The Employment Dynamics of Disadvantaged Women: Evidence from the SIPP*

John C. Ham
University of Maryland and the National University of Singapore, IFAU, IFS, IRP (UW-Madison) and IZA

Xianghong Li
York University

Lara Shore-Sheppard
Williams College and NBER

Revised February 2014

* This paper was previously circulated under the title “Analyzing Movements over Time in Employment Status and Welfare Participation while Controlling for Seam Bias using the SIPP.” We are grateful for the comments received at the IRP Summer Research Workshop, as well as seminars at many universities and branches of the Federal Reserve Bank. These suggestions significantly improved the paper. Sandra Black, Richard Blundell, Mary Daly, Marcus Frölich, Soohyung Lee, Jose Lopez, Bruce Meyer, Robert Moffitt, Geert Ridder, Melvin Stephens, Lowell Taylor and Tiemen Woutersen made very helpful remarks on an earlier draft, as did an Editor and two anonymous referees. Eileen Kopchik provided, as usual, outstanding programming in deciphering the SIPP. This research has been supported by grants from the National Science Foundation (SES-0136928 and SES-0627968). Any opinions, findings and conclusions, or recommendations in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. We are responsible for any errors. Contact information: John Ham (corresponding author), echejc@nus.edu.sg, Department of Economics, Faculty of Arts and Social Sciences, National University of Singapore, AS2 #06-02, 1 Arts Link, Singapore 117570, phone: (65) 8323 1085 fax: (65) 6775 2646; Xianghong Li, xli@econ.yorku.ca, Department of Economics, York University, 4700 Keele Street, Vari Hall 1068, Toronto, ON M3J 1P3, Canada, phone: (416) 736-2100 ext. 77036, fax: (416) 736-5987; Lara Shore-Sheppard, Lara.D.Shore-Sheppard@williams.edu, Department of Economics, Williams College, 24 Hopkins Hall Drive, Williamstown, MA 01267, phone: (413) 597-2226, fax: (413) 597-4045.
ABSTRACT

Estimating models of employment dynamics for disadvantaged families is becoming increasingly important. We estimate multi-state, multi-spell duration models describing the employment dynamics of disadvantaged single mothers, and conduct a rich set of counterfactual policy analyses. Unlike much of the current literature, we consider both left-censored and fresh spells, since the former dominate our data, and we allow for time-changing explanatory variables, since the latter are known to improve considerably the empirical identification of the model. We find that the most important policy effects occur in left-censored spells, indicating the importance of including them in our analysis. The Survey of Income and Program Participation (SIPP) is a natural data set for this purpose. However the SIPP suffers from “seam bias” - a situation where too few transitions are observed within the reference periods and too many transitions are reported between the reference periods. Seam bias is likely to be especially serious for duration models because both the start date and the end date of an employment or non-employment spell may be misreported. In this paper we introduce and estimate misreporting models that correct for seam bias in the SIPP. We compare the results from our misreporting models with those obtained from the much simpler ad-hoc ones currently used in the literature, and find that both of the current methods do a relatively poor job of dealing with seam bias.
1. Introduction

Models describing movements on and off welfare for disadvantaged families have been widely studied by labor economists, but there has been substantially less work studying movements in and out of employment for the heads of these families. However, employment dynamics have increased in importance due to a new focus on work in policies targeting low-income families, including welfare reform, expansions in the Earned Income Tax Credit and expansions in Medicaid coverage. In this paper we fill this gap in the literature by estimating multi-state, multi-spell duration models describing the employment dynamics of disadvantaged single mothers. We then use our estimates to conduct a rich set of counterfactual policy analyses. We focus on the period 1986-1995, a time of fundamental change in antipoverty policy in the United States. During this time, federal and state policy moved from a regime focused on cash assistance and programs with built-in disincentives to work, to a new regime focused on encouraging work through changes in the structure of cash assistance and greater provision of in-kind benefits such as child care subsidies and health insurance.

Unlike many other studies, our paper considers spells in progress at the beginning of the sample as well as those that start during the sample period. Such left-censored spells dominate the employment histories of disadvantaged mothers and thus are absolutely crucial to understanding the employment dynamics of such women. Further, we find that the most important policy effects occur in left-censored spells, indicating the importance of including them in our analysis.

We use the Survey of Income and Program Participation (SIPP), since it is the only data set that provides both monthly data on employment and a rich set of covariates. However, many panel surveys in general, and the SIPP in particular, suffer from “seam bias” - too few transitions observed within reference periods and too many transitions reported between reference periods. Seam bias is likely to be especially important for multi-spell duration models since both the start date and the end date of a spell may be misreported. Previous approaches to seam bias generally take one of two ad-hoc forms. One approach is to use only the last month observation from each retrospective period (known as a “wave” in the SIPP and

---

1 See our literature review in Section 2.

2 In economics, this data problem was noted first by Czajka (1983) for benefit receipt in the U.S. Income Survey Development Program. Since then, seam effects have been documented for various longitudinal surveys in many North American and European countries, e.g. the SIPP, the Current Population Survey, the National Longitudinal Surveys, the Panel Study of Income Dynamics, the Canadian Survey of Labour and Income Dynamics, the European Community Household Panel Survey, and the British Household Panel Survey. Lemaitre (1992) concludes that all current longitudinal surveys appear to be affected by seam problems, regardless of differences in the length of recall periods or other design features. The Census Bureau, which collects the SIPP, has long recognized this problem and has attempted to reduce it in the SIPP, most recently by incorporating “dependent interviewing” procedures in the 2004 panel of the survey. Notwithstanding the adoption of such procedures, which explicitly link the wording of current interview questions to information provided in the preceding interview, seam bias continues to be a substantial problem in the SIPP (Moore 2008).

3 Abrevaya and Hausman (1999) offer a very general approach to mismeasurement of duration within the context of a single spell model, but their approach cannot be used for our purposes since it does not allow for all of the hazard function parameters to be estimated.
covering four months), dropping the other months. Two of the most prominent applications of this approach are in Grogger (2004) and Ham and Shore-Shappard (2005), and in what follows we refer to this as the GHS approach. A second approach to seam bias is to use the monthly data and to include a dummy variable for the last month of the reference period (e.g. Blank and Ruggles 1996 and Fitzgerald 2004); the appropriateness of this approach depends on whether the problems introduced by seam bias can be effectively eliminated by the use of a last month dummy.

