negative consumption. Second, theoretical considerations and/or the method in which the data are collected may truncate the distribution. The first of these considerations may lead to the regression with the lognormal distribution developed by Amemiya [3]; the last may lead to the combination of probit and regression analysis

[1] The authors wish to thank Robert E. Hall for helpful comments, Thomas Moore and Frederick Nold for computational assistance and the Office of Research of the California Department of Social Welfare for providing the data. The research for the first author was supported by National Science Foundation Grant GS-2635 at the Institute for Mathematical Studies in the Social Sciences at Stanford University. The second author was supported by US Department of Labor, Manpower Administration, Grant No. 51-06-73-06 at Stanford University. The authors are indebted to Jon Peck and Dennis J. Aigner for calling their attention to errors in equations (2.11) and (2.14) in the earlier version of the paper and to the referee for his valuable comments which resulted in a considerable improvement of the paper.

developed by Tobin [10]. When both considerations are relevant simultaneously, the lognormal analog of Tobit analysis, or the truncated lognormal model, is an appealing prior specification.

For example, consider a study of the determinants of the duration of welfare dependency using data from a five-year survey of welfare recipients. The dependent variable in such a relationship, the number of months during the five-year period the household received welfare, must be positive and has an upper limit of sixty.[2] In addition, there is reason to suspect both a positively skewed distribution and a piling up of the density at the upper limit. Ordinary least squares takes none of these considerations into account; Tobit analysis can be easily modified to account for both the lower and upper truncation, but cannot account for the skewness of the distribution.

To be sure, the lognormal distribution is not the only possible means to take account of the non-negative and skewed dependent variable. But we have used the lognormal distribution because it has been extensively and successfully used to represent non-negative skewed random variables in many fields of application, including economics. For a good discussion of these applications the reader may consult Aitchison and Brown [1]. Another feature of the lognormal distribution, which is

[2] We conceive of this relationship as imbedded in a larger model which also explains the probability of being on welfare. Since data is available only for welfare families, we cannot estimate such a relationship. Our results therefore may be interpreted as estimating a relationship determining the duration of welfare dependency conditional upon being on welfare. See Boskin [6] for a discussion of this situation in the analogous case of market labor supply.

intuitively attractive, is that its variance is proportional to the square of its mean. The gamma distribution has a similar shape to lognormal and has the same feature regarding the variance, but we have not used it because the truncated gamma distribution is computationally much more cumbersome than the truncated lognormal.

Turning now to our second purpose, an empirical study of welfare recipients, we note the large number of conflicting hypotheses about the behavior of welfare recipients and explanations of the rapid increase in welfare rolls. Most of these can be reduced to estimates of the response of welfare recipients to the changes in their budget constraint caused by the welfare system. Welfare induces two changes in the budget constraint facing recipients: it guarantees a certain income at the zero earnings level and it taxes earnings by reducing the welfare payment a fraction of a dollar for each dollar earned. That is, welfare imposes the usual income and wage effects on the work-leisure choice. We present below estimates of, and tests of hypotheses about, the economic and demographic variables affecting the duration of welfare dependency.

Section 2 presents the truncated lognormal regression model together with the likelihood function and its first- and second-order derivatives.

Section 3 discusses the iterative procedure used to calculate the maximum likelihood estimates, the asymptotic distribution of the estimator, and the problem of an appropriate initial estimator.

Section 4 describes the data, and the definition, and generation, of the variables used in the empirical part of the study.

Section 5 presents our empirical results, documenting some important, and in some ways surprising, findings about welfare recipients. We present estimates of, and formal tests of hypotheses about, the determinants of the duration of welfare dependency. We also present a comparison of the estimates from our truncated lognormal procedure with the results from ordinary least squares and Tobit estimation.

Section 6 offers a brief summary and conclusion.

2. The Model and the Maximum Likelihood Estimates

We first define a sequence of lognormal random variables, $\{y_t^*\}$, and then define y_t as a random variable obtained by truncating y_t^* at a certain value. We define $\{y_t^*\}$ as follows:

(2.1) $\{y_t^*\}$ is independent lognormal with $Ey_t^* = \beta' x_t$ and

$$Vy_t^* = \eta^2(\beta' x_t)^2,\ t = 1,2,\ldots,T\ ,$$

where x_t is a K-component vector of known constants, β is a K-component vector of unknown parameters, and η^2 is a scalar unknown parameter. We assume that $\beta' x_t > 0$ for all t. Equivalently, we have

(2.2) $\{\log y_t^*\}$ is independent normal with $E\log y_t^* = \log \beta' x_t - \frac{\sigma^2}{2}$

and $V\log y_t^* (\overset{\text{def}}{=} \sigma^2) = \log(1+\eta^2)$.

Next we define y_t as

$$(2.3) \quad \begin{aligned} y_t &= y_t^* \quad \text{if } y_t^* < \alpha \\ &= \alpha \quad \text{if } y_t^* \geq \alpha , \end{aligned}$$

where α is a known positive constant.[3]

Our statistical problem is the estimation of a (K+1)-component vector of unknown parameters, (β', η^2), or equivalently, (β', σ^2) on the basis of the observations $y_1, y_2, \ldots, y_T$. We will consider the estimation of (β', σ^2) for the sake of mathematical convenience. We choose the maximum likelihood method of estimation.

Let S_1 be the set of t for which $y_t = \alpha$ and S_2 be that for which $y_t < \alpha$. Then, the likelihood function of $y_1, y_2, \ldots, y_T$ is given by

$$(2.4) \quad L = \prod_1 P[y_t^* \geq \alpha] \cdot \prod_2 g_t(y_t) ,$$

where g_t is the lognormal density given by

$$(2.5) \quad g_t(y_t) = \frac{1}{\sqrt{2\pi}\sigma y_t} \exp\left[- \frac{\left(\log y_t - \log \beta' x_t + \frac{\sigma^2}{2}\right)^2}{2\sigma^2} \right] ,$$

and $\prod_1$ is the product over $t \in S_1$ and $\prod_2$ is over $t \in S_2$. We note that

$$(2.6) \quad P[y_t^* \geq \alpha] = P[\log y_t^* \geq \log \alpha] = F\left(-\log \alpha + \log \beta' x_t^\circ - \frac{\sigma^2}{2}\right) ,$$

[3] In our case α is obviously a known constant. However, in some other applications it may be more appropriate to assume α to be an unknown parameter to be estimated. This would be an important model to consider but is beyond the scope of the present paper.

where F is the distribution function of $N(0, \sigma^2)$. Therefore, the log likelihood function is, aside from constants,

$$(2.7) \qquad \log L = \sum_1 \log F_t - \frac{1}{2}\sum_2 \log \sigma^2 - \frac{1}{2\sigma^2}\sum_2 v_t^2$$

where we use the abbreviations

$$(2.8) \qquad F_t = F(\mu_t) = F(-\log \alpha + \log \beta' x_t - \frac{\sigma^2}{2})$$

and

$$(2.9) \qquad v_t = \log y_t - \log \beta' x_t + \frac{\sigma^2}{2} .$$

The maximum likelihood equations are obtained by equating the partial derivatives of (2.7) with respect to β and σ^2 to 0 as follows: We have

$$(2.10) \qquad \frac{\partial \log L}{\partial \beta} = \sum_1 \frac{f_t}{F_t} \frac{1}{\beta' x_t} x_t + \frac{1}{\sigma^2} \sum_2 \frac{v_t}{\beta' x_t} x_t = 0 ,$$

where $f_t = f(\mu_t)$ is the density of $N(0, \sigma^2)$ evaluated at μ_t, and

$$(2.11) \qquad \frac{\partial \log L}{\partial \sigma^2} = -\frac{1}{2\sigma^2}\sum_1 \frac{\mu_t f_t}{F_t} - \frac{T_2}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_2 v_t^2 - \frac{1}{2}\sum_1 \frac{f_t}{F_t} - \frac{1}{2\sigma^2}\sum_2 v_t = 0 ,$$

where T_2 is the number of elements in S_2. The maximum likelihood estimates of β and σ^2 are defined as the roots of (2.10) and (2.11).

The solution of (2.10) and (2.11) must be obtained by an iterative procedure, which we will explain in the next section.

We will need the following second-order derivatives both in the iterative procedure and in obtaining the asymptotic variance-covariance matrix of the maximum likelihood estimates. We have

$$\text{(2.12)} \qquad \frac{\partial^2 \log L}{\partial\beta\partial\beta'} = -\frac{1}{\sigma^2}\sum_1 \frac{\mu_t f_t F_t + \sigma^2 f_t(f_t+F_t)}{F_t^2(\beta' x_t)^2} x_t x_t' - \frac{1}{\sigma^2}\sum_2 \frac{1+v_t}{(\beta' x_t)^2} x_t x_t' ,$$

$$\text{(2.13)} \qquad \frac{\partial^2 \log L}{\partial\sigma^2\partial\beta} = \frac{1}{2\sigma^4}\sum_1 \frac{(\mu_t+\sigma^2)(\sigma^2 f_t^2+\mu_t f_t F_t)-\sigma^2 f_t F_t}{F_t^2 \beta' x_t} x_t + \frac{1}{2\sigma^4}\sum_2 \frac{\sigma^2-2v_t}{\beta' x_t} x_t ;$$

and

$$\text{(2.14)} \qquad \frac{\partial^2 \log L}{\partial(\sigma^2)^2} = -\frac{1}{4\sigma^4}\sum_1 \frac{f_t}{F_t^2}[(\mu_t+\sigma^2)^2(\frac{\mu_t F_t}{\sigma^2}+f_t)-(3\mu_t+2\sigma^2)F_t]$$

$$+ \frac{T_2}{2\sigma^4} - \frac{T_2}{4\sigma^2} + \frac{1}{\sigma^4}\sum_2 v_t - \frac{1}{\sigma^6}\sum_2 v_t^2 .$$

3. <u>The Iterative Procedure</u>

Let $\theta = (\beta', \sigma^2)'$. Then, the well-known iterative procedure called the Newton-Raphson method is defined as follows: given an initial estimate $\hat{\theta}_1$, we define $\hat{\theta}_2$ by

$$\text{(3.1)} \qquad \hat{\theta}_2 = \hat{\theta}_1 - \left[\frac{\partial^2 \log L(\hat{\theta}_1)}{\partial\theta\partial\theta'}\right]^{-1} \frac{\partial \log L(\hat{\theta}_1)}{\partial\theta} .$$

After obtaining $\hat{\theta}_2$, one iterates again to obtain $\hat{\theta}_3$, continuing the process until $\hat{\theta}_n$ converges within the prescribed bounds.

Let $\hat{\hat{\theta}}$ be the estimate obtained by the above procedure. Then, clearly, $\hat{\theta}$ is a root of equations (2.10) and (2.11). Unfortunately, however, there is no assurance that $\hat{\theta}$ will correspond to the root for which log L attains the global maximum since (2.10) and (2.11) are highly nonlinear and therefore very likely to have multiple roots. We have not been able to obtain a useful set of conditions on the initial estimate for the iteration to converge to such a root. We might add that it is extremely difficult to obtain a practically verifiable set of conditions for the convergence of an iterative procedure in nonlinear models in general and consequently there are few such results in the literature. For general discussions of iterative procedures, the reader is referred to Jacoby, Kowalik, and Pizzo [9] and Goldfeld and Quandt [7].

If a consistent estimator were available as the starting value of the iteration, the convergence to the global maximum likelihood estimator would be more readily attained. However, we have not been able to find a practical consistent estimator for our model. Amemiya [2] proposed a simple consistent estimator for the truncated normal regression model but his estimator does not extend to the truncated lognormal model.

Given these circumstances, the best strategy seems to be to try a few different starting values. If they all converge to the same value and if moreover that value is a reasonable one in view of our *a priori* knowledge, we can get some assurance that the value obtained is indeed the global maximum likelihood estimator. In the empirical results discussed in Section 5 we have used the ordinary least squares, the weighted least squares, and a certain arbitrary value as the starting value of the iteration and all of them have produced the same reasonable value.

The ordinary least squares estimator of β in the regression of y_t on x_t using only those observations for which $y_t < \alpha$ is defined as

$$(3.2) \qquad \hat{\beta}_1 = (\sum_2 x_t x_t')^{-1}(\sum_2 y_t x_t)$$

The corresponding initial estimator of σ^2 is given by

$$(3.3) \qquad \hat{\sigma}_1^2 = \log(1 + \hat{\eta}_1^2)$$

where

$$(3.4) \qquad \hat{\eta}_1^2 = \frac{1}{T_2} \sum_2 \left[\frac{y_t - \hat{\beta}_1' x_t}{\hat{\beta}_t' x_t}\right]^2$$

The formula (3.4) is suggested by an inspection of (2.1). The weighted least squares estimator of β is defined as

$$(3.5) \qquad \beta^* = \left[\sum_2 \frac{1}{(\hat{\beta}_1' x_t)^2} x_t x_t'\right]^{-1} \sum_2 \frac{1}{(\hat{\beta}_1' x_t)^2} y_t x_t$$

The corresponding estimator of η^2 must be modified accordingly using (3.4).[4] These estimates are neither unbiased nor consistent.

Now, suppose that $\hat{\theta}$ obtained by the iterative procedure is indeed the value of θ that globally maximizes the likelihood function. Then we can expect that under reasonable conditions $\hat{\theta}$ is consistent and asymptotically normal with the asymptotic variance-covariance matrix approximated by

$$-\left[\frac{\partial^2 \log L}{\partial\theta\partial\theta'}\bigg|_{\hat{\theta}}\right]^{-1}$$

Amemiya [2] proved the consistency and the asymptotic normality of the maximum likelihood estimator for the truncated normal regression model. Because of a close similarity between the log likelihood functions of the two models, the same can be proved for our model by following each step of Amemiya's proofs closely. Thus, a hypothesis about θ can be tested using the asymptotic normal distribution.

A hypothesis about θ may be also tested by using the likelihood ratio test. Let the null hypothesis be $g_i(\theta) = 0$, $i = 1,2,\ldots,K$. Then, as is well known,

$$-2\log \frac{\sup_{\theta\in S} L(\theta)}{\sup_{\theta} L(\theta)}$$

[4] For more discussion of regression with the lognormal dependent variable, see Amemiya [3].

is asymptotically distributed as χ^2 with the degrees of freedom equal to K, the number of restrictions.

As one can see from (3.1), the iterative procedure requires the evaluation of the first-order and second-order partial derivatives given in (2.10) - (2.14). For computational purposes it will be convenient to make the following substitutions in these equations:

(3.6) $$F(\mu_t) = \Phi(\frac{\mu_t}{\sigma}) \quad ,$$

$$F(\mu_t) = \frac{1}{\sigma}\, \phi(\frac{\mu_t}{\sigma}) \quad ,$$

where Φ and ϕ are the distribution and density functions respectively of the standard normal variate. The ordinate of Φ is calculated by means of expansion.

It is possible that in early stages of iteration $\beta_n' x_t$ may become negative for some t. Then, log $\beta_n' x_t$ cannot be defined. To avoid this difficulty, we will arbitrarily put log $\beta_n' x_t = 0$ whenever $\beta_n' x_t$ is negative. This procedure did not cause a serious problem in our empirical example discussed in the following sections because we had $\beta_n' x_t > 0$ for all t for n larger than a certain finite integer n_o. For, then, we can pretend as though we started the iteration with $\hat{\theta}_{n_o}$ so that the subsequent iteration is exactly the Newton-Raphson.

4. The Data

The data used in this study are survey data from the State of California AFDC five-year survey. The data follow individual households over the period 1965-70. The data include information on the total time on welfare, recidivism, and characteristics of household members such as age, race, sex, education, work experience, health, income by source, assets, and amount of aid. The data cover 658 households, and are particularly important as one of the few examples of data on individuals over time. These families all came on welfare in the first month of the study.

The variables used in the empirical study include:

TOA, time on aid, the total number of months during the five-year period during which the household received payments under AFDC. This is the dependent variable of the model and we assume it to be statistically independent among individuals given the values of the independent variables. Of course, this variable can only take on values between one and sixty.

W, the expected hourly market wage facing the household head. This wage is imputed on the basis of a hedonic regression of wages on personal characteristics from the 1967 Survey of Economic Opportunity, and adjusted for the employee component of the payroll tax and the individual income tax. The procedure was developed by Hall [8] and is also discussed in Boskin [5].

U, the expected duration (in weeks) of unemployment facing the head of the household. This is imputed in a manner analogous to W, and was also developed by Hall.

NWY, non-wage income, encompasses income from sources other than earnings, including imputed income (at 12%) to consumer durables.

A, a dummy variable taking the value one when the head of the household is twenty-five years old or under as of the start of the survey.

H, a dummy variable taking the value one when the head of the household reported that an adverse health condition affected employability as of the start of the survey.

PS, a dummy variable taking the value one if a child of preschool age, five years old or under, was present in the household for at least three of the five years of the study.[5]

5. Empirical Results

Our empirical results reveal some interesting insights into welfare dependency, as well as methodological points of interest. A first point of interest is that only 113 of the 658 households were on AFDC for the entire period. This amounts to just seventeen percent of the sample. In 1970, just 213 of the 658, or just thirty-two percent, were receiving AFDC payments despite a serious deterioration in employment prospects over the five-year period. Just under thirty percent of the households came back on welfare after having been off aid

[5] Of course, a very small number of families without a child of preschool age at the start of the survey may have had children born during this period. No information is available on this question; it is assumed no such births occurred.

at least once. These data suggest a continuous turnover of the welfare population rather than the familiar stereotype of a permanently entrenched welfare population. A very large number of cases show up on the welfare rolls only for a brief period of time (almost half the sample was on aid for no more than one year). What then are the factors influencing the duration of welfare dependency?[6]

We start out by presenting our results inclusive of all variables described above. The wage, unemployment and non-wage income variables describe the economic constraints facing the welfare recipient in the labor market.[7] The health, age and pre-school child represent demographic factors which may tend to keep the head of the household from working.

Table 1 presents the results based on our truncated lognormal regression model. Overall, the equation does quite well, the standard error being less than one-tenth of the mean of the left-hand variable. However, four of the individual coefficients, those for NWY, H, PS, and A are both quite small and, given this, not measured precisely enough to be considered different from zero. That is, in each case, we accept the hypothesis, on the basis of the t-test implicit in (3.2),

[6] We repeat the proviso of footnote 2.

[7] It should be pointed out that a small percentage of welfare recipients earned a modest amount while on welfare. That is, a minor fraction of the adjustment to the equilibrium quantity of labor supply took place while on welfare. The vast majority, however, took place by moving on and off the welfare rolls; given the high implicit tax on earnings under AFDC (see Barr and Hall [4]), this is perfectly rational.

that the coefficient equals zero. As noted above, we may also employ a likelihood ratio test to test (more powerfully) the composite hypothesis that all four coefficients equal zero. The statistic $-2 \ln \lambda$, where λ is the ratio of the maximum of the likelihood function over the restricted to the unrestricted parameter space, is distributed as χ^2_K, where K is the number of restrictions. In our case we have four degrees of freedom. The restricted maximum is generated by the regression reported in Table 2. We note that $-2 \ln \lambda$ is 5.2, which is about half the critical value of the χ^2_4 at the 5% level. We thus find no evidence for age, health, pre-school children, or non-wage income influencing the duration of welfare dependency. The coefficient on the expected duration of unemployment, U, suggests a modest increase in the expected time on welfare as expected unemployment rises; the implied elasticity, calculated at median values is about three-fourths.

The results reported in Tables 2 and 3 also reflect the collinearity of the constant term and the other dummy variables, especially PS. When these are dropped from the estimation equation, the estimated coefficient for the constant rises substantially.

Table 1

Truncated Lognormal Regression: Dependent Variable: TOA

Variable	Coefficient	Standard Error
C	48.19	20.6
W	-18.81	6.06
NWY	0.02×10^{-2}	0.10×10^{-2}
U	5.93	2.86
H	- 3.71	8.60
PS	6.71	13.09
A	0.46	0.75

$s^2 = 2.19$

convergence achieved after fifteenth iteration

Table 2

Truncated Lognormal Regression: Dependent Variable: TOA

Variable	Coefficient	Standard Error
C	62.00	12.91
W	-16.90	5.32
U	5.73	2.86

$s^2 = 2.12$

convergence achieved after fifteenth iteration

Table 3

Alternative Specifications of the Regression of TOA on W

	Variable	Coefficient	Standard Error
A. Truncated lognormal:			
	C	78.67	15.14
	W	-19.92	6.83
$s^2 = 2.17$			
B. Tobit: (with the upper truncation)	C	45.10	4.86
	W	-10.26	3.08
$s^2 = 1.81$			
C. Ordinary Least Squares: (non-limit observations)	C	25.27	2.91
	W	- 4.38	1.53
$s^2 = 0.66$			

Our major substantive conclusion is that expected market wage rates do exert an important influence on the duration of time on aid. As expected wage rates rise, time on aid falls off sharply. The wage elasticity, evaluated at the median wage and time on aid, is over unity. The interpretation of this result is not difficult. The implied wage elasticity of labor supply suggests that higher-wage recipients remain on aid for a shorter time as part of a long-run adjustment of labor supply. The lower the wage, of course, the more likely it is that remaining on welfare is the income-maximizing strategy. Our results suggest that the way to induce AFDC heads of households to work in the market[8] is to raise their expected wage, for example by a wage subsidy or negative payroll tax.

[8] We leave aside the question of whether, or when, this is desirable social policy.

We also find a modest effect of expected unemployment on welfare dependency.

Finally, it is instructive to compare the estimates of the wage elasticity of the duration of welfare dependency derived from different estimation procedures. Regression results are reported in Table 3 [9] and the wage elasticities derived in Table 4. The point estimates differ markedly, the ordinary least squares estimate being about one-quarter, and the Tobit estimate about one-half, of the truncated lognormal estimate. Three regressions are graphed in Figure 1. They show that both OLS and Tobit underestimate the expected value of TOA over the relevant range of W. That OLS using only non-limit observations underestimates the expected value of TOA can be rigorously explained as follows: We have

$$(5.1) \qquad E(y_t^*|y_t^* < \alpha) = \frac{1}{1-F_t}\int_0^\alpha y g_t(y)dy$$

where g_t and F_t are defined in (2.5) and (2.8) respectively. But, using the substitution $z = \log y$ to evaluate the integral, the right-hand side of (5.1) can be shown to be equal to

$$\frac{F(\log\alpha - \log\beta' x_t - \frac{\sigma^2}{2})}{F(\log\alpha - \log\beta' x_t + \frac{\sigma^2}{2})} \cdot \beta' x_t \quad ,$$

[9] The Tobit and ordinary least squares regressions on the full set of variables each yielded virtually identical estimates of the wage coefficient as those reported in Table 3.

assumes normal when it is truly lognormal, one will underestimate $P[y_t^* \geq \alpha]$. That means Tobit will attach too little weight to the information contained in the observations at α and too much weight to that contained in the observations below α. Hence the result we wanted to explain.

6. Conclusion

We have presented an estimation technique which may be valuable in econometric studies when the dependent variable satisfies certain criteria and applied it to an empirical problem which is not without interest itself.

Our major substantive conclusion is that the wage elasticity of the duration of welfare dependency is probably in the neighborhood of unity. This result lends strong support to those who advocate a program which increases the expected market wage of potential welfare recipients as a method of inducing participation in market work.

In addition, a comparison with ordinary least squares and Tobit estimation suggests that when the dependent variable must be non-negative, yet has an upper limit (as is natural in studies of the allocation of time), the data tend to flatten out the regression line when estimated by these methods in order to account for such phenomena.

Bibliography

[1] Aitchison, J. and J.A.C. Brown, The Lognormal Distribution (Cambridge: Cambridge University Press, 1969).

[2] Amemiya, T., "Regression Analysis When the Dependent Variable is Truncated Normal," Econometrica, forthcoming.

[3] Amemiya, T., "Regression Analysis When the Variance of the Dependent Variable is Proportional to the Square of Its Expectation," Journal of the American Statistical Association, forthcoming.

[4] Barr, N. and R. Hall, "The Taxation of Earnings Under Public Assistance," Working Paper No. 85, M.I.T., 1972.

[5] Boskin, M., "A Conditional Logit Model of Occupational Choice," Journal of Political Economy, forthcoming (March 1974).

[6] Boskin, M., "The Economics of the Labor Supply," forthcoming in G. Cain and H. Watts, eds., Labor Supply and Income Maintenance (New York: Rand McNally, 1973).

[7] Goldfeld, S. M. and R. E. Quandt, Nonlinear Methods in Econometrics (Amsterdam: North-Holland Publishing Company, 1972).

[8] Hall, R., "Wages, Income and Hours of Work in the U.S. Labor Force," forthcoming in G. Cain and H. Watts, eds., Labor Supply and Income Maintenance (New York: Rand McNally, 1973).

[9] Jacoby, S.L.S., J. S. Kowalik and J. T. Pizzo, Iterative Methods for Nonlinear Optimization Problems (Englewood Cliffs: Prentice-Hall, Inc., 1972).

[10] Tobin, J., "The Estimation of Relationships for Limited Dependent Variables," Econometrica, 26(January, 1958), 24-36.

A MARKOV MODEL OF TURNOVER IN AID-TO-FAMILIES WITH DEPENDENT CHILDREN

by

Michael J. Boskin and Frederick C. Nold

Technical Report No. 125

March 1974

Prepared Under

NATIONAL SCIENCE FOUNDATION GRANT GS-39906

THE ECONOMICS SERIES

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES
Fourth Floor, Encina Hall
Stanford University
Stanford, California
94305

A MARKOV MODEL OF TURNOVER IN AID-TO-FAMILIES WITH DEPENDENT CHILDREN*

by

Michael J. Boskin and Frederick C. Nold**

1. Introduction

Among the many badly mistaken popularly held views about welfare, especially Aid-To-Families-With-Dependent Children (AFDC), is the view that the population of recipients is more or less permanently entrenched in a welfare dependency status. We shall present, and analyze, data which suggest that nothing could be further from the truth. There is an enormous turnover in the welfare population: new families come on welfare and go off continuously; most families stay on welfare for periods of far shorter duration thatn is commonly supposed; and finally, there is a substantial amount of intermittent recidivism. This paper will be devoted to a study of this turnover in AFDC.

Aside from enabling us to dispel an incorrect popular view of AFDC, a study of turnover has important implications for economic policy and its administration. There are important social costs and benefits

---

*This paper represents a revised version of Memorandum No. 150 of the Center for Research in Economic Growth, Stanford University, 1973.

**The authors wish to thank Takeshi Amemiya and T. W. Anderson for valuable advice, Leonard Carlson for research assistance and Thomas Moore for computational assistance. The authors were supported, respectively, by U.S. Department of Labor, Manpower Administration, Grant No. 51-06-73-06 and by National Science Foundation Grant GS-39906 at the Institute for Mathematical Studies in the Social Sciences, Stanford University. The authors wish to thank the Manpower Administration and NSF for this support.

associated with the turnover in the recipient population even when compared to a situation with little turnover but the same number of family-months on AFDC. First, and probably most obvious, is the substantially higher cost of administering a program with substantial turnover. Less obvious, but possibly more important, are the costs and benefits to society associated with the frequent turnover in the allocation of work between home and market (see Becker [1965]) implied by the turnover in AFDC. Benefits associated with such turnover include the efficient use of time in job search (see Phelps, et. al. [1970]) and other types of human capital activities. Costs include the imputed time costs to the mother in moving from one use of her endowment of time to another and the necessity of continuously reacquiring specific human capital on each job. These factors should weigh heavily in any policy decisions likely to affect turnover (see the essay by Harberger [1971b] for a discussion of the basic principles).

There is also a duality between turnover and duration of time on welfare. After all, a family accumulates time on welfare only by coming on and (for some number of periods) failing to go off again. Thus, we gather information about duration on (and off) welfare by analyzing turnover.[1/] There are, of course, important benefits and costs associated with changes in the duration of time on welfare. If we induce welfare recipients out of work in the home into work in the market, the social opportunity cost is not, as is implicitly argued by some, zero. After one accounts for a variety of special circumstances (see the discussion in Harberger [1971a]), the relevant social opportunity cost is the prevailing wage rate (or the minimum wage) or something very close to it. The value of the time spent by AFDC mothers working in the home (raising and caring for their children, for example) should not be ignored in such policy decisions.

In the present paper, we present and analyze quasi-longitudinal data on welfare families. In doing so we hope to provide a more accurate description of the AFDC population than is currently available, develop a statistical technique for analyzing welfare turnover and duration, and provide some empirical estimates of the economic determinants of welfare dependency and turnover which will provide some insight into the behavior of the welfare population and some potential input into intelligent policy in this area.

Toward this end, Section 2 presents a two-state (on or off welfare) Markov model of welfare turnover; Section 3 describes the data used in the empirical anlaysis; Section 4 reports the empirical results of the study of welfare turnover; Section 5 offers a brief conclusion; and the Appendix develops the maximum likelihood estimators of our model, together with their distributions.

2. A Two State Markov Chain Model of Welfare Dependency

We shall adopt the model that the movements of individual s, s = 1,...,S, between state 1, on welfare, and state 2, off welfare, are stochastic and governed by the following probabilities:

$$\Pr\left\{\begin{array}{l}\text{individual } s \text{ moves from state 1 at time } t\\ \text{to state 2 at time } t+1\end{array}\right\} = \alpha_s$$

$$\Pr\left\{\begin{array}{l}\text{individual } s \text{ moves from state 2 at time } t\\ \text{to state 1 at time } t+1\end{array}\right\} = \beta_s$$

Note that we have implicitly introduced four distinct assumptions: (i) the probabilities of transitions between the states are independent of time;[2/] (ii) the transition probabilities are not dependent upon which states were occupied before t;[3/] (iii) t is taken to be discrete; (iv) only one movement can occur in a unit of time.[4/] These assumptions lead us to adopt the following convention: just after the beginning of each time period a Bernoulli trial is conducted in which the probability of transition is determined by the state the individual occupies. The outcome of that trial determines the state the individual will occupy until the beginning of the next period, at which time another trial is performed.

We must now take cognizance of our particular initial conditions. As we shall indicate below, a person become part of the survey we use by entering state 1 in the first month of the sample period. For reasons of mathematical convenience, we shall label the starting point of the survey $t = -1$. Then, at the beginning of the next period, $t = 0$, each individual is in state 1. Let $P_i^s(t)$ be the probability that individual s occupies state i at time t; $i = 1,2$, for all t. We have $P_1^s(0) = 1$, $s = 1,\ldots,S$ and $P_1^s(t) + P_2^s(t) = 1$, for all t. Individual

s can be in state i at time $t+1$, $t \geq 0$ in one of two ways: he could have entered state i on or before time t and remained there, or he could have just entered state i from state j, $i \neq j$. Hence, we have for $t \geq 0$

$$P_1^s(t+1) = (1-\alpha_s)P_1^s(t) + \beta_s P_2^s(t) \quad ,$$

$$P_2^s(t+1) = \alpha_s P_1^s(t) + (1-\beta_s)P_2^s(t) \quad .$$

Using the initial conditions, these difference equations can be solved uniquely yielding:

$$P_1^s(t) = (1-\alpha_s-\beta_s)^t + \frac{\beta_s}{\alpha_s+\beta_s}[1 - (1-\alpha_s-\beta_s)^t] \quad ,$$

$$P_2^s(t) = \frac{\alpha_s}{\alpha_s+\beta_s}[1 - (1-\alpha_s-\beta_s)^t] \quad , \quad t \geq 0 \quad .$$

We need these probabilities in the Appendix to calculate the expected values of several random variables. However, they can also help us answer a question in this section: the expected percentage of a time period of length T individual s will spend off welfare. Let $T_2^s(t)$ be the amount of time s spends off welfare up to time t and define $\Delta T_2^s(t) = T_2^s(t) - T_2^s(t-1)$. Then,

$$P_r\{\Delta T_2^s(t) = 1\} = \alpha_s P_1^s(t-1) + (1-\beta_s)P_2^s(t-1) \quad .$$

Using the probabilities derived above we obtain

$$\mathcal{E}T_2^s(T) = \sum_{t=1}^{T} [\alpha_s P_1^s(t-1) + (1-\beta_s)P_2^s(t-1)]$$

$$= \frac{\alpha_s}{\alpha_s+\beta_s}(T-1) + \frac{\alpha_s(\alpha_s+\beta_s-1)}{(\alpha_s+\beta_s)^2}[1 - (1-\alpha_s-\beta_s)^T] \ .$$

For large T, the expected percentage of time s is off welfare is approximately $\alpha_s/(\alpha_s+\beta_s)$.

Note that

$$\lim_{T\to\infty} P_2^s(T) = \lim_{T\to\infty} \frac{\mathcal{E}T_2^s(T)}{T} = \frac{\alpha_s}{\alpha_s+\beta_s}$$

$$\lim_{T\to\infty} P_1^s(T) = \lim_{T\to\infty} \frac{T-\mathcal{E}T_2^s(T)}{T} = \frac{\beta_s}{\alpha_s+\beta_s} \ .$$

Another question concerns the expected duration of a stay in a state. Once the individual is in state i, the probabilities of his making a transition are fixed and the duration of his stay follows the geometric distribution. For example, suppose individual s has just entered state 1 and let Y_s denote the duration of stay of individual s in that state. Then

$$\Pr\{Y_s = t\} = \alpha_s(1-\alpha_s)^{t-1} \ .$$

The mean and variance of Y_s are α_s^{-1} and $(1-\alpha_s)\alpha_s^{-2}$, respectively.

Note that for small α_s the mean and standard deviation of Y_s have approximately the same value.

Finally, we are interested in studying turnover. In the context of this model, turnover means a transition between states. Using methods similar to those employed above, we can establish that the expected number of times individual s leaves welfare is approximately $[\alpha_s\beta_s/(\alpha_s+\beta_s)]T$, for large T, where T is the length of time period we are considering.

The model may be summarized as follows. Once we know an individual's α_s and β_s we can obtain the distributions of the random variables which can be used to characterize the individual's welfare dependency. We have shown ~~that expected~~ duration, expected transitions and expected proportion of time in each state are simple functions of α_s and β_s. It should be noted that if α_s and β_s are increased by equal proportional amounts, the expected number of transitions increases by approximately that amount, but the expected proportion of time in the states, which is a function of the relative magnitudes of α_s and β_s, remains unchanged.

We shall assume that the probabilities of the transitions for each individual are functions of that individual's socioeconomic characteristics,[5/] and we adopt the logistic functional form for this dependence:

$$\alpha_s = \frac{\exp(\theta' X_s)}{1+\exp(\theta' X_s)} = f(\theta' X_s), \quad s = 1,\ldots,S ,$$

and

$$\beta_s = \frac{\exp(\Gamma' Z_s)}{1+\exp(\Gamma' Z_s)} = f(\Gamma' Z_s) \quad , \quad s = 1,\ldots,S \quad ,$$

where X_s and Z_s are, respectively, $K\times 1$ and $P\times 1$ vectors of exogenous variables which measure the socioeconomic characteristics of individual s, Θ and Γ are $K\times 1$ and $P\times 1$ vectors of unknown parameters, and S is the total number of individuals in the sample.

We estimate Θ and Γ in our logistic regression model with a maximum-likelihood procedure described in detail in the Appendix. Before proceeding to discuss the results of this estimation, we turn to a discussion of the data which, we shall see, played a major role in determining the way we modelled our problem.

3. The Data

The data used in this study are survey data from the State of California AFDC five-year survey. The data follow individual households over the period 1965-70 and include information on the total time on welfare, recidivism, and characteristics of household members such as age, race, sex, education, work experience, health, income by source, assets, and amount of aid. The data cover 440 households, and are particularly important as one of the few examples of data on individuals over time.[6/] These families all came on welfare in the first month of the study. In the data described below, we treat each month of the survey as a potential transition.

The variables used in the empirical study include:

T_{11}, the number of times the individual remained on welfare over the sample period.

T_{22}, the number of times the individual remained off welfare over the sample period.

T_{12}, the number of times the individual went off welfare during the five years of the study.

T_{21}, the number of times the individual came on welfare.

T , the total number of potential transitions off and on welfare, which is equal to sixty, one for each month of the survey.

The remaining variables were constructed from the basic data provided in the survey:

W , the expected hourly market wage facing the household head. This wage is imputed on the basis of a regression of wages on personal characteristics from the 1967 Survey of Economic Opportunity, and adjusted for the employee component of the payroll tax and the individual income tax. The independent variables in the wage equation include age, sex, race, education, location, union membership, health and interactions among these variables. The procedure is discussed in more detail in Boskin [1974 or Hall [1973]. The wage is also used in the form of a series of dummy variables representing wage categories (see the discussion in Section 4).

U , the expected duration (in weeks) of unemployment facing the head of the household. This is imputed in a manner analogous to W, and is also discussed in Boskin [1974] or Hall [1973]. Again, unemployment is sometimes entered categorically, rather than continuously.

NWY, non-wage income, encompasses income from sources other than earnings, including imputed income (at 12%) to consumer durables; sometimes entered as a dummy variable dividing the sample at the mean non-wage income.

A , a dummy variable taking the value one when the head of the household is twenty-two years old or under as of the start of the survey.

H , a dummy variable taking the value one when the head of the household reported that an adverse health condition affected employability as of the start of the survey.

PS , a dummy variable taking the value one if a child of pre-school age, five years old or under, was present in the household for at least three of the five years of the study.

It is instructive to examine the frequency distribution of the number of times individuals came onto welfare (which either equals, or exceeds by one, the number of times they went off) during the period. Presented in Table 1 below, these figures quickly dispel a popular misconception about welfare: that the population is rather firmly entrenched in a permanent welfare dependency status. Indeed, we note that only seventeen percent of this sample remained on welfare for the entire period. Fully twenty-seven percent went off and came back on again during the period. Combined with a median duration of the total time on welfare during the period of just fourteen months, we may infer that the popular conception is badly mistaken: there is an enormous turnover in the AFDC rolls; the flow of people over even a period as short as fixe years is much larger than the stock at any point in time, and many families remain on welfare for comparatively short periods. These phenomena are anlayzed in more detail below.

Table 1

Distribution of Number of Times Came on Welfare

Number of Times	% of Sample	% Time on Welfare
1	73	37
2	20	47
3 or more	7	50

4. Empirical Results

Our empirical results reveal some extremely interesting insights into welfare turnover and dependency. We shall present our results in several stages. First, we present the maximum likelihood estimates of the logistic regression of the probabilities of leaving and entering welfare status as functions of a variety of economic and demographic variables.7/ Second, we use these results to derive predicted values and standard errors of the transition probabilities for representative persons with specified characteristics. Finally, we present estimates of the predicted duration on and off welfare (each time a transition is made) for the same individuals.

The logistic regressions were run with the variables in continuous and categorical form. We report in Table 2 the results in the categorical form, as they are more revealing.8/

The estimated coefficients all have the expected sign. The most striking, and most important, result is that persons facing an expected market wage rate below the minimum wage are much less likely to leave welfare and much more likely to return to welfare than persons with a wage larger than the minimum wage. When entered continuously, the wage variable worked in the same directions but not as decisively. Dividing up the above the minimum wage group into narrower groups failed to produce any evidence of differential wage effects above the minimum wage. This result, of course, is consistent with a basic proposition of economic theory: the minimum wage tends to preclude workers with a value of marginal product less than the minimum from employment in the market (see the classic discussion by Stigler [1946] and the excellent analysis by Welch [1973]).9/,10/

Individuals with an expected duration of unemployment of less than two weeks tend to leave welfare more readily than those with

Table 2

Transition Probability Regressions[a]

Variable	(leave welfare) θ	(return to welfare) γ
Constant	-2.819 (0.082)	-3.496 (0.077)
Non-white	-0.790 (0.123)	0.515 (0.135)
Wage $\leq$ minimum wage	-0.500 (0.084)	0.366 (0.099)
Expected unemployment $\leq$ 2 wk/yr	0.237 (0.094)	--
Non-wage income > mean	0.232 (0.138)	--

Dummy variables for age less than twenty-three, unskilled occupation, presence of a pre-school child, and adverse health affecting employability were included in preliminary regressions. The estimated coefficients of these variables were all small and were not measured precisely enough to be considered different from zero using the usual test. The likelihood ratio test that all the coefficients of these omitted variables are zero yielded a χ^2 statistic of 16.0, which is less than the 5% critical value of $\chi^2(10) = 18.3$; hence we accept the hypothesis.

[a]Standard errors in parentheses are from the estimated information matrix. Convergence was achieved in seven iterations.

expected unemployment greater than two weeks per year. While there are several factors influencing the range over which unemployment insurance and welfare complement, or substitute for, each other, these results are consistent with the view that, *ceteris paribus*, a lower search cost for employment tends to induce persons out of welfare into the market.

Persons with a non-wage income greater than the mean tend to have a higher probability of leaving than those below the mean.11/ To the extent that job search requires some financing out of the potential worker's own resources, this result is as expected. There is, of course, the counter effect of leisure--or work in the home--being a normal good with a positive income effect. The net result we find supports (very weakly) the view that the former effect dominates the latter. We find no evidence of a non-wage income effect on the probability of re-entering welfare.

We note also that nonwhites have a lower probability of leaving welfare and a higher probability of returning than whites, *ceteris paribus*.12/

We also note that the criteria upon which welfare recipients are exempt from registering for the recently-enacted work incentive program (WIN)--primarily adverse health or a pre-school age dependent--do not appear to have any noticeable effect on the probability of leaving or re-entering welfare. These results, which are briefly summarized in the footnote to Table 2, are available in detail from the authors.

Table 3 presents point estimates and estimated standard errors of the probability of moving on or off welfare in any given month, and the steady-state probability of being on welfare for individuals with different characteristics.[13] The estimates are quite precise. The estimated monthly probabilities of leaving welfare range from about two to nine percent. The estimated monthly probabilities of returning range from about three to seven percent. The estimated steady-state probabilities of being on welfare in any period range from about twenty-five to eighty-one percent.[14] The range of all three estimates is thus quite substantial.

As expected, given the results reported in Table 2, nonwhites facing a wage less than the minimum wage, an expected duration of unemployment greater than two weeks, with a nonwage income less than the mean, have the lowest estimated probability of leaving, and the highest estimated probability of returning to, welfare. These results combine to predict a steady-state probability of being on welfare of over eighty percent. At the other extreme, whites with wages above the minimum, non-wage income above the mean and expected unemployment below two weeks have the highest estimated probability of leaving welfare, 8.7%, and the lowest estimated probability of returning, 2.9%; these data combine to predict a steady-state probability of being on welfare of 25.2%. The other groups have estimated probabilities between these two extremes. We see that the probability of leaving welfare increases, the probability of returning falls, and the steady-state probability of being on welfare

Table 3

Monthly Transition Probabilities[a]

Group	Probability of Leaving Welfare (α) (in percent)	Probability of Entering Welfare (β) (in percent)	Approximate Probability of Being on Welfare at Time, T, for Large T[b] (in percent)
Wage ≤ $1.60			
White			
Expected Unemployment ≤ 2 weeks/year			
Non-Wage Income ≤ Mean	4.39 (.257)	4.19 (.278)	48.84 (2.21)
Non-Wage Income > Mean	5.47 (.718)	4.19 (.278)	43.36 (3.61)
Expected Unemployment > 2 weeks/year			
Non-Wage Income ≤ Mean	3.49 (.331)	4.19 (.278)	54.51 (2.87)
Non-Wage Income > Mean	4.37 (.664)	4.19 (.278)	48.95 (4.14)
Nonwhite			
Expected Unemployment ≤ 2 weeks/year			
Non-Wage Income ≤ Mean	2.04 (.263)	6.82 (.876)	76.98 (3.22)
Non-Wage Income > Mean	2.56 (.460)	6.82 (.876)	72.71 (4.38)
Expected Unemployment > 2 weeks/year			
Non-Wage Income ≤ Mean	1.61 (.198)	6.82 (.876)	80.84 (2.75)
Non-Wage Income > Mean	2.03 (.357)	6.82 (.876)	77.06 (3.85)

Table 3 (continued)

Group	Probability of Leaving Welfare (α) (in percent)	Probability of Entering Welfare (β) (in percent)	Approximate Probability of Being on Welfare at Time T, for Large T (in percent)
Wage > $1.60			
White			
Expected Unemployment ≤ 2 weeks/year			
Non-Wage Income ≤ Mean	7.03 (.480)	2.94 (.691)	29.50 (5.09)
Non-Wage Income > Mean	8.71 (1.135)	2.94 (.691)	25.25 (5.07)
Expected Unemployment > 2 weeks/year			
Non-Wage Income ≤ Mean	5.63 (.448)	2.94 (.691)	34.31 (5.59)
Non-Wage Income > Mean	7.00 (.969)	2.94 (.691)	29.58 (5.68)
Nonwhite			
Expected Unemployment ≤ 2 weeks/year			
Non-Wage Income ≤ Mean	3.32 (.433)	4.83 (.588)	59.28 (4.31)
Non-Wage Income > Mean	4.15 (.737)	4.83 (.588)	53.79 (5.35)
Expected Unemployment > 2 weeks/year			
Non-Wage Income ≤ Mean	2.64 (.286)	4.83 (.588)	64.68 (3.73)
Non-Wage Income > Mean	3.31 (.540)	4.83 (.588)	59.41 (4.92)

[a]Standard errors in parentheses are for the estimated expected values and are calculated using the asymptotic distribution theory developed in the Appendix.

[b]One minus these entries yield point estimates of the approximate probability of being off welfare at T, for large T.

falls as we move from nonwhites to whites, below the minimum wage to above, more than two weeks of expected unemployment to less, and less than the mean nonwage income to more.

Table 4 presents estimates of the expected duration of stay on and off welfare (each time a transition to that state is made) for the same groups of individuals. Since expected duration in a state is given by the reciprocal of the probability of leaving that state, the results follow closely those reported above. The estimated expected duration on welfare (each time on welfare) ranges from about eleven months to over sixty months. the estimated expected duration off welfare (each time a person comes off) ranges from about fifteen months to about thirty-four months. As we by now expect, expected duration on welfare is higher for nonwhites than whites, for persons facing wages below the minimum than above, for persons facing expected unemployment greater than two weeks per year than for less, and for persons with non-wage income less than the mean compared to those with more. The expected duration off welfare, once off, is larger for whites than nonwhites, and for those with wages above the minimum than below.

5. Conclusion

We have analyzed a new body of longitudinal data on welfare recipients in order to gain some insight into turnover in the welfare population. In doing so, we developed a statistical model of transition among discrete states which is potentially applicable to many economic

Table 4

Expected Duration in Months of Stay on Welfare[a]

	Expected Duration On Welfare Per Time On Welfare ($1/\alpha$)	Expected Duration Off Welfare Per Time Off Welfare ($1/\beta$)
Wage ≤ $1.60		
White		
Expected Unemployment ≤ 2 weeks/year		
Non-Wage Income ≤ Mean	22.80 (1.335)	23.88 (1.586)
Non-Wage Income > Mean	18.28 (2.398)	23.88 (1.586)
Expected Unemployment > 2 weeks/year		
Non-Wage Income ≤ Mean	28.62 (2.712)	23.88 (1.586)
Non-Wage Income > Mean	22.89 (3.475)	23.88 (1.586)
Nonwhite		
Expected Unemployment ≤ 2 weeks/year		
Non-Wage Income ≤ Mean	49.04 (6.323)	14.67 (1.883)
Non-Wage Income > Mean	39.07 (7.004)	14.67 (1.883)
Expected unemployment > 2 weeks/year		
Non-Wage Income ≤ Mean	61.86 (7.618)	14.67 (1.883)
Non-Wage Income > Mean	49.23 (8.682)	14.67 (1.883)

Table 4 (continued)

	Expected Duration On Welfare Per Time On Welfare ($1/\alpha$)	Expected Duration Off Welfare Per Time Off Welfare ($1/\beta$)
Wage > $1.60		
White		
Expected Unemployment ≤ 2 weeks/year		
Non-Wage Income ≤ Mean	14.22 (0.971)	33.98 (7.995)
Non-Wage Income > Mean	11.48 (1.496)	33.98 (7.995)
Expected Unemployment > 2 weeks/year		
Non-Wage Income ≤ Mean	17.75 (1.412)	33.98 (7.995)
Non-Wage Income > Mean	14.28 (1.977)	33.98 (7.995)
Expected Unemployment ≤ 2 weeks/year		
Non-Wage Income ≤ Mean	30.14 (3.936)	20.70 (2.526)
Non-Wage Income > Mean	24.10 (4.280)	20.70 (2.526)
Expected Unemployment > 2 weeks/year		
Non-Wage Income ≤ Mean	37.92 (4.105)	20.70 (2.526)
Non-Wage Income > Mean	30.22 (4.939)	20.70 (2.526)

[a]Standard errors in parentheses are for the estimated expected values and are calculated using the asymptotic distribution theory developed in Section 3.

phenomena.[15/] We have discovered that the popular notion of the welfare population as more or less permanently entrenched on welfare is erroneous. An enormous amount of turnover occurs in the welfare population and the average duration of time on welfare once on welfare is relatively modest.[16/]

In analyzing the probabilities ~~of leaving~~ leaving, and returning to, welfare our main substantive conclusion is that persons facing a wage below the minimum wage are much less likely to leave welfare, much more likely to return, stay off welfare for shorter periods, stay on welfare for longer periods, and are much more likely to be on welfare in the steady state than those facing wages above the minimum wage.

We have also found that non-wage income, expected unemployment and race affect welfare dependency status in a manner consistent with economic theory and anectdotal evidence.

APPENDIX

Maximum Likelihood Estimators of Transition Probabilities and Related Statistics

For each individual s, we observe T_{ij}^s, the number of times individual s made a transition from state i to state j, i,j = 1,2.

We now develop the maximum likelihood estimates of Θ and Γ.[17/] The probability associated with any given individual's set of T_{ij}^s is

$$P(T_{11}^s,T_{12}^s,T_{21}^s,T_{22}^s|\Theta,\Gamma) = K(T_{11}^s,T_{12}^s,T_{21}^s,T_{22}^s)\cdot[1 - f(\Theta'X_s)]^{T_{11}^s}\cdot$$

$$\cdot[f(\Theta'X_s)]^{T_{12}^s}[1 - f(\Gamma'Z_s)]^{T_{22}^s}\cdot[f(\Gamma'Z_s)]^{T_{21}^s} ,$$

where $K(T_{11}^s,T_{12}^s,T_{21}^s,T_{22}^s)$ is a function which does not depend on Θ or Γ.[18/] We assume the individuals move independently so that the probability of the whole set of observations on S individuals is proportional to

$$\prod_{s=1}^{S}\left\{[1-f(\Theta'X_s)]^{T_{11}^s}[f(\Theta'X_s)]^{T_{12}^s}[1-f(\Gamma'Z_s)]^{T_{22}^s}[f(\Gamma'Z_s)]^{T_{21}^s}\right\} .$$

Hence the log of the likelihood function is

$$\ln \mathcal{L}(\Theta,\Gamma|T_{11}^s,T_{12}^s,T_{21}^s,T_{22}^s,\ s = 1,\ldots,S) = k + \sum_{s=1}^{S}\{T_{11}^s\ln\ [1-f(\Theta'X_s)]+$$

$$+T_{12}^s\ln\ [f(\Theta'X_s)]+T_{22}^s\ln\ [1-f(\Gamma'Z_s)]+T_{21}^s\ln\ [f(\Gamma'Z_s)]\} .$$

The first partial derivatives of $\ln \mathcal{L}$ are:

$$\frac{\partial \ln \mathcal{L}}{\partial \theta_i} = \sum_{s=1}^{S} [T_{12}^s x_{is} - (T_{11}^s + T_{12}^s) x_{is} f(\Theta' X_s)] \quad , \quad i = 1,\ldots,K$$

$$\frac{\partial \ln \mathcal{L}}{\partial \gamma_i} = \sum_{s=1}^{S} [T_{21}^s z_{js} - (T_{22}^s + T_{21}^s) z_{js} f(\Gamma' Z_s)] \quad , \quad j = 1,\ldots,P \ .$$

The maximum likelihood estimators, denoted $\hat{\Theta}' = [\hat{\theta}_1,\ldots,\hat{\theta}_K]$ and $\hat{\Gamma}' = [\hat{\gamma}_1,\ldots,\hat{\gamma}_P]$, are obtained by equating these first partial derivatives to zero and solving for the θ_i's and γ_j's. Because these equations are nonlinear in the parameters, we chose the Newton-Raphson iterative method of solution. This method uses the matrix of second partial dervatives of $\ln \mathcal{L}$ which we denote $[m_{ij}]$. The $(K+P)^2$ elements of $[m_{ij}]$ are defined as:

$$m_{ij} = \frac{\partial^2 \ln \mathcal{L}}{\partial \theta_i \partial \theta_j} = \sum_{s=1}^{S} \{-(T_{11}^s + T_{12}^s) x_{is} x_{js} f(\Theta' X_s)[1 - f(\Theta' X_s)]\} \ ,$$

$$i,j = 1,\ldots,K \ ,$$

$$m_{i,j+K} = m_{i+K,j} = \frac{\partial^2 \ln \mathcal{L}}{\partial \theta_i \partial \gamma_j} = 0 \quad , \quad i = 1,\ldots,K \quad , \quad j = 1,\ldots,P \ ,$$

$$m_{i+K,j+K} = \frac{\partial^2 \ln \mathcal{L}}{\partial \gamma_i \partial \gamma_j} = \sum_{s=1}^{S} \{-(T_{22}^s + T_{21}^s) z_{is} z_{js} f(\Gamma' Z_s)[1 - f(\Gamma' Z_s)]\} \ ,$$

$$i,j = 1,\ldots,P \ .$$

Note that $[m_{ij}]$ is block diagonal. Hence, we can estimate Θ and Γ separately, simplifying the estimation procedure and saving computation time.

Distributions of the Estimators and Related Statistics

The log of the likelihood function is a well-behaved function of the unknown parameters, and to show that the conditions are met for the maximum likelihood estimators derived above to be B.A.N. (Best Asymptotically Normal) is relatively easy.[19/] Hence, letting

$$\psi = \begin{bmatrix} \Theta \\ \Gamma \end{bmatrix} ,$$

$$\hat{\psi} = \begin{bmatrix} \hat{\Theta} \\ \hat{\Gamma} \end{bmatrix} ,$$

$$I(\psi) = -\frac{1}{S}\mathcal{E}[m_{ij}]|_{\psi} ,$$

we have

$$\sqrt{S}(\hat{\psi}-\psi) \overset{d}{\to} N(O_1, I^{-1}(\psi)) ,$$

where $N(\ ,\ )$ denotes the multivariate normal distribution and O_1 is a $(K+P)\times 1$ vector of zeros. Note that we assume $S \to \infty$ for the asymptotic distribution theory. Evaluation of $I(\psi)$, which is usually called the Fisher information matrix, requires evaluation of $\mathcal{E}T_{ij}^{s}$; $i,j = 1,2$. As in Section 2 we must take account of the initial conditions as well as how the data were collected. Using techniques similar to those employed in Section 2, we have for $s = 1,\ldots,S$:

$$\mathcal{E}(T_{11}^{s}) = \frac{(1-\alpha_s)\beta_s}{\alpha_s+\beta_s}(T_s-1) + \frac{\alpha_s(1-\alpha_s)}{(\alpha_s+\beta_s)^2}(1-\xi_s^{T_s}) ,$$

$$\mathcal{E}(T_{22}^s) = \frac{-(1-\beta_s)\alpha_s}{\alpha_s+\beta_s}(T_s-1) + \frac{\alpha_s(\beta_s-1)}{(\alpha_s+\beta_s)^2}(1-\xi_s^{Ts}) ,$$

$$\mathcal{E}(T_{12}^s) = \frac{\alpha_s\beta_s}{\alpha_s+\beta_s}(T_s-1) + \frac{\alpha_s^2}{(\alpha_s+\beta_s)^2}(1-\xi_s^{Ts}) ,$$

$$\mathcal{E}(T_{21}^s) = 1 + \frac{\alpha_s\beta_s}{\alpha_s+\beta_s}(T_s-1) - \frac{\alpha_s\beta_s}{(\alpha_s+\beta_s)^2}(1-\xi_s^{Ts}) ,$$

where $\xi_s = 1-\alpha_s-\beta_s$. To obtain the estimate of the variance coavariance matrix of $\hat{\psi}$ we need only calculate $\frac{1}{S}\, I^{-1}(\psi)|_{\psi=\hat{\psi}}$. The elements of this matrix will be close to the elements of $-[m_{ij}]^{-1}\Big|_{\psi=\hat{\psi}}$, which is generated at the last iteration of the Newton-Raphson method.

We are also interested in the distribution of functions of the elements of $\hat{\psi}$. We have indicated above that $\hat{\psi}$ is a B.A.N. estimator of ψ. If $g(\psi)$ is a function whose first partial derivatives with respect to the elements of ψ exist, then

$$\sqrt{S}(g(\hat{\psi}) - g(\psi)) \xrightarrow{d} N(0_1, D_g'(\psi) I^{-1}(\psi) D_g(\psi))$$

where $D_g' = [\partial g(\psi)/\partial\theta_1, \ldots \partial g(\psi)/\partial\theta_K, \partial g(\psi)/\partial\gamma_1, \ldots, \partial g(\psi)/\partial\gamma_P]$. Thus, $g(\hat{\psi})$ is a B.A.N. estimator of $g(\psi)$.[20/]

We are specifically interested in six functions of the elements of ψ: α_s, α_s^{-1}, β_s, β_s^{-1}, $\beta_s(\alpha_s+\beta_s)^{-1}$, and $\alpha_s(\alpha_s+\beta_s)^{-1}$. We derive estimates of these functions by substituting the elements of $\hat{\psi}$ for their unknown counterparts. The result above applies, and we obtain the following asymptotic distributions, $s = 1,\ldots,S$:

$$\sqrt{S}(\hat{\alpha}_s-\alpha_s) \xrightarrow{d} N\left(0,\alpha_s^2(1-\alpha_s)^2[X_s'O_2']I^{-1}(\psi)\begin{bmatrix}X_s\\O_2\end{bmatrix}\right),$$

$$\sqrt{S}(\hat{\alpha}_s^{-1}-\alpha_s^{-1}) \xrightarrow{d} N\left(0,\exp(-2\theta'X_s)[X_s'O_2']I^{-1}(\psi)\begin{bmatrix}X_s\\O_2\end{bmatrix}\right),$$

$$\sqrt{S}\left(\frac{\hat{\beta}_s}{\hat{\alpha}_s+\hat{\beta}_s}-\frac{\beta_s}{\alpha_s+\beta_s}\right) \xrightarrow{d} N\left(0,[a(\theta,\Gamma)X_s'\ \ b(\theta,\Gamma)Z_s']I^{-1}(\psi)\begin{bmatrix}a(\theta,\Gamma)X_s\\b(\theta,\Gamma)Z_s\end{bmatrix}\right),$$

where O_2 is a $P\times 1$ vector of zeros,

$$a(\theta,\Gamma) = \frac{-[1+\exp(\Gamma'Z_s)]\exp(\theta'X_s+\Gamma'Z_s)}{[\exp(\theta'X_s)+\exp(\Gamma'Z_s)+2\exp(\theta'X_s+\Gamma'Z_s)]^2},$$

$$b(\theta,\Gamma) = \frac{[1+\exp(\theta'X_s)]\exp(\theta'X_s+\Gamma'Z_s)}{[\exp(\theta'X_s)+\exp(\Gamma'Z_s)+2\exp(\theta'X_s+\Gamma'Z_s)]^2},$$

Results parallel to those given above hold for $\hat{\beta}_s$, $\hat{\beta}_s^{-1}$ and $\hat{\alpha}_s(\hat{\alpha}_s+\hat{\beta}_s)^{-1}$, *mutatis mutandi*. Numerical values for the variances of the estimated functions are obtained by evaluating the theoretical asymptotic variances at $\hat{\psi}$, the maximum likelihood estimates. Again note that the limit is taken on S.

Footnotes

1/ The basic model for studying the allocation of time to different activities is by now so familiar as to render it repetitious to present it here. Utility maximization subject to a budget constraint on full income (and perhaps some production relationships) produces demand functions for the use of time in different activities which depend upon wage rates and nonwage income. These variables will play an important role in the empirical anlaysis presented below.

2/ We are forced to adopt this convention because of the method of data collection (see the discussion in Section 3). In principle, of course, we would prefer to adopt a more general model and test the assumption of stationarity.

3/ This first-order assumption might also be relaxed with a richer body of data.

4/ An alternative model which divides the sample into two groups--those who would never make a transition regardless of the stimulus (stayers) and those whose transition probabilities follow a first-order stationary Markov process--has been suggested by Blumen et. al. [1955], developed statistically by Goodman [1961] and applied to the interesting case of transition in and out of poverty status by McCall [1971]. It is worthwhile both to examine our data, and to point out a potential problem with the interpretation of our results, in light of the mover-stayer approach.

First, we have divided our sample into groups based on the eight dichotomous variables reported in Table 2 and followed the procedure suggested by Goodman [1961] to test the hypothesis that the proportion of stayers in each group is zero. We rejected this hypothesis for eighteen percent of the groups; a strict interpretation of the mover-stayer model would suggest that we had about eight percent stayers in our sample. However, the test is an asymptotic one, and the length of time involved (five years) is short relative to the time spanned by many economic decisions (fertility and marriage, divorce and remarriage). The eight percent can only be interpreted as an upper bound on the proportion of stayers. For example, we note that roughly ten percent of those individuals on welfare for the first fifty-five months of the sample period left on or before the sixtieth month. It would be surprising if this exodus from AFDC would suddenly cease after month sixty. Thus, the eight percent of our sample suspected of being stayers probably contain a substantial number of movers. Further, the suspected stayers are spread rather evenly across the sixteen groups reported in Tables

3 and 4. Hence, the estimated transition probabilities look very similar for the whole sample and for the sample purged of suspected stayers. All of these results are available from the authors upon request.

Second, if indeed there is a proportion of the population who are truly stayers in welfare status, our sampling procedure--examining data for families which came onto welfare in 1965--may systematically exclude such stayers, since they never have a chance to come on in 1965 if they have been on welfare continuously since some previous date. We are thus sampling from a population depleted of stayers. Since at least seventy percent of the current California welfare population entered (or re-entered) welfare from 1965 to the present, our model at worst would be inappropriate for whatever proportion of the, at most, thirty percent of the current welfare population which was already on welfare in 1964 are truly stayers. This objection should be noted in interpreting our results. We may only be describing turnover in the new additions (since 1965) to welfare--although these do make up the overwhelming majority of welfare recipients.

While our data can shed no direct light on the pre-1965 group, one indirect inference is possible. While the mover-stayer model would predict depletion of the pool of potential stayers by 1965, each newly eligible age cohort would contain some potential stayers. Our results, however, suggest that the estimated proportion of suspected stayers and the transition probabilities are very similar for our sample of welfare recipients above and below twenty-three years of age.

The net effect of the above reasoning is to lead us to consider the mover-stayer model unnecessary for our sample, but to interpret our results cautiously as perhaps applying only to the seventy (or more) percent of welfare recipients entering welfare since 1965.

5/ The approach of making the transition probabilities a function of personal characteristics has been used successfully in several previous studies of labor market phenomena. See Hall [1973] and McCall [1971] for two interesting applications. Also see the classic study by Orcutt, et. al. [1961].

6/ The original data contained 653 observations. We eliminated the observations for families with missing information and also for the few families eligible under the unemployed father program. Hence, our data refer to female headed households (the overwhelming majority of the welfare caseload) only.

7/ Since our sample includes only persons who have been on welfare for at least one month, our results must be interpreted as conditional upon that fact. We do not estimate the probability that someone who has never been on welfare will eventually come on aid. Further, recall the proviso of footnote 4.

8/ The continuous results tended to look very much like the categorical results. Since the latter impose less a priori structure and allow the data to play a greater role in determining the results, we eschew any attempt to report the former. They are available upon request from the authors.

9/ The most likely explanation for leaving welfare, other than obtaining employment, is marriage, remarriage or other alliance with an employed male. Our results are consistent with the notion that low-wage women are less valuable marriage partners than high-wage women and hence have a lower probability of forming such an alliance.

10/ Since there was a general increase in wage rates throughout this period, a few workers with wages below the minimum in 1965 may have been above the minimum by 1970. Further the coverage of the minimum wage, and its level, were extended slightly during the period. Average welfare payments and consumer prices also increased. The net effect of these considerations is probably to lead us to underestimate the wage effect on transition probabilities. However, examination of the time pattern of leaving welfare (for those who left and never came back--the only group about whom we can infer this information) and of the group with wages between the old and new minimum provides no evidence that this effect is quantitatively important for our sample.

11/ This variable passes the one-tailed test at the 5% level marginally.

12/ Several interactions of the race and other variables were tried, but none produced any evidence of interation effects.

13/ Recall that we are invoking asymptotic distribution theory.

14/ Recall the proviso of footnotes 4 and 7.

15/ When a first-order discrete time stationary Markov process is a reasonable assumption. See footnotes 2 and 3.

16/ Recall the proviso about our sample noted in footnotes 4 and 7.

17/ The vectors Θ and Γ can also be estimated in other ways. The best alternative to the maximum likelihood method is Berkson's minimum logit chi-squared method. (See Chapter 3 of Cox [1970].)

We decided to employ the maximum likelihood method despite higher computation costs and the asymptotic equivalence of the estimators because the large number of extreme observations in our sample would force adoption of rather ad hoc correctives were we to use Berkson's method.

18/ See P. Billingsley [1961, p. 13] or Anderson and Goodman [1957, p. 91].

19/ See S. Zacks [1971, p. 247] for a concise statement of the conditions under which the maximum likelihood estimator is also B.A.N.

20/ See S. Zacks [1971, p. 249] for a heuristic development of this result.

References

Amemiya, T. and M. Boskin [1974], "Regression Analysis When the Dependent Variable is Truncated Lognormal, with an Application to the Determinants of the Duration of Welfare Dependency," International Economic Review (forthcoming, June).

Anderson, T. W. and L. A. Goodman [1957], "Statistical Inference About Markov Chains," Annals of Mathematical Statistics, Vol. 28, pp. 89-110.

Becker, G. [1965], "A Theory of the Allocation of Time," Economic Journal, Vol. 75 (September), pp. 493-517.

Billingsley, P. [1961], "Statistical Methods in Markov Chains," Annals of Mathematical Statistics, Vol. 32, pp. 12-40.

Blumen, I., M. Kogan and P. McCarthy [1955], The Industrial Mobility of Labor as a Probability Process, Cornell University.

Boskin, M. [1974], "A Conditional Logit Model of Occupational Choice," Journal of Political Economy, (forthcoming, March).

Cox, D. R. [1970], Analysis of Binary Data, Methuen & Co. Ltd., London.

Goodman, L. A. [1961], "Statistical Methods for the Mover-Stayer Model," Journal of the American Statistical Association, pp. 841-868.

Hall, R. [1973], "Turnover in the Labor Force," Brookings Papers on Economic Activity, No. 3, pp. 709-764.

Harberger, A. C. [1971a], "On Measuring the Social Opportunity Cost of Labor," International Labour Review (June), pp. 559-579.

Harberger, A. C. [1971b], "Three Basic Postulates for Applied Welfare Economics: An Interpretative Essay," Journal of Economic Literature Vol. IX (September), pp. 785-797.

Kendall, M. G. and A. Stuart [1961], The Advanced Theory of Statistics, Vol. 2, Hafner Publishing Co., New York.

McCall, J. [1971], "A Markovian Model of Income Dynamics," Journal of the American Statistical Association, pp. 439-447.

Orcutt, G. H., M. Greenberger, J. Korbel and A. Rivlin [1961], Microanalysis of Socioeconomic Systems: A Simulation Study, Harper & Brothers.

Phelps, E. S., et. al. [1970], Microeconomic Foundations of Employment and Inflation Theory, W. W. Norton & Co., New York.

Stigler, G. [1946], "The Economics of Minimum Wage Legislation," American Economic Review, Vol. 36 (June).

Welch, F. [1973], "Minimum Wage Legislation in the United States," unpublished mimeo.

Zacks, S. [1971], The Theory of Statistical Inference, John Wiley & Sons Inc.