Institute for Research on Poverty
Discussion Paper no. 1046-94

A Hazard Model for Welfare Durations with Unobserved Location-Specific Effects

John Fitzgerald
Department of Economics
Bowdoin College

October 1994

This research was undertaken while John Fitzgerald was an American Statistical Association/ NSF/Census Fellow at the U.S. Bureau of the Census. The program is supported by NSF grant SES 87-13643 and the Census Bureau. All opinions and conclusions are the author's and do not reflect the views of the NSF, the Census Bureau, or the Institute for Research on Poverty.

Abstract

Many papers have investigated how personal characteristics and environmental variables affect welfare durations of unmarried mothers. This paper estimates proportional hazard models for welfare durations that allow for either fixed state or fixed labor market area effects. Conditioning on residence location by fixed effects can limit the impact of three types of potential bias. (1) Estimates of the effects of personal characteristics can be biased owing to the omission of relevant local area variables. (2) Estimates of the impact of state welfare benefit levels are biased because they proxy for other unmeasured attributes of the state, in particular, the entire state welfare system. Conditioning on state fixed effects limits this bias to the extent that we can use time variation within states to estimate the benefit level effect. (3) With state fixed effects, we can better estimate the impact of local conditions, such as unemployment rates, because they also may have been picking up omitted state-level effects. The models are estimated by the Cox partial likelihood method with time-varying covariates. Data come from the 1984 and 1985 panels of the Survey of Income and Program Participation. I find that some personal characteristics (being black or Hispanic, education) have greater impact after controlling for location-specific effects.

A Hazard Model for Welfare Durations
with Unobserved Location-Specific Effects

I. INTRODUCTION AND PROBLEM

Models of welfare dynamics have generally estimated the effects on duration of welfare receipt of personal characteristics of the recipient, state AFDC benefit levels, and some environmental covariates such as unemployment rates. This paper estimates a proportional hazard model for welfare durations that allows for fixed state or local labor market effects in order to provide more complete control for state policy and environmental influences.

Conceptual models for welfare durations usually assume that a woman on AFDC chooses between the options of staying on or getting off welfare. In these discrete choice models, a woman chooses the current option that provides a larger present value of her expected future utility. The non-welfare option is associated with increasing earnings (getting a job, increasing current work hours) or marrying. The expected returns on these options can vary through time, producing a sequence of decisions which result in welfare spells. For example, see Blank (1989).

Within this framework, the exit from AFDC would depend on the relative value of the welfare option compared to job or marriage options. This in turn depends on (a) personal characteristics such as the mother's age and education (which affect her wage and job options), number and age of her children (which affect the value of home production and cost of child care), the availability of other income when off welfare (property income or child support); (b) policy parameters such as the benefit level and other state welfare program characteristics; and (c) environmental variables that reflect the job market, marriage market, and so on. Most models have depended upon the state AFDC benefit maximum to proxy the state welfare system. Most also depend on state-level measures of unemployment rates and environmental variables, with some using local-area measures.[1]

Given the limited number of state or local area variables that can be included in such models, omitted variables are likely. A key example would be administrative practices of the state of residence: these important practices may not be captured by the inclusion of state benefit levels alone. Another example would be community attitudes toward welfare. This paper estimates a model of welfare duration that allows for unobserved fixed state effects, and a separate model that allows for unobserved fixed local effects of a labor market area (described below). Such models have the potential to remove biases due to unobserved, time-invariant characteristics of the states or localities. Conditioning on location via fixed effects can remove three potential biases: (1) the bias on personal characteristics such as education level and race due to omission of local or state traits, (2) bias on the effects of included state-level variables such as AFDC benefit levels due to omission of other state characteristics, and (3) bias on local area variables such as unemployment rates due to omitted state effects.

I compare three models: first, a proportional hazard model without fixed effects that includes some local area covariates; second, the same model with fixed state effects; and third, a model with fixed local effects of a labor market area. The models are estimated using the Cox partial likelihood method and allow for time-varying covariates. The next section describes this model. Section III describes the data taken from the 1984 and 1985 panels of the Survey of Income and Program Participation (SIPP). Section IV discusses the results, and Section V shows fit and sensitivity tests. A brief conclusion follows.

Throughout the paper, the term "labor market area," or LMA, refers to an aggregate of counties. The definition is based on work by Tolbert and Killian (1987), who divide the United States into 382 labor market areas based on the relative strength of commuting ties among counties. Compared to some other definitions, these LMA's have the advantage that they exhaust all counties in the United States and can cross state boundaries. The LMA's tend to look like the more familiar SMSA's in urban areas, but rural counties are also grouped.

II. THE FIXED EFFECT PARTIAL LIKELIHOOD MODEL

I specify a proportional hazard model where the baseline hazard includes a common fixed effect for persons in the same location. In this section I refer to a "location-specific effect," which could be either state-specific or specific to a labor market area. The model is estimated by Cox's partial likelihood method.

Both Chamberlain (1985) and Kalbfleish and Prentice (1980) discuss the possibility of using partial likelihood to eliminate fixed effects as nuisance parameters. The development in this paper closely follows that of Ridder and Tunali (1989, 1990), who estimate by partial likelihood a mortality hazard for children with a fixed family-specific effect. Their model allows time-varying covariates and formally develops the conditions under which such an approach is appropriate.