Here we propose and implement a model of misreporting based on the concept of telescoping from the statistics literature to correct for seam bias; the more general forms we implement allow for heterogeneity across individuals in terms of misreporting. We work within the context of a discrete time multi-spell, multi-state model with unobserved heterogeneity. We allow each hazard function to depend on a host of policy variables (such as welfare benefits, a variety of welfare program features introduced by states to induce women to leave welfare, and the minimum wage) as well as business cycle conditions, current duration, and demographic variables. Unlike many studies, we allow for the relevant explanatory variables to change during and across spells, satisfying the sufficient condition for identifying discrete time duration models with unrestricted duration dependence and unobserved heterogeneity (Hausman and Woutersen 2013). Since we use our correction for misreporting within the context of a much richer economic model than those considered by authors who have previously addressed misreporting, we face identification issues not encountered in earlier work. As a result we need to constrain the parameters describing misreporting parameters to be the same in left-censored and fresh employment (or non-employment) spells to obtain identification without restricting duration dependence parameters.

Using our estimates, we calculate the effect, and the corresponding confidence interval, of individually changing each policy variable on i) the expected duration of each type of labor market spell and ii) the fraction of time spent in employment in the short run, medium run, and long run. Effects on the fraction of time spent in employment have rarely been reported in the literature and ours is the first study to provide correct confidence intervals for these effects. As noted above, we find that the effects of changing policy or demographic variables are much larger in left-censored spells than in fresh spells. This finding is important because previous studies generally ignore left-censored spells, calculating only the effects of a policy change on the hazard function or on the expected duration of fresh spells. Thus, our estimates (with appropriate confidence intervals) will be much more useful for policy makers.

4There is a third approach to seam bias which we do not consider. This approach involves collapsing the employment indicator for monthly data into an employment indicator for the (four-month) reference period. The reference period employment indicator is set to 1 if the fraction of time that respondents are employed or on welfare in the reference period is above some threshold (e.g. Acs, Philips, and Nelson 2003 and Ribar 2005). However, the fraction of time threshold for determining the employment indicator will be quite arbitrary. Further, this approach is likely to result in the loss of short spells, and there are likely to be many short employment spells for disadvantaged women.

5 Ham and LaLonde (1996) is the first duration study to estimate these effects, but they do not calculate the respective confidence intervals. Osikominu (2013) estimates similar effects but calculates inconsistent confidence intervals for them.
We begin by discussing previous work on estimating the determinants of employment dynamics (Section 2). In Section 3 we describe our sample and the seam bias problems in the SIPP, and then review the commonly used seam bias correction methods in the literature, focusing on the GHS approach where only the last months of the reference period are used. We present our econometric model and describe our counterfactual policy analyses in Section 4, and describe our seam bias correction approach in Section 5. We present the empirical results, including the policy effects, in Section 6. We find that there are important differences between our policy effect estimates and those from the GHS approach for all of the effects we consider, and there are substantial differences between our estimated policy effects and those from the last month dummy approach for the demographic variables. We conclude the paper in Section 7.

2. The Employment Dynamics of Disadvantaged Women

Transitions into and out of employment are of crucial importance to policy makers, as they determine unemployment rates, poverty rates, and the overall well-being of low-income families. A clear understanding of employment dynamics is essential for policymaking. For example, policy makers are likely to be very interested in the determinants of employment duration, since short employment spells prevent disadvantaged individuals from acquiring on-the-job human capital. Our target population is single mothers with a high school education or less, a group that has long been the focus of public policy. We estimate monthly transition rates into and out of employment for an important period in US policy towards low-income single mothers, 1986-1995, which culminated in the replacement of Aid to Families with Dependent Children (AFDC) with Temporary Assistance to Needy Families (TANF). This period covered a fundamental shift in the structure of US policy towards single mother families, with an increased emphasis on work and somewhat greater provision of in-kind benefits including health insurance for children via Medicaid and child care subsidies.

While there is a large European literature on employment dynamics for disadvantaged men and women (see e.g. Ridder (1986), Ham, Svejnar and Terrell (1998), Cockx and Ridder (2003), Sianesi (2004), Micklewright and Nagy (2005), Johansson and Skedinger (2009), and Blundell, Francesconi and Van der Klaauw (2010)), relatively few North American studies have focused explicitly on employment dynamics of the disadvantaged; exceptions are Ham and LaLonde (1996), Eberwein, Ham and LaLonde (1997), and Aaronson and Pingle (2006). The related topic of welfare dynamics for less-educated women has been examined in several U.S. and Canadian studies, e.g. Blank and Ruggles (1996), Grogger (2004), Card and Hyslop (2005), and Keane and Wolpin (2002, 2007). However, employment dynamics and welfare dynamics will differ, as single mothers can work and collect welfare simultaneously, and can certainly be out of employment and off welfare simultaneously.

We focus on both describing the relationship between demographic factors and employment and non-employment durations, and on examining how key welfare policy changes implemented during this
period affected employment and non-employment duration. We also examine the influence of two important labor market factors, the minimum wage and the unemployment rate. The demographic factors we study include age, education, race, ethnicity, and number of children under the age of 6. Knowledge of how such factors affect employment duration is critical for accurate targeting of policy. To the extent that women who have not completed high school have shorter employment spells and longer non-employment spells, for example, children in such families are likely to experience greater reductions in well-being from policies promoting work. During the period that we study, many states experimented with such policies, receiving waivers from the federal government to permit welfare program experimentation. We follow the literature (see e.g. Blank 2002) and characterize the waiver policies as being either positive incentives to work (“carrot” policies) or changes intended to raise the cost of not working (“stick” policies). We consider an increase in the amount of earnings a woman is allowed to keep before losing welfare benefits (an increased earnings disregard) as a “carrot” policy, while as “stick” policies we include: implementation of a time limit for welfare receipt, a work requirement for welfare receipt, a family cap (denial of additional benefits or a reduction of benefits for women who have additional children while on welfare), and stronger requirements for participation in training and job readiness programs. Such policies may have had impacts on employment and non-employment duration, but little is known about the magnitude of such effects. Finally, we consider the magnitude of effects of higher unemployment rates as a measure of macroeconomic conditions likely to affect employment and non-employment durations, and effects of higher minimum wages.

3. **Seam Bias and a Commonly Used Correction Method in the Literature**

3.1 **Seam Bias and the SIPP Data**

Our primary data come from the 1986-1993 panels of the SIPP. The SIPP was designed to provide detailed information on income sources and levels, labor market outcomes, and program participation. Our sample is restricted to single mothers who have at most twelve years of schooling. Because we investigate employment status, we consider women only between the ages of 16 and 55.6 Although researchers investigating welfare durations often smooth out one-month spells, we use the original data with all one-month spells intact because employment status is often very unstable among low-educated women and it is common for them to have very short employment and non-employment spells.7 Because we use state level variables such as

---

6 Respondents are chosen based on their education and age at the beginning of the panel. If a single mother marries in the middle of the survey, we keep the observations before the marriage and treat the spell in progress at the time of marriage as right-censored. Dropping these women entirely does not affect our results.

7 Hamersma (2006) investigates a unique Wisconsin administrative data set containing information from all Work Opportunity Tax Credit (WOTC) and Welfare-to-Work Tax Credit applications. The majority of WOTC-certified workers in Wisconsin are either welfare recipients or food stamp recipients. She finds that over one-third of certified workers have fewer than 120 annual hours of employment (job duration), while another 29 percent of workers have fewer than 400 annual hours. Only a little over one-third of workers have annual employment of more than 400 hours. These administrative data show that a significant share of employment spells are less than one month among
maximum welfare benefits, minimum wage rates, unemployment rates, and whether the state obtained a welfare waiver as discussed above, we exclude women from nine small states (Maine, Vermont, Iowa, North Dakota, South Dakota, Alaska, Idaho, Montana, and Wyoming) as these states are not separately identified in the 1986-1993 SIPP panels.

The SIPP uses a rotation group design, with each rotation group consisting of about a quarter of the entire panel, randomly selected. For each calendar month, members of one rotation group are interviewed about the previous four months (the reference period or wave), and all rotation groups are interviewed over the course of any four-month period. Calendar months are thus equally distributed among the months of the reference period. We call the four months within each reference period month 1, month 2, month 3 and month 4. We will also refer to month 4 as the last month. The rotation design guarantees that approximately 25% of transitions should occur in months 1, 2, 3, and 4 respectively. However, summary statistics show that for our sample, 46% of all job transitions (i.e. from non-employment to employment and vice-versa) are reported to occur in the last month, and this percentage is far greater than the 25% one would expect. This phenomenon of too many transitions across waves and too few transitions within waves is observed for most variables in the SIPP (see e.g. Young 1989, Marquis and Moore 1990, Ryscavage 1988 and Moore 2008).

The problem of seam bias is not unique to the SIPP, as we noted previously, but is common in panel survey data. Many of the employment and welfare duration studies for Canada, the U.S. and the UK cited above used panel surveys. Consequently, these studies often were forced to confront (or ignore) the seam bias problem. On the other hand, many of the continental European studies cited above used administrative data. We would expect seam bias, and misreporting in general, to be less of a problem in administrative data, although to the best of our knowledge this issue has not been thoroughly investigated. Further, Chakravarty and Sarkar (1999) note that there appears to be seam bias at the end of the month in monthly administrative financial data, and Johansson and Skedinger (2009) argue that there is misreporting in Swedish administrative data on disability status. In future it would be interesting to investigate whether there is substantial seam bias in European administrative labor market data; until this is done we believe our results will be of most use to researchers studying the Canadian, U.S. and UK labor markets.

In our examination of employment dynamics, we follow Heckman and Singer (1984a) and allow the hazard functions for left-censored spells to differ from the hazard functions for fresh spells for time spent both in and out of employment. Some researchers “eliminate” the complications raised by left-censored spells by excluding them from their analyses, using only fresh spells. We do not proceed this way for two reasons. First, left-censored spells constitute the great majority of all spells in our data; thus, using only fresh spells will not give an accurate picture of the employment dynamics of a typical woman in our sample.
Second, there is an important issue of potential selection bias regarding which women are observed in a fresh spell, and we analyze left-censored and fresh spells jointly to correct for this selection.

Table 1 provides summary statistics for the first month of each spell for our sample of employment and non-employment spells. Panel A shows that 3528 single mothers have left-censored non-employment spells while 1889 single mothers have fresh non-employment spells; note that these two groups are not mutually exclusive. The single mothers in left-censored non-employment spells are on average observably more disadvantaged: they are more likely to be Hispanic, less likely to have a full twelve years of schooling, less likely to have had a previous marriage, are somewhat younger, and are more likely to be disabled or have missing disability information than those in fresh non-employment spells. Also, the single mothers in left-censored non-employment spells tend to have more children, and their youngest children tend to be younger, compared to those in fresh non-employment spells. These differences reflect the fact that many of those experiencing fresh non-employment spells were in employment at the baseline, while all of those in left-censored spells were, by definition, in non-employment at the baseline.

Panel B shows that 3826 single mothers have left-censored employment spells while 2000 single mothers have fresh employment spells (again the two groups are not mutually exclusive). Single mothers having left-censored employment spells tend to be less disadvantaged than those having fresh employment spells. Specifically, they are older, less likely to be minority group members, more likely to have twelve years of schooling, more likely to have had a previous marriage, less likely to be disabled, and tend to have both fewer children and older children than those in fresh employment spells. Again these differences reflect the fact that many of those experiencing fresh employment spells were in non-employment at the baseline, while all of those in left-censored spells were, by definition, in employment at the baseline.

3.2. Problems Arising from Measuring Spells Using Only the Last Month Observations

As discussed in the introduction, a very common solution to the seam bias problem in the SIPP data is the GHS approach of using only the last month observation from each wave, dropping the three other months in the wave. Two reasons have been given for using the last month data. First, most transitions take place between waves, i.e. in the last month. However, as we discuss below, this reason ignores the fact that almost one-half of completed fresh spells are lost by using only the last month data. Further, we show below that information on the timing of transitions that occur in months other than the last month also is lost, potentially introducing severe distortions to the true employment patterns.

Second, some researchers, e.g. Grogger (2004), have expressed concern that many of the off-seam transitions are simply the result of an imputation procedure by the Census Bureau, making their use suspect. In the SIPP, monthly data are imputed when a sample member either refuses to be interviewed or is unavailable for that interview (and a proxy interview cannot be obtained), or when someone who enters a
sample household after the start of the panel leaves the household during the reference period. The Census Bureau indicates whether a monthly observation is imputed using a variable INTVW, which equals 1 or 2 if a self or proxy interview is obtained (hence the data are not imputed), and 3 or 4 if the respondent refuses to be interviewed or left, respectively (hence the data are imputed). Using this variable, we compare the frequency of transitions at the seam in the imputed and non-imputed data (see Appendix Table A1). We find that approximately 50% of the transitions take place in month 4 in the non-imputed data, but about 81% take place in month 4 in the imputed data. Hence, imputation reduces the fraction of off-seam transitions rather than increasing them.

Under the last month data approach, spells are constructed by acting as if we observed only the last month data for each wave. When there is a status change from the previous month 4 to the current month 4, month 4 of the current wave is coded as the end of a spell. Here we construct three examples representative of our data to illustrate the problems that may arise when adopting this approach. In these examples, which are shown in Figures 1.1 – 1.3, we let \( U, U', E \) and \( E' \) denote a fresh non-employment spell, a left-censored non-employment spell, a fresh employment spell and a left-censored employment spell, respectively. In each figure, the numbers above the line indicate the survey months and the numbers below the line represent the reference period months.

The first example illustrated in Figure 1.1 assumes that a respondent has four spells. The first spell is a left-censored non-employment spell ending in a month 1, the second is a fresh employment spell ending in a month 3, the third is a fresh non-employment spell ending in another month 3, and the last spell is a right-censored fresh employment spell. Using only the last month data, we would treat this respondent’s work history as consisting of a left-censored non-employment spell lasting 32 months and a right-censored employment spell lasting four months. We would lose both a two-month fresh employment spell and a 24-month fresh non-employment spell. In addition, we would miscalculate the spell length of both the left-censored and right-censored spells.

In the next example, shown in Figure 1.2, we keep everything else the same as in Figure 1.1 and only shift the ending point of the second spell to month 2 of a reference period. Now the second fresh employment spell lasts for five months with a month 4 in the middle of the spell. For such a case, using only last month data would not lead to the omission of the second and third spells, only to the miscalculation of the length of all four spells.

Finally, our last example assumes that a respondent has three spells, as shown in Figure 1.3. The first spell is a left-censored non-employment spell ending in month 3 of the first reference period while the

---

8 All adults in sampled households at the start of the panel are considered original sample members and are followed to any new address. Someone entering a sample household after the start of the panel is interviewed as a member of the household, but not followed if he or she leaves. In that case, the remaining months of a reference period after the departure will be imputed, making such observations useless for our study.
second is a completed fresh employment spell, and the third spell is a fresh non-employment spell censored at the end of the sample. Using only last month data, we would record this respondent’s work history as a left-censored employment spell and a fresh non-employment spell. From this example, it is clear that we will lose all left-censored spells less than or equal to three months in length by switching to the last month data. In addition, using the last month data will lead both to miscalculating spell length and to misclassifying spell type for left-censored spells.

To recap, the above three examples show that by using only the last month data, we could lose some spells, miscalculate the length of spells, and misclassify the spell type. The problem is more severe with short spells that are less than four months in duration and that do not cover a month 4. Further, it is clear that in general, using only the last month observations will lead to an overestimate of the length of left-censored spells. However, for fresh spells, using only the last month data may underestimate or overestimate the length of an observed fresh spell, since both the start and end of a fresh spell may be mismeasured.

Of course, the above three examples compare the last month data to the true duration data, while in practice we do not know the true duration of spells. Thus, the relevant comparison is the last month data versus the monthly data contaminated by seam bias. However we note that even in the presence of seam bias, we observe many more short spells in our monthly data than in the last month only data. If some short spells are in fact omitted due to seam bias, switching to using only the last month data will only accentuate this problem.

To shed more light on the comparison between the monthly data and the last month only data, we examine the number of completed spells and the empirical survivor functions for each data type. First, comparing the number of completed spells (which provide the empirical identification for the parameters of the hazard functions), we find that shifting from using monthly data to the last month data results in the loss of about 47% of fresh employment spells, 48% of fresh non-employment spells, 20% of left-censored employment spells, and 18% of left-censored non-employment spells. Our Online Appendix Table A2 presents distributions of spell length and total number of spells of monthly data versus last month data by spell type. However, it is difficult to ascertain the effect of using the last month data on spell lengths from Online Appendix Table A2 because of the right-censored spells. Instead, we investigate this issue using the empirical survivor functions. Online Appendix Figures A.1 and A.2 show that using only the last month data will increase the estimated survivor function for left-censored employment and non-employment spells by a considerable amount. Online Appendix Figures A.3 and A.4 show that this phenomenon is even more pronounced for fresh employment and non-employment spells. Thus shifting from the monthly data to only the last month data leads to both omitting spells and overestimating the spell lengths.

9 An accurate measure of the frequency with which individuals fail to report short spells can be obtained only from matched administrative-survey data, which are not available for this sample.

10 Our Online Appendix is found at http://dept.econ.yorku.ca/~xli/HomePage_files/appendix_seam_bias_hls.pdf.
4. Econometric Model

4.1 Duration Models

In this section, we describe our multi-spell multi-state duration models without the complication caused by the seam bias problem, which will be incorporated in Section 5. We consider four types of spells and as in Section 3, we let $U'$ and $U$ represent left-censored and fresh non-employment spells respectively, and let $E'$ and $E$ represent left-censored and fresh employment spells respectively. For each spell, we assume that the hazard function of individual $i$ takes the following form:

$$
\lambda_{ik}(t | X_i(\tau + t), \theta_{ik}) = \frac{1}{1 + \exp\left\{ -h_k(t) - X_i(\tau + t)\beta_k - \theta_{ik} \right\}}, \quad k = U, U', E, E' \quad (4.1)
$$

where $t$ denotes current spell duration, $h_k(t)$ denotes duration dependence, $\tau$ denotes the calendar time of the start of the spell, $X_i(\tau + t)$ denotes a (possibly) time changing explanatory variable, and $\theta_{ik}$ denotes unobserved heterogeneity. We follow common practice and specify the unobserved heterogeneity corresponding to these four types of spells through a vector $\theta = (\theta_U, \theta_{U'}, \theta_E, \theta_{E'})$, and assume that $\theta$ is distributed independently across individuals and is fixed across spells for a given individual. Following McCall’s (1996) multivariate generalization of the Heckman and Singer (1984b) approach, we let $\theta$ follow a discrete distribution with points of support $\theta_1, \theta_2, ..., \theta_J$ (where, e.g., $\theta_1 = (\theta_{U1}, \theta_{U'1}, \theta_{E1}, \theta_{E'1})$) and associated probabilities $p_1, p_2, ..., p_J$ respectively, where $p_J = 1 - \sum_{j=1}^{J-1} p_j$.

We illustrate our multi-spell multi-state model using a hypothetical work history in Figure 2. Again, in Figure 2 the numbers above the line are the survey months and the numbers below the line are reference period months. At this point, we focus only on the observed history in the top panel; the other three scenarios below related to seam bias correction will be discussed in Section 5. Here the individual starts in a left-censored non-employment spell, which she exits 5 months after the beginning of the sample. She then has a fresh employment spell of 7 months, and finally a fresh non-employment spell, which is censored after 24 months. After omitting the individual index, the contribution to the likelihood function is

$$
L = \sum_{j=1}^{J} p_j \left[ \lambda_{U'}(5 | \theta_{U'j}) \prod_{r=1}^{4} \left( 1 - \lambda_U(r | \theta_{U'j}) \right) \right] \left[ \lambda_E(7 | \theta_{Ej}) \prod_{r=1}^{6} \left( 1 - \lambda_E(r | \theta_{Ej}) \right) \right] \left[ \prod_{r=1}^{24} \left( 1 - \lambda_U(r | \theta_{Uj}) \right) \right]. \quad (4.2)
$$

Note that $\tau$ will be individual specific, but we omit its subscript for ease of exposition.
4.2 Expected Durations and the Effect of Changing an Explanatory Variable on the Expected Durations

One of our two primary counterfactual measures is the effect of changing an independent variable on the expected duration of each type of spell, as this is a parameter with natural policy interpretation and comparable across studies using different levels of time aggregation (including continuous time). To begin, we note that, for example, conditional on the unobserved heterogeneity $\theta_E$ the probability that a fresh employment spell lasts $t$ months is given by the density function

$$f_E(t|\theta_E) = \lambda_E(t|\theta_E) \prod_{r=1}^{t-1} (1 - \lambda_E(r|\theta_E)).$$  \hspace{1cm} (4.3)

The expected duration for a fresh employment spell is therefore

$$E_D = \sum_{j=1}^{J} \sum_{t=1}^{\infty} t \cdot f_E(t|\theta_E) \cdot p_j.$$  \hspace{1cm} (4.4)

Since there is no guarantee that this expected duration will be finite, we instead calculate a truncated expected duration as follows:

$$E_D^{trunc} = \sum_{j=1}^{J} \sum_{t=1}^{T^*} t \cdot f_E(t|\theta_E) + S(T^*|\theta_E) \cdot T^* \cdot p_j.$$  \hspace{1cm} (4.5)

We choose $T^* = 60$; obviously $E_D^{trunc}$ is a lower bound of $E_D$.\textsuperscript{12} We can calculate (4.5) for each individual for any given level of the explanatory variables, and then take the sample average for each type of spell.\textsuperscript{13} To conduct the counterfactual analysis, we set an explanatory variable at two different levels while setting the other variables to their actual values; the difference between the two expected durations represents the effect of the respective change in that particular explanatory variable. Since the expected duration is a differentiable function of the parameters, with non-zero, bounded derivatives, we use the delta method to obtain confidence intervals for our estimate of the effect of changing the explanatory variables on the expected durations.

4.3 Simulating the Effects of Changing an Explanatory Variable on the Fraction of Time Spent Employed

Our second primary counterfactual measure is the effect of changing an independent variable on the fraction of time that an individual is employed, since this combines in a sensible way the effect of changing an explanatory variable on all four hazard rates. We use simulations to predict the fraction of time spent in

\textsuperscript{12} The longest panel in our data lasts 40 months.

\textsuperscript{13} To investigate whether the out-of-sample durations are having a disproportionate impact on estimated expected duration, we also follow Eberwein, Ham and LaLonde (2002) and i) freeze the hazard function for durations longer than 15 months at 15 months for fresh spells and ii) freeze the hazard function for durations longer than 25 months at 25 months for left-censored spells. We find that this does not affect our estimated effects.
employment over several time spans, e.g. 3-year (short run), 6-year (medium run) and 10-year (long run) horizons, as well as to examine how these fractions change with changes in the macroeconomic, demographic and public policy variables. For each individual in our sample, we first simulate an employment/non-employment history over a particular time horizon and then calculate the sample fraction of time employed based on the individual simulated histories. Note that the simulated employment histories and fractions depend on the parameter estimates for all four hazard functions. To construct the counterfactual for an explanatory variable we simulate the model for two different levels of an explanatory variable while keeping the other variables at their actual values; the difference between the two expected fractions of time employed represents the effect of changing that particular explanatory variable. Since the simulated fractions are discontinuous functions of the parameter estimates, the delta method cannot be used to calculate consistent estimates of the relevant confidence intervals. For such a case, researchers often form \((1 - \alpha)\) confidence intervals by taking a large number of draws from the asymptotic distribution of the parameter estimates, simulating the model using each draw of the parameter estimates, and then trimming the bottom and top \(0.5\alpha\) of the function values (the estimated employment fractions for our case). We refer to this approach as the AD-bootstrap. However, Woutersen and Ham (2013) show that the AD-bootstrap is generally not valid and produces confidence intervals that are too small. Instead Woutersen and Ham (2013) propose an alternative approach, the CI-bootstrap, which is appropriate in this setting. They show that consistent 95% confidence sets for these functions can be obtained by taking M draws from the asymptotic distribution of the parameter estimates, and calculating the function for only the draws within the 95% confidence interval. We describe our use of CI-bootstrap approach in more detail in Online Appendix 3.

5. Correcting for Seam Bias

In this section we outline a parametric approach to correct for seam bias which is consistent with misreporting arising due to telescoping, the phenomenon observed in the statistics literature whereby individuals report changes as occurring further in the past than actually happened (Sudman and Bradburn 1974, p.69). We first describe the behavioral assumptions that we rely on to identify the seam bias correction parameters. Then we discuss using our seam bias correction in the context of a single employment spell as a platform for investigating the identification issues. We show that we can identify seam bias parameters in our model for the fresh employment spells. However, we cannot identify separate misreporting parameters for the left-censored employment spells unless we restrict the duration dependence in such spells. This identification problem has not arisen in previous empirical studies considering misclassification because our econometric model in the absence of seam bias is much richer than those in previous studies considering misclassification (a multi-state duration model vs. a dynamic probit model.) To identify our model, we thus assume that fresh and left-censored employment (non-employment) spells share the same misreporting
parameters and show that this assumption seems reasonably compatible with the data. Since our model becomes significantly over-identified given this assumption, we are able to consider several extensions of our misreporting model. Finally, we consider several quite different models of misreporting, but reject them as being inconsistent with observed patterns in aggregated SIPP data.

5.1 Notation and Behavioral Assumptions

We first set up our notation before discussing our assumptions. Let $M^\text{obs} = (m, l)$ represent a spell reported to end in month $m$ of reference period (wave) $l$.\(^{14}\) (Note that $M^\text{obs} = (m, l)$ could be contaminated by seam bias.) Further, let $M^\text{true} = (m, l)$ represent the case where a spell truly ended in month $m$ of reference period $l$. The variable $m$ in $M^\text{obs}$ and $M^\text{true}$ can take five possible values: 1, 2, 3, 4, or 0, where $m = 1, 2, 3$ or 4 indicates a transition in months 1, 2, 3 or 4, respectively, and $m = 0$ indicates a right-censored spell ending with the survey. For our sample $l$ takes on the values from 1 to $L$, where $L$ is the number of waves, which depends on the specific SIPP panel being used.\(^{15}\) For example, $M^\text{obs} = (4, 4)$ indicates that a transition was reported to have occurred in month 4 of reference period 4, while $M^\text{true} = (3, 5)$ denotes a different transition actually occurred in month 3 of reference period 5.

In the SIPP a respondent is first asked whether she had a job or business at any time during the previous four-month period; if the answer is yes, the respondent is then asked whether she had a job or business during all weeks of the period. Further questions are asked of individuals who report some time employed and some time not employed to determine the timing of their periods of employment and non-employment. For example, suppose an individual continues a spell of non-employment into a given wave and does not have a job for months 1 and 2 of the new wave, but she gets a job in month 3 that continues into month 4 of the wave. Given the interview structure, under telescoping this individual may report that she has a job for the whole wave based on the fact she has a job in month 4, which is the month closest to (right before) the interview month. For this particular example, a non-employment spell ending in month 2 of the current reference period will be reported to end in month 4 of the previous reference period. Goudreau, Oberheu and Vaughan (1984) document this as the most common type of misreporting behavior for AFDC benefit receipt in the Income Survey Development Program.\(^{16}\)

To model the misreporting behavior, we make the following four assumptions:

\(^{14}\) The end of a spell can occur either because a transition took place or because the individual reached the end of the sample period.
\(^{15}\) There are seven waves in the 1986 and 1987 panels, six in the 1988 panel, eight in the 1990 and 1991 panels, ten in the 1992 panel, and nine in the 1993 panel.
\(^{16}\) Their study is conducted by comparing respondents’ reports obtained from interviews with administrative record information.
A1) the respondents report all transitions that occurred during reference period \( l \) as occurring either in the true month or month 4 of reference period \( l-1 \);

A2) a respondent reports a transition that actually occurred in months 1, 2 or 3 of reference period \( l \) as taking place in month 4 of reference period \( l-1 \) with some pre-specified (but unknown) probabilities \( \gamma_1 \), \( \gamma_2 \) and \( \gamma_3 \) respectively;

A3) if a transition truly happened in month 4 of a reference period, the respondent reports it as occurring in that month; and

A4) the true transition rate for a given duration does not depend on the reference month, \( m \).

Given these behavioral assumptions, we have the following conditional probabilities:

\[
P(M_{\text{obs}} = (m, l) \mid M_{\text{true}} \neq (m, l)) = 0, \ m = 1, 2, 3; \tag{5.1}
\]

\[
P(M_{\text{obs}} = (4, l-1) \mid M_{\text{true}} = (m, l)) = \gamma_m, \ m = 1, 2, 3; \tag{5.2}
\]

\[
P(M_{\text{obs}} = (4, l) \mid M_{\text{true}} = (4, l)) = 1 \quad \text{and} \tag{5.3}
\]

\[
P(M_{\text{obs}} = (0, l) \mid M_{\text{true}} = (0, l)) = 1. \tag{5.4}
\]

5.2 Correcting for Seam Bias in a Single Employment Spell Setting

To illustrate the method in the simplest way, we first explore the problem involving a single spell of employment. Based on our behavioral assumptions, it is straightforward to derive the likelihood function given that we observe \( M_{\text{obs}} = (m, l) \), and the reported length of the spell, \( \text{dur}_{\text{obs}} \), both of which potentially have been contaminated by seam bias. Let \( L_i(K) \) denote the contribution to the likelihood function for a complete spell of \( K \) months without considering seam bias:

\[
L_i(K) = \sum_{j=1}^{J} p_j \left[ \lambda_i(K \mid \theta_j) \prod_{t=1}^{K-1} (1 - \lambda_i(t \mid \theta_j)) \right],
\]

where \( \lambda_i, \theta, p \) are the hazard function, unobserved heterogeneity support point and associated probability parameters for a single spell, and their multiple spell counterparts are defined in Section 4.1. For notational simplicity, we drop the individual subscript \( i \) in the following discussion.

---

17 This assumption is the consequence of the survey design. For ease of interviewing, the entire sample is randomly split into four rotation groups, and one rotation group (one-fourth of the sample) is interviewed each calendar month. Each rotation group in a SIPP panel is interviewed once every four months about employment and program participation during the previous four months.

18 Our assumptions rule out the possibility that individuals forget about very short spells that fall between two interviews. However, without administrative data linked to the SIPP data, we have no way of verifying the validity of this assumption.

19 Here we assume that seam bias affects only the end date, and not the start date, of a spell. We relax this assumption when we consider multiple spell data.
Taking seam bias into consideration, the likelihood contribution for a complete spell of \( K \) months that is reported to end in months 1, 2, 3, or 4 of reference period \( l \) is given by \( ^{20} \)

\[
P\left( M^{\text{obs}} = (1, l), \text{dur}^{\text{obs}} = K \right) = (1 - \gamma_1) L(K); \tag{5.5}
\]

\[
P\left( M^{\text{obs}} = (2, l), \text{dur}^{\text{obs}} = K \right) = (1 - \gamma_2) L(K); \tag{5.6}
\]

\[
P\left( M^{\text{obs}} = (3, l), \text{dur}^{\text{obs}} = K \right) = (1 - \gamma_3) L(K) \quad \text{and} \tag{5.7}
\]

\[
P\left( M^{\text{obs}} = (4, l), \text{dur}^{\text{obs}} = K \right) = \gamma_1 L(K + 1) + \gamma_2 L(K + 2) + \gamma_3 L(K + 3) + L(K), \tag{5.8}
\]

where the \( \gamma \)'s are the seam bias parameters defined by equation (5.2). A natural question to ask is whether the model is identified without restricting the form of the duration dependence, and we now turn to this issue.

### 5.2.1 Identification with Fresh Spell Data

At first glance, it may appear that we have to restrict the form of the duration dependence to identify our model. However, this is not the case, at least for fresh spells. For simplicity we consider a model for employment spells with duration dependence but no explanatory variables and no unobserved heterogeneity (The argument for non-employment spells is identical.). One could estimate the parameters of this simplified model using the analog principle (Manski 1994) by comparing sample moments and their probability limits, which we will refer to as population moments.

The data provide us with the empirical hazards for spells ending at duration \( t \) in reference month \( m \), where \( m = 1, 2, 3, 4 \), and these are our sample moments; let \( p_m(t) \) denote their population counterparts. To obtain analytical expressions for these population moments, we first assume that in the population there are \( N_t \), \( N_{t-1} \), \( N_{t-2} \), and \( N_{t-3} \) individuals having current durations equal to \( t \), \( t-1 \), \( t-2 \), and \( t-3 \), respectively in month 4 over all reference periods. (\( N_t \), \( N_{t-1} \), \( N_{t-2} \), and \( N_{t-3} \) may be thought of as being large but finite for now, as they will drop out of the population moments below.) Using the terminology of the duration literature, for \( t \geq 4 \), \( N_t \), \( N_{t-1} \), \( N_{t-2} \), and \( N_{t-3} \) represent the total number of individuals at risk at durations \( t \), \( t-1 \), \( t-2 \), and \( t-3 \) respectively in month 4. Let \( \lambda(t) \) denote the hazard function at spell duration \( t \).

As discussed in Section 5.1, we assume that some of the transitions in months 1, 2, and 3 of the following reference period are heaped into the current month 4, so \( N_t \), \( N_{t-1} \), \( N_{t-2} \), and \( N_{t-3} \) are not contaminated by seam bias. For month 1, the number of individuals who actually enter month 1 at duration \( t \)

\[ ^{20} \text{This is shown in Appendix 1 of this paper.} \]

\[ ^{21} \text{This derivation requires that for each } t \geq 4, N_t > 0 \text{ and } \lambda(t) < 1. \]
is \( N_{t-1} \cdot \left[ 1 - \lambda(t-1) \right] \). This is the number of individuals at risk at duration \( t \) in month 1, i.e. the difference between the number entering the previous month 4 at duration \( t-1 \) and the number who leave in the previous month 4 at duration \( t-1 \). However, note that the number observed to be at risk consists of those actually at risk minus the sum of:

1. The number of individuals who actually left in month 1 at duration \( t \) but were observed to exit in the previous month 4 due to seam bias, \( N_{t-1} \cdot \left[ 1 - \lambda(t-1) \right] \cdot \lambda(t) \cdot \gamma_1 \); and

2. The number of individuals who actually left in month 2 at duration \( t+1 \) but were observed to exit in the previous month 4 due to seam bias, \( N_{t-1} \cdot \left[ 1 - \lambda(t-1) \right] \cdot \lambda(t) \cdot \lambda(t+1) \cdot \gamma_2 \); and

3. The number of individuals who actually left in month 3 at duration \( t+2 \) but were observed to exit in the previous month 4 due to seam bias, \( N_{t-1} \cdot \left[ 1 - \lambda(t-1) \right] \cdot \lambda(t) \cdot \lambda(t+1) \cdot \lambda(t+2) \cdot \gamma_3 \),

where \( \gamma_1, \gamma_2 \) and \( \gamma_3 \) are the seam bias parameters defined by equation (5.2). Of those actually at risk in month 1, a fraction \( \lambda(t) \) actually leave, but only a fraction \( \lambda(t) \cdot (1 - \gamma_1) \) are observed to have left in month 1, and the rest \( \lambda(t) \cdot \gamma_1 \) were reported to have left in the previous month 4 due to seam bias. Thus, the number observed leaving equals \( N_{t-1} \cdot \left[ 1 - \lambda(t-1) \right] \cdot \lambda(t) \cdot (1 - \gamma_1) \). Given this, we can easily formulate the first population moment condition for the fraction observed to leave in month 1 at duration \( t \) (after deleting the common factor \( N_{t-1} \cdot \left[ 1 - \lambda(t-1) \right] \) from both the numerator and denominator) as

\[
p_1(t) = \frac{\lambda(t) \cdot (1 - \gamma_1)}{1 - \lambda(t) \cdot \gamma_1 - \left[ 1 - \lambda(t) \right] \cdot \lambda(t+1) \cdot \gamma_2 - \left[ 1 - \lambda(t) \right] \cdot \lambda(t+1) \cdot \lambda(t+2) \cdot \gamma_3}.
\] (5.9)

For month 2 and month 3, the population moment conditions are

\[
p_2(t) = \frac{\lambda(t) \cdot [1 - \gamma_2]}{1 - \lambda(t) \cdot \gamma_2 - \lambda(t+1) \cdot \gamma_3},
\] (5.10)

\[
p_3(t) = \frac{\lambda(t) \cdot [1 - \gamma_3]}{1 - \lambda(t) \cdot \gamma_3}.
\] (5.11)

The derivations of (5.10) and (5.11) are very similar to those of (5.9), and thus are included in the Online Appendix 1.

For month 4, under our assumptions the number of individuals who actually (and were observed to) enter month 4 at duration \( t \) is \( N_t \). The number of individuals who actually leave in month 4 at duration \( t \) equals \( N_t \cdot \lambda(t) \). However, due to seam bias, we also observe leaving in month 4:
1. $N_i \cdot [1 - \lambda(t)] \cdot \lambda(t+1) \cdot \gamma_1$ individuals who actually left in month 1 of the next reference period but reported leaving in month 4,

2. $N_i \cdot [1 - \lambda(t)] \cdot [1 - \lambda(t+1)] \cdot \lambda(t+2) \cdot \gamma_2$ individuals who actually left in month 2 of the next reference period but reported leaving in month 4 and

3. $N_i \cdot [1 - \lambda(t)] \cdot [1 - \lambda(t+1)] \cdot [1 - \lambda(t+2)] \cdot \lambda(t+3) \cdot \gamma_3$ individuals who actually left in month 3 of the next reference period but reported leaving in month 4.

The number of individuals observed leaving in month 4 equals the sum of $N_i \cdot \lambda(t)$ and the three terms above. Thus, after canceling out $N_i$ from both the numerator and denominator, we have

$$p_4(t) = \lambda(t) + [1 - \lambda(t)] \cdot \lambda(t+1) \cdot \gamma_1 + [1 - \lambda(t)] \cdot [1 - \lambda(t+1)] \cdot \lambda(t+2) \cdot \gamma_2 + [1 - \lambda(t)] \cdot [1 - \lambda(t+1)] \cdot [1 - \lambda(t+2)] \cdot \lambda(t+3) \cdot \gamma_3.$$ (5.12)

If we equate population and sample moments, we have only four equations in seven unknowns, $\lambda(t), \lambda(t+1), \lambda(t+2), \lambda(t+3), \gamma_1, \gamma_2$, and $\gamma_3$. However, we can obtain additional moment conditions by considering, for example, $p_1(t+1), p_2(t+2), p_3(t+3), p_4(t+1)$, and $p_5(t+1)$. Since we can repeat this and add many more moments without introducing new parameters, the model is significantly over-identified.\(^{22}\) Of course, the number of moments being greater than or equal to the number of unknowns is only a necessary, but not sufficient, condition for identification. To pursue this issue further, for several exactly identified models, where the number of moment conditions equals the number of unknowns, we used several reasonable sets of values for the empirical moments and let Maple solve for the parameters. For each set of empirical moments, we found only one set of real solutions for the parameters when we restricted them to the unit interval.\(^{23}\)

### 5.2.2 Non-Identification of Left-censored Spells’ Misreporting Parameters

As noted in Section 3, we allow left-censored spells to have different hazard rates than the fresh spells. Here we show that without any auxiliary information, the parameters for the left-censored spells are not identified unless we restrict the duration dependence in such spells, for example, assuming no duration dependence. Since we measure the duration of these spells from the start of the sample, we will only observe a spell of duration 1, 5, 9, 13… ending in month 1 of the reference period, a spell of length 2, 6, 10, 14… ending in month 2, a spell of length 3, 7, 11, 15… ending in month 3, and a spell of length 4, 8, 12, 16… ending in

\(^{22}\) Here we are abstracting from an endpoint issue. In the last month of the sample, which will be month 4, we would not expect any misreported transitions, since there are no future spells to misreport from. Thus, we assume no misreporting in the last month of the sample, and this also will aid in identification.

\(^{23}\) The same arguments can be used if we add time constant observed explanatory variables to the model since one can always stratify the data by demographic group.
month 4. Let $\lambda^k(t)$ represent hazard function and $\gamma^k_k$, $k=1,2,3$ represent seam bias correction parameters for the left-censored spells. For $t=4, 5, 6, 7$, we only have four moments in seven unknowns $(\lambda^k(4), \lambda^k(5), \lambda^k(6), \lambda^k(7)$ and the three seam bias correction parameters). Unfortunately, if we add four more population conditions for $t=8, 9, 10, 11$, we also will add four more unknowns $\lambda^k(8), \lambda^k(9), \lambda^k(10), \lambda^k(11)$. Thus, we now have 8 moments for 11 unknowns, and the identification problem remains.  

If, however, we restrict the duration dependence by, for example, assuming no duration dependence, $\lambda^k(t)$ reduces to the duration constant hazard $\lambda^k$ and the misreporting parameters are now identified in the left-censored spells. For example for $t=4, 5, 6, 7$, we now have four equations and four unknowns. If we add more population conditions, for example, $t=8, 9, 10, 11$, we will have more equations and no additional unknowns. Thus, assuming no duration dependence, the left-censored spell hazard functions and misreporting parameters are over-identified.

But assuming no duration dependence in left-censored spells runs contrary to most empirical evidence on such spells for disadvantaged women, so we must find a more reasonable identification assumption. Our data appear to offer support for the alternative identifying assumption that the fresh and the left-censored spells of the same type share the same misreporting parameters, i.e. $\gamma_k = \gamma^k$, for $k=1, 2, 3$.

In Appendix Table A2 we present the distribution of the transitions across the four interview months for each type of spell. It shows that the distributions for the left-censored and fresh employment spells are similar, and so are the distributions for the left-censored and fresh non-employment spells. For example, 52% and 53% of the transitions out of left-censored and fresh employment spells respectively are reported to occur in month 4, while the relevant statistics for the left-censored and fresh non-employment spells are 36% and 42% respectively.

Investigating identification theoretically with unobserved heterogeneity, seam bias and duration dependence is beyond the scope of this paper. However we note two points here. First, because we have time changing explanatory variables (that are not a function of individual choices), our discrete time model in the absence of seam bias satisfies the sufficient condition for global identification of a discrete time model with duration dependence and unobserved heterogeneity – see Hausman and Woutersen (2013). Second, we had no trouble inverting the Hessian for any of our estimated models, indicating that we at least have local

\[ Note that this identification problem would disappear if we had information on the actual start date of the left-censored spells since the argument used for the fresh spells would now be applicable to the left-censored spells. 

\[ Of course, the absence of duration dependence in the left-censored spells is a sufficient condition for identification, not a necessary one. However, we did not attempt to find weaker conditions on the duration dependence in the left-censored spells that would be both necessary and sufficient, because we believe it will be much harder to judge the reasonableness of such conditions than the reasonableness of i) the assumption of no duration dependence in the left-censored spells or ii) the assumption (described next) of equality of misreporting parameters for left-censored and fresh spells of the same type. \]
identification with unobserved heterogeneity, misreporting, and duration dependence. We suspect one reason for this is that, given our assumptions, for some observations (spells ending in reference months 1, 2 or 3) we know the exact length of the spell.

5.3 Correcting for Seam Bias in a Multiple-Spell Model

In a multiple-spell duration model, correcting for seam bias complicates the likelihood function dramatically, because adjusting a response in one spell involves shifting not only the end of that spell but also the start of the subsequent spell. As a practical matter, this is a serious complication for us as respondents in our sample have up to seven spells, and a respondent can have several spells ending in month 4 in her history.

We demonstrate how to incorporate seam bias correction into the likelihood function by revisiting the example of Figure 2 in Section 4. This example covers all essential problems for estimating multi-spell, multi-state duration models with seam bias. Recall that the respondent reports a three-spell history: the first spell is a left-censored non-employment spell ending in month 1 of reference period 2, the second is a fresh employment spell reported to end in month 4 of reference period 3, and the third is a fresh non-employment spell, which is censored at the end of the sample. Since the first spell ends in month 1 of the reference period, we assume it is reported correctly. However, since the second spell is reported to have ended in month 4 of the third reference period, the reported history could be true, but there are also three additional possible histories A, B, and C, that are illustrated in Figure 2. Specifically, the second spell could actually have ended in month 1 of the following reference period, which implies that we would need to reduce the duration of the subsequent (censored) spell by one month. Alternatively, it also could have ended in month 2 or 3 of the following reference period, in which case we would need to shorten the length of the subsequent spell by two or three months respectively.

We denote the misreporting parameters for employment and non-employment spells by \((\gamma^E_1, \gamma^