Let t_{ij} denote the (uncensored) spell length by the ith woman in location j. Define the hazard as

(1) $$\Psi\left(t_{ij} \middle| X_{ij}(t_{ij}), v_j\right) = h\left(X_{ij}(t_{ij})\right)\lambda_j(t_{ij}, v_j)$$

where

(2) $h(X_{ij}(t)) = \exp(B'X_{ij}(t_{ij}))$ and X denotes the matrix of potentially time varying covariates; B is a vector of unknown coefficients; $\lambda_j(t_{ij},v_j)$ is the baseline hazard in location j, which is allowed to depend on the unobserved fixed effect v_j.

Ridder and Tunali state four conditions that ensure that the partial likelihood model with time-varying covariates is appropriate. These are met in my context where we have independent spells by women grouped by location. Condition 1 is independent failure rates—there is no interaction between exit rates by women in the same location, conditional on covariates. Condition 2 states that the covariates are exogenous. I tried to select largely exogenous covariates (age, race, ethnicity, education, other income, age and number of children), although it might be argued that children and education are choice variables.[2] Condition 3 states that the censoring process is independent of the welfare exit process. This clearly holds for spells censored by the end of the panel, but is more questionable for sample attriters. Fitzgerald and Zuo (1991) suggest that attrition may not be a problem for models of welfare spells in SIPP. Condition 4 is that the (conditional) exit rates for a location do not depend on the number of women (spells) in that location. The last assumption is necessary to avoid bias because the estimation is restricted to locations that include two or more spells.

The intuition behind likelihood construction for the partial likelihood model is straightforward. Each person (spell) contributes a piece to the overall likelihood that answers the following question: Given that an exit occurred at time S, what is the probability that the exit occurs by the actual case with length S rather than any of the other cases that are still at risk at time S?

In notation, let $\mathbb{R}_{ij}$ be the risk set for person (spell) *i* in location *j*. This is the set of all persons who live in *j* who have not exited prior to t_{ij} (including censored cases). The contribution of an individual to the likelihood is

(3)
$$L_i(\boldsymbol{B}) = \frac{\Psi(t_{\boldsymbol{ij}} | \cdot)}{\sum_{k \in \mathbb{R}_{\boldsymbol{ij}}} \Psi(t_{\boldsymbol{ij}} | \cdot)}$$

(4)

$$= \frac{h\left(X_{ij}\left(t_{ij}\right)\right)}{\sum\limits_{k \in \mathbb{R}_{ij}} h\left(X_{kj}\left(t_{ij}\right)\right)}$$

(5)

$$= \frac{\exp\left(\boldsymbol{B}' X_{ij}\left(t_{ij}\right)\right)}{\sum\limits_{k \in \mathbb{R}_{ij}} \exp\left(\boldsymbol{B}' X_{kj}\left(t_{ij}\right)\right)} .$$

Everything is measured at t_{ij}, the spell length for person *i*. Note that the underlying hazard λ_j has canceled out, along with its implicit fixed effect v_j. In essence, the risk set for a person includes only those in her location, and the estimation makes comparisons only among those who live in location *j*. The likelihood for the whole sample is

(6)

$$L = \prod_{i=1}^{N} L_i(\boldsymbol{B})$$

where $L_i(B)$ is from (3).

I maximize the log likelihood:[3]

(7) $$\mathfrak{L} = \sum_{i=1}^{N} \ell n\, L_i(\boldsymbol{B}) = \sum_{i=1}^{N} \left\{ \boldsymbol{B}' X_{ij}(t_{ij}) - \ell n \sum_{k \in \mathbb{R}_{ij}} \exp\left(\boldsymbol{B}' X_{ij}(t_{ij})\right) \right\}.$$

A word should be said about the use of time-varying covariates. I assumed that the proper way to line up the covariates across persons was by elapsed time in spell. Thus for the likelihood contribution by person *i* with spell length t_{ij}, the risk set uses covariates from elapsed time in spell t_{ij} for all persons in the risk set. For example, if person *i*'s spell has length 3 then the risk set uses covariates for the third month of each spell.

III. DATA AND VARIABLES

A. SIPP

My data come from the Survey of Income and Program Participation (SIPP), a longitudinal sample of households representing the noninstitutionalized population of the United States. It includes monthly information on income, use of government programs, labor force participation, and demographic characteristics. Interviews are conducted every four months asking about activity in the previous four months. Each year a new panel is introduced. Each panel potentially gives 32 months of data collected from eight interviews.[4] I worked with the 1984 and 1985 Longitudinal Research Files (Panels) which have been longitudinally edited for consistency (SIPP, 1989, pp. B-1 to B-19). The 1984 panel includes about 20,000 households and spans June 1983 to March 1986.[5] The 1985 panel includes about 15,000 households and spans October 1984 to July 1987. For more details on SIPP, see Nelson, McMillen, and Kasprzyk (1985).

B. Welfare Recipiency

I began with a subsample of unmarried women with children (female heads of families) who received welfare at any time during a panel. This group was chosen because female heads are of primary policy interest, and because the welfare data on this group may be more reliable.[6] A woman is coded as a recipient if she reports receiving either AFDC or General Assistance. This definition includes women who misreport their AFDC receipt as General Assistance, a known problem (Marquis and Moore, 1989). Based on earlier work and an administrative data check, this definition more accurately identifies the AFDC population of female heads in SIPP than using AFDC receipt alone.[7]

A spell of welfare receipt is defined as the length of time that a woman continuously receives welfare income (AFDC or General Assistance). One month gaps of nonreceipt were ignored to produce a continuous spell over the gap. A spell can occur at any time during a panel. To avoid econometric difficulties in working with left-censored spells, I used only complete and right-censored spells. Thus my spells are those of new entrants, that is, the first observed (complete or right-censored) spell of receipt.

Persons who miss interviews during the panel or refuse to answer specific items may have data imputed to them. All imputed recipiency data are treated as missing in the analysis. Persons who missed interviews were considered censored at that interview.[8]

C. Variables

Table 1 displays the means and definitions of the variables. Most are self explanatory.

Local area unemployment is used to capture the strength of the local labor market, and local area sex ratio is used to crudely proxy marriage prospects. These variables were matched to SIPP individuals using county-of-residence information available on internal Census files.[9] The LMA variables are weighted averages of the counties within the LMA.[10] The unemployment data, LUNEM, are annual rates (1983-87) computed for the LMA based on county-level data from the

Bureau of Labor Statistics. The county data on LMA Sexratio (LSEXRT) came from the 1988 City County Data Book.

A few notes are in order. The urban residence dummy, URBAN, indicates residence in a large SMSA (population greater than 250,000). One expects that welfare use is more common, hence less stigmatized, in the anonymity of larger urban areas. Other income was included to show outside income possibilities such as child support; I included it as a dummy variable because including it as a linear continuous variable always produced a small coefficient with a large standard error. State welfare program information came from the U.S. House of Representatives, Committee on Ways and Means (1987). As is commonly done, I used the maximum benefit level for a family of four, AFDCMAX, as the benefit measure.[11] I control for family size through number of children, NKID. A dummy for having children aged less than 6, YKID, reflects increased value of home time and increased cost of child care if working. The median welfare duration of 11-12 months shown in Table 1 agrees with Ruggles (1989) and Long and Doyle (1989).[12]

TABLE 1

Definition, Mean, and Standard Deviation of Variables for Spells

Variable Name	Definition	Mean	Std Dev
AGE	(at spell beginning)	28.7	9.21
BLACK	(dummy = 1 if black)	0.36	0.48
HISP	(dummy = 1 if Hispanic)	0.12	0.33
EDU	(years of education completed)	10.76	2.42
NKIDS	(total children aged 18 or younger)	1.79	1.05
YKID	(dummy = 1 for children below 6)	0.67	0.47
OTHDUM	(dummy = 1 for positive property income, child support or alimony)	0.13	0.34
URBAN	(dummy =1 if live in SMSA with population more than 250,000)	0.61	0.48
LSEXRT	(LMA male per 100 female)	94.53	3.58
LUNEM	(LMA unemployment rate)	8.07	2.77
AFDCMAX	(AFDC benefit for family of four, in $100's)	4.16	1.48
P84	(dummy = 1 if from 1984 panel data)	0.65	0.47
COMPLETE	(dummy = 1 if complete spell)	0.495	0.50
Median Duration (months)[a]		11-12	
Sample Size		533	

Note: First observed spell from pooled 1984 and 1985 panels of SIPP. Welfare recipiency is either AFDC or General Assistance.

[a]Median Duration from estimated Kaplan-Meier survivor function.

IV. RESULTS

The partial likelihood estimates are shown in Table 2. The first column shows coefficients for the model without fixed effects. The second shows the exponentiated coefficient, which helps interpret the size of the coefficient. In a proportional hazard model, the exponentiated coefficient shows the proportional change in the hazard for a one-unit change in the covariate. For example, the hazard for blacks is only .71 as high as for non-blacks.

In addition to race, Hispanic origin, number of children, and presence of young children all have substantial, negative coefficients. AFDC benefits have a well-estimated, moderate-sized effect (recall that its units are $100). The local variables have moderate-sized, fairly well estimated coefficients. For example, a one percent rise in LUNEM lowers the hazard by 4 percent. Thus persons in labor market areas with unemployment that is 2.38 percent above the mean LUNEM (one standard deviation above) would have about 10 percent lower hazards.

The second model includes no fixed effects and only 9 coefficients. It is included for comparison to the local fixed-effect model below.

Before turning to the fixed-effect results, let me briefly discuss whether we expect coefficients to become larger or smaller in absolute value after conditioning on the location effects. A simple omitted-variable interpretation would suggest that conditioning on location would attenuate coefficients on variables such as benefit levels because low-benefit states may be those that have tougher administration that would also give rise to shorter spells. But there is another effect that complicates. Ridder and Verbakel (1984) show that unobserved heterogeneity uncorrelated with covariates biases

TABLE 2

Partial Likelihood Estimates of Proportional Hazard Model of Welfare Duration

Variable	No Fixed Effect 11 coeff. Coeff.	Exp (Coeff)	No Fixed Effect 9 coeff. Coeff.	Exp (Coeff)	State Fixed Effect 11 coeff. Coeff.	Exp (Coeff)	Local Fixed Effect 9 coeff. Coeff.	Exp (Coeff)
Age	.00582	1.01	.00429	1.00	.00825	1.01	.0145	1.01
	(.00894)		(.00887)		(.00973)		(.0123)	
Black	-.343**	.71	-.406***	.67	-.414***	.67	-.563***	.57
	(.139)		(.137)		(.172)		(.222)	
Hispanic	-.318	.73	-.376*	.69	-.444*	.64	-.420	.66
	(.222)		(.221)		(.262)		(.306)	
Education	.0470	1.05	.0446	1.05	.0437	1.04	.0897**	1.09
	.0293		(.029)		(.0330)		(.0404)	
Number of Kids	-.136**	.87	-.127*	.88	-.106	.90	-.133	.88
	(.0656)		(.065)		(.0719)		(.0890)	
Presence of Kids	-.253	.78	-.234	.79	-.242	.79	-.149	.86
	(.161)		(.160)		(.179)		(.221)	
Other Income Dummy	-.0385	.96	-.0182	.98	.0109	1.01	-.0490	.95
	(.191)		(.190)		(.216)		(.277)	
ADFCmax	-.121***	.89	-.135***	.87	-.464	.63	-.384	.68
	(.0490)		(.0470)		(.549)		(.389)	

(table continues)

TABLE 2, continued

	No Fixed Effect 11 coeff.		No Fixed Effect 9 coeff.		State Fixed Effect 11 coeff.		Local Fixed Effect 9 coeff.	
Variable	Coeff.	Exp (Coeff)	Coeff.	Exp (Coeff)	Coeff.	Exp (Coeff)	Coeff.	Exp (Coeff)
Urban	-.319***	.73			-.305*	.74		
(.136)				(.182)				
LMA Unemployment	-.0458* (.0246)	.96	-.0296 (.0241)	.98	-.00636 (.0378)	.99	-.00323 (.0110)	1.00
LMA Sexratio	.0300* (.0192)	1.03			.0566* (.0329)	1.06		
Sample Size	533		533		527		438	
Number of Completed Spells								
Number of Fixed Locations					38		88	
Log Likelihood	-1482		-1487		-579.8		-290.9	
χ^2 Test for No Fixed Effect					8.23 (11 d.f.)		8.79 (9 d.f.)	

Notes: Standard errors in parentheses. Sample of unmarried mothers from 1984–1985 panels of SIPP. Asterisks indicate that coefficient is significantly different from zero at the 10 percent (*), 5 percent (**), or 2 percent (***) levels.

coefficients toward zero in proportional hazard models. Conditioning on the fixed effect removes some heterogeneity and thus the coefficients should become larger in absolute value.

The results for the model with fixed state effects are shown in the third pair of columns. My results tend to show the latter effect mentioned above. The coefficients on black and Hispanic are larger in magnitude; conditioning on state of residence, blacks and Hispanics have longer spells. The coefficient on AFDC benefits also becomes larger in size, but the standard error becomes relatively larger. This coefficient is estimated using only the within-state variation over time in benefit levels; when real benefits rise over time, those within a state will stay on welfare longer. There is apparently not enough variation over the years 1983 to 1987 to get a precise estimate. But the coefficient on unemployment is near zero, and poorly estimated; our labor market area unemployment rates may not vary enough within states to get more precision. Another way of saying this is that a state’s economic condition may be fairly homogeneous over the time span analyzed. The coefficient on sex ratio becomes larger in size, conditioning on state. Overall, the qualitative results are similar to the model with no fixed effects.

More formally, one can test whether the state fixed effects are significant. The test statistic is based on the difference between the coefficient vector with and without fixed effects weighted by the appropriate variance matrix.[13] Ridder and Tunali (1990) derive this Hausman-type statistic and show that it has a Chi-squared distribution. The statistic for state fixed effect, shown at the bottom of Table 2, has a value of 8.23, which indicates that fixed effects are not statistically significant at usual significance levels. This suggests that the coefficients in the models, as a group, are not very different. Thus the results without fixed effects are not misleading.[14]

The last two columns show the model with fixed labor market effects, that is, using only variation within labor market areas. This removes further heterogeneity and the coefficients again generally rise in absolute value, especially for personal characteristics. Black continues to have a well-estimated coefficient and shows an even larger negative effect on exits. Hispanic origin continues to have the large effect that it had in the fixed state-effect model. The coefficient on education doubles and is well estimated. Within a labor market, higher education levels have a big impact, larger than in previous studies. AFDC benefits again have a larger coefficient than without fixed effects, but it is poorly estimated, most likely owing to lack of variation. The LUNEM coefficient is poorly estimated, again probably owing to lack of variation. The remaining coefficients change little. The χ^2 test for no fixed effects is shown at the bottom of the table, and is again not significant (at 5%).

Overall, the effects of race and ethnicity are very important after conditioning on labor market area. One might have suspected that conditioning more fully on local environmental characteristics by fixed effects would have reduced the black/white difference in exit rates--since black areas tend to have less robust labor markets--but this is apparently overbalanced by the rise in the size of the coefficients due to the reduction in heterogeneity. Generally, personal characteristics are <u>more</u> important when one controls for location heterogeneity.[15]

V. ROBUSTNESS AND FIT

To evaluate the fit of the models, I used plots of the generalized residuals as described by Lawless (1982, pp. 365–366).[16] The underlying hazard from the proportional hazard model can be recovered, and one can form an estimate of the integrated hazard, say $\hat{H}_0(t)$. A generalized residual $\hat{e}_i = \hat{H}_0(t)\exp(B'x_i)$ can be formed for each spell. In the absence of censoring, the e_i's should look like a random sample from the unit exponential distribution. In the presence of censoring, one constructs a set of censored residuals and uncensored residuals. These can be combined to estimate a (product-limit) survivor function. This survivor function should be consistent with the underlying unit exponential.

Figure 1 shows the log of the survivor function from the residuals plotted against the value of the residual. If the proportional hazard model is adequate, the residuals shown fall along a straight line with slope -1 that corresponds to the unit exponential. The plots do not show large departures. There appears to be growing departure at high values of residuals, but we must remember that the sampling error grows as we move out along the axis as well. Overall, I conclude that there is not evidence from this test of a bad fit for the proportional hazard model.

To check sensitivity, I ran a discrete-time proportional hazard model as developed by Prentice and Gloeckler (1978) (without fixed effects) for comparison to the assumed continuous-time model. This discrete specification is developed by integrating a continuous-time proportional hazard model into discrete intervals. The model is explained in the appendix. It allows a stepwise hazard that is very flexible. Since it is discrete, it handles ties in the spell length in a natural way and provides a check on my treatment of ties in the continuous-time models of the last section.[17] Table 3 shows the results. The coefficients and standard errors are in close agreement with the no-fixed-effect model of Table 2, confirming that my assumption of a continuous-time model in Table 2 is not misleading.

The second check allows for heterogeneity in the discrete model. I estimated the discrete hazard model, without fixed effects, allowing for an individual specific heterogeneity component to multiply the proportional hazard. I followed the approach of Meyer (1988), who develops the likelihood for this model. The likelihood is formed by conditioning on the individual component and then integrating it out over its assumed distribution, taken to be gamma. (See the Appendix for details.) To the extent heterogeneity is important, not correcting for heterogeneity produces an underlying hazard that overestimates the true state dependence due to welfare use (e.g., Lancaster, 1979). In this model, I found that the variance of the gamma distribution tends to zero, indicating little unmeasured heterogeneity. The estimate of sigma has a large standard error, and we cannot reject that sigma equals zero. The estimated coefficients from this model are extremely close to those

FIGURE 1

TABLE 3

Discrete Hazard for Welfare Duration: Complementary Log-Log Specification with No Local-Specific Fixed Effects

Variable	Complementary Log-Log	Complementary Log-Log with Gamma Heterogeneity
Constant	-4.47**	-2.030***
	(1.87)	(.162)
Age	.0062	-.0756
	(.0090)	(.0901)
Black	-.357***	-.333**
	(.139)	(.138)
Hispanic	-.321	-.315
	(.222)	(.223)
Education	.0491*	.599**
	(.0290)	(.288)
Number of Kids	-.141**	-.178***
	(.0656)	(.0676)
Presence of Child < 6	-.261	-.243
	(.161)	(.158)
Other Income Dummy	-.0646	-.0798
	(.191)	(.194)
AFDCmax	-.125***	-.124**
	(.0490)	(.0507)
Urban	-.331**	-.358***
	(.136)	(.136)
LMA Unemployment	-.0480*	-.0465*
	(.0246)	(.0254)
LMA Sexratio	.0308*	-.0302
	(.0192)	(.195)
T2 (5-8 months)	-.447***	-.445***
	(.159)	(.159)

(table continues)

TABLE 3, continued

Variable	Complementary Log-Log	Complementary Log-Log with Gamma Heterogeneity
T3 (9-12 months)	-.486***	-.484***
	(.189)	(.186)
T4 (13-16 months)	-1.10***	-1.081***
	(.302)	(.298)
T5 (17-20 months)	-.734**	-.717**
	(.328)	(.326)
T6 (20+ months)	-.890**	-.880***
	(.365)	(.366)
Sigma	--	.0154
		(.291)
Sample Size (Persons)	533	533
Log Likelihood	-980.6	-978.9

Notes: Standard errors in parentheses. Sample of unmarried mothers from 1984, 1985 panels of SIPP. Asterisks indicate that coefficient is significantly different from zero at the 10 percent (*), 5 percent (**), or 1 percent (***) level.

from the model assuming zero heterogeneity. This is consistent with work cited by Meyer (1988) that heterogeneity is generally not important when one allows a very flexible functional form for the underlying hazard.

VI. CONCLUSION

The results establish that personal characteristics are important determinants of welfare durations even after one controls for location by fixed effects. In particular, with labor market area fixed effects, being black or Hispanic continues to result in lower welfare exit hazards. Thus the race variable is not simply picking up bad labor market attributes. Education has a large positive effect on exit rates even among those who live in the same labor market area. The biases on personal characteristics due to omission of fixed state or local effects appears to be limited to the characteristics of race and ethnicity and education; other coefficients on age, number of children, presence of young children, and other income show little change when fixed effects are added.

AFDC benefits may have a larger effect after controlling for the fixed effects, but the standard errors for AFDC benefits in the fixed-effect models are so large that they inspire little confidence. Measured local area variables have moderate impacts without fixed effects. Conditional on state fixed effects, unemployment rates have a small and poorly estimated effect, most likely due to lack of variation within states. Local sex ratios have a larger positive effect on exits after controlling for state-level effects. Both AFDC benefit and measured local area effects suffer from lack of within-state variation, which could potentially be solved by data with more time variation. A final problem is that the geographic area used for local measures, the LMA, may be too large an aggregate to accurately measure local influences. Nevertheless, by virtue of its definition, it is the appropriate level of aggregation for judging labor market strength.

Overall, the data suggest that standard estimates of the effects of some personal characteristics (race, Hispanicity, education) are biased owing to unmeasured state or local effects. Other personal traits are not biased. AFDC benefit levels may have a larger impact once one controls for other unmeasured state effects, but the data are not strong enough on this point to say more. Even though the fixed-effects technique did not produce strikingly different results in this application, it offers an improved way of dealing with state or local heterogeneity in policy evaluations.

Appendix
Discrete Hazard Model and Heterogeneity

This appendix describes a complementary log-log form of discrete hazard. Assume the underlying continuous hazard has a proportional hazards form $\lambda_i(t) = \lambda_0(t)\exp(B'X_i(t))$ with baseline hazard $\lambda_0(t)$. If the data are discrete, then exit probabilities for each interval can be computed by integrating $\lambda_i(t)$ over each interval. This is the approach of Prentice and Gloeckler (1978). See discussion in Allison (1982) or Meyer (1988).

The discrete hazard becomes

$P(t) = 1-\exp(-\exp(\alpha(t) + B'X(t)) = 1-\exp(-h(t))$

where $h(t) = \exp(\alpha(t) + B'X(t))$ and $\alpha(t)$ represents the baseline hazard as a piecewise linear function, which allows flexibility.

The sample log likelihood becomes

(1)

$$\mathcal{L}(\alpha, B) = \sum_{i=1}^{N} \delta_i \left\{ \log\left[1 - \exp\left(-h(t_i)\right)\right]\right\} + \sum_{t=1}^{t_i - 1} \exp(-h(t))$$

where

$$\delta_i = \begin{cases} 1 \text{ for complete spells} \\ 0 \text{ for censored spells} \end{cases}$$

and

$$t_i - 1 = \begin{cases} \text{month before exit for complete spells} \\ \text{last observed month for censored spells} \end{cases}$$

The text also presents results that allow for unobserved heterogeneity. Meyer (1988) extends the discrete hazard model to include heterogeneity, and I adopt his method. He argues that once one adopts a flexible semi-parametric baseline hazard, the exact choice of a parametric distribution for the heterogeneity may not matter. I use a flexible hazard similar to Meyer's and a parametric distribution for heterogeneity.

Assume that heterogeneity affects the underlying hazard multiplicatively:

$$\lambda_i(t) = \theta_i \lambda_0(t) \exp\left(B' X_i(t)\right)$$

where θ_i is a draw from some distribution $F(\theta_i)$, assumed independent of $X_i(t)$ and the censoring mechanism. The log likelihood is obtained by conditioning on the unobservable θ_i and then integrating it out over its distribution. Meyer shows that the resulting log likelihood is

(2)

$$\mathcal{L} = \sum_{i=1}^{N} \log \left\{ \int_\theta \exp\left[-\theta \sum_{t=1}^{t_i - 1} h(t) \right] dF(\theta) - \delta_i \int_\theta \exp\left[-\theta \sum_{t=1}^{t_i} h(t) \right] dF\theta \right\}$$

Meyer makes the usual convenient assumption that θ has a gamma distribution with mean one and variance σ^2. In this case, the integration can be done in closed form yielding the log likelihood

(3)

$$\mathcal{L} = \sum_{i=1}^{N} \log \left\{ \left[1 + \sigma^2 \sum_{t=1}^{t_i - 1} h(t) \right]^{-\sigma^{2}} - \delta_i \left[1 + \sigma^2 \sum_{t=1}^{t_i} h(t) \right]^{-\sigma^{2}} \right\}.$$

Endnotes

[1]Examples using state-level variables include Bane and Ellwood (1983), Ellwood (1986), O'Neill, Bassi, and Wolf (1987), Ruggles (1989), Long and Doyle (1989), or Fitzgerald (1991). Two papers use local-area variables measured below the state level. Blank (1989) uses SMSA unemployment rates in her work with SIME/DIME data, but, accordingly, she has data only from Seattle and Denver. Fitzgerald (1994) uses nationwide data from the Survey of Income and Program Participation linked to county and labor-market-area data on environmental variables.

[2]This does not cause a problem unless these decisions are made jointly with the welfare participation decision.

[3]The estimation is performed using Newton-Raphson iterations in GAUSS software. The standard errors are computed from the inverse of the estimated Hessian. Consistency of the estimators can be justified by two asymptotic arguments. Chamberlain (1985, p. 24) cites Cox's argument for the case of multiple spells per individual where the asymptotics depend on the number of spells per person increasing. In my context this corresponds to the number of persons per location increasing, which a larger SIPP sample would provide. When dealing with labor market areas, my sample actually has relatively more locations than persons per location. This corresponds to Ridder and Tunali (1990) who prove consistency in a child mortality model with family fixed effects as the number of families increase.

[4]Half of the 1984 panel was interviewed nine times, and half eight, with 15 percent of the sample cut at interviews 5 and 6. The longitudinal research files contain information from eight interviews.

[5]The SIPP uses a rotating, staggered interview design whereby one-fourth of the sample is interviewed each month. Thus the calendar time span of the panel exceeds 32 months, but not all persons are interviewed at the ends.

[6]Problems with misreporting of recipiency have been documented by Coder and Ruggles (1988) and others. SIPP includes many married couples on AFDC with earnings in states where they would ordinarily be ineligible. Further, many men report receiving AFDC who would also be ineligible. To guard against misreporting, consistency checks were performed to ensure that a sample woman was categorically AFDC eligible, i.e., unmarried and a parent or guardian. To prevent timing of reported events within a spell to cause me to drop spells on this account, I allowed one month of slippage within a spell where a woman could report marriage or no children.

[7]I reason that unmarried women with children who report receiving General Assistance are most likely receiving AFDC. An administrative record check supports this assumption. Kent Marquis and Jeff Moore of the Census Bureau kindly prepared an analysis comparing recorded receipt of AFDC from state administrative records for a four-state convenience sample, to reported receipt of (a) AFDC alone and (b) AFDC or General Assistance. To the extent possible the analysis worked with unmarried adult women with children in their households. Under definition (b), the analysis showed a large drop in false reports of non-receipt, where SIPP showed no AFDC receipt and the "true" recipiency from administrative records showed receipt, from 35 percent to 5 percent. Definition (b) does lead to a slight rise in false reports of receipt (to 6 percent from 3 percent), but this does not outbalance the former error reduction.

[8]Work by Fitzgerald and Zuo (1991) suggests that results using imputed data would be quite similar.

[9]I had access to internal files because I was a Census employee while participating in the ASA/NSF/Census Fellowship program.

[10]If a county had a missing data item, data from the remaining counties within the LMA were used to get LMA values.

[11]While it might be more accurate to use the benefit adjusted by family size, this would add some endogeneity to the benefit measure since family size could potentially depend on the benefit level.

[12]An earlier paper, Fitzgerald (1991), obtained a longer median length of 20 months for two reasons: (1) the earlier work coded out up to three-month gaps while the current sample codes out only one-month gaps, and (2) the earlier work did not include reported General Assistance cases, which tend to have shorter spells, while this paper includes such cases. The hazard models for the two samples are very similar.

[13]The statistic is defined as

$$T = (\hat{\boldsymbol{B}}_s - \hat{\boldsymbol{B}}_u) V(\hat{\boldsymbol{B}}_s - \hat{\boldsymbol{B}}_u)^{-1} (\hat{\boldsymbol{B}}_s - \hat{\boldsymbol{B}}_u)$$

where B_s is the coefficient vector from the "stratified" model with fixed effects, and B_u is the coefficient vector for the "unstratified" model with no fixed effects. Ridder and Tunali (1990, pp. 23–25) prove that the variance matrix can be computed simply as

$$V(\hat{\boldsymbol{B}}_s - \hat{\boldsymbol{B}}_u) = V(\hat{\boldsymbol{B}}_s) - V(\hat{\boldsymbol{B}}_u),$$

where the right-hand-side terms are the estimated variance/covariance matrices from the separate estimations. The statistic T is shown to be distributed $\chi^2(p)$ where p is the number of coefficients in the model.

[14]Ridder and Tunali do not investigate the power of their test. In their own example, they get a significant test statistic only after dropping statistically insignificant variables from their model. They state (1990, p.27) that these variables adversely affect the power of their test.

[15]I investigated one other possible explanation. As mentioned earlier, a location must have at least 2 spells, 1 complete, to contribute to the likelihood. As we move from the fixed-effect model to that of the LMA fixed effect, we lose some sample spells. To check that this was not biasing my results (in violation of condition 4 from Section II), I ran the no-fixed-effect model on the smaller sample that contributes to the LMA model. The smaller sample tends to be more urban and more black. I found that the coefficient on black became somewhat smaller in absolute value for the restricted sample. Thus the variation in sample does *not* explain why the coefficient on black is bigger for the fixed effect LMA model. That is, holding sample constant, there is even a larger difference between the no-fixed-effect and LMA fixed-effect models.

[16]Residual tests are described in many sources. For example, see Lancaster (1990, Ch. 11).

[17]In the previous estimations, ties were handled by allowing each tied observation to contribute to the likelihood as if it were a single observation. This is essentially Peto's approximation (Kalbfleish and Prentice, 1980, p. 74).

References

Allison, Paul D. 1982. "Discrete-Time Methods for the Analysis of Event Histories." In *Sociological Methodology*, edited by Samuel Leinhart. San Francisco: Jossey-Bass.

Bane, Mary Jo, and David Ellwood. 1983. *The Dynamics of Dependence: The Route to Self-Sufficiency.* Report prepared for U.S. Department of Health and Human Services: Urban Systems Research and Engineering, Inc.

Blank, Rebecca. 1989. "Analyzing the Length of Welfare Spells." *Journal of Public Economics* 39: 245–274.

Chamberlain, Gary. 1985. "Heterogeneity, Omitted Variable Bias, and Duration Dependence." In *Longitudinal Analysis of Labor Market Data*, edited by James Heckman and Burton Singer. Cambridge, U.K.: Cambridge University Press.

Coder, John, and Patricia Ruggles. 1988. "Welfare Recipiency as Observed in the SIPP." *SIPP Working Paper Series* No. 8818, U.S. Bureau of the Census, Washington, D.C.

Ellwood, David. 1986. *Targeting 'Would Be' Long-Term Recipients of AFDC.* Report prepared for U.S. Department of Health and Human Services. Princeton, N.J.: Mathematica Policy Research.

Fitzgerald, John. 1991. "Welfare Durations and the Marriage Market: Evidence from SIPP." *Journal of Human Resources* 26 (Summer): 546–561.

Fitzgerald, John. 1994. "Local Labor Markets and Local Area Effects on Welfare Duration." *Journal of Policy Analysis and Management* (forthcoming Winter 1995).

Fitzgerald, John, and Xeujin Zuo. 1991. "Alternative Samples for Welfare Durations in SIPP: Does Attrition Matter?" *Proceedings of the Seventh Annual Research Conference.* Washington, D.C.: U.S. Bureau of the Census.

Kalbfleisch, John D., and Ross L. Prentice. 1980. *Wiley Series in Probability and Mathematical Statistics.*, *The Statistical Analysis of Failure Time Data.* New York: John Wiley and Sons, Inc.

Lancaster, Tony. 1979. "Econometric Methods for the Duration of Unemployment." *Econometrica* 47: 939–956.

Lancaster, Tony. 1990. *The Economic Analysis of Transition Data*. New York: Cambridge University Press.

Lawless, Jerald F. 1982. *Statistical Models and Methods for Lifetime Data*. New York: John Wiley and Sons, Inc.

Long, Sharon K., and Pat Doyle. 1989. "Welfare Recipiency and Welfare Recidivism: An Analysis of Short-term Dynamics." Unpublished manuscript, presented at the Allied Social Science Association Meetings.

Marquis, Kent, and Jeff Moore. 1989. "Response Errors in SIPP: Preliminary Results." In *Proceedings of the Fifth Annual Research Conference*. Washington, D.C.: U.S. Bureau of the Census.

Meyer, Bruce D. 1988. "Semiparametric Estimation of Hazard Models." Unpublished manuscript, Northwestern University, Evanston, Ill.

Nelson, D., D. McMillen, and D. Kasprzyk. 1985. "An Overview of the SIPP, Update 1." *SIPP Working Paper Series* No. 8401, U.S. Bureau of the Census, Washington, D.C.

O'Neill, June, Laurie Bassi, and Doug Wolf. 1987. "The Duration of Welfare Spells." *Review of Economics and Statistics* 69, no. 2 (May): 241–248.

Prentice, R. L., and L. A. Gloeckler. 1978. "Regression Analysis of Grouped Survival Data with Applications to Breast Cancer Data." *Biometrics* 34: 57–67.

Ridder, Geert, and Insan Tunali. 1989. "Analysis of Related Durations: A Semi-parametric Approach." *Cornell University Working Paper* No. 425, Ithaca, N.Y.

Ridder, Geert, and Insan Tunali. 1990. "Family Specific Factors in Child Mortality: Stratified Partial Likelihood Estimation." Unpublished manuscript, University of Groningen, The Netherlands.

Ridder, Geert, and W. Verbakel. 1984. "On the Estimation of Proportional Hazards Model in the Presence of Unobserved Heterogeneity." Unpublished manuscript, Institute of Actuarial Science and Econometrics, University of Amsterdam.

Ruggles, Patricia. 1989. "Welfare Dependency and Its Causes: Determinants of the Duration of Welfare Spells." *SIPP Working Paper Series* No. 8908, U.S. Bureau of the Census, Washington, D.C.

SIPP. 1989. *Survey of Income and Program Participation (SIPP) 1984 Full Panel Microdata Research File.* Prepared by Data User Services Division. Washington, D.C.: U.S. Bureau of the Census.

Tolbert, Charles N., and Molly Sizer Killian. 1987. *Labor Market Areas for the United States.* Staff Report No. AGES870721. Washington, D.C.: Agriculture and Rural Economy Division, Economic Research Service, U.S. Department of Agriculture.

U.S. House of Representatives, Committee on Ways and Means. 1987. *Green Book.* Washington, D.C.: U.S. Government Printing Office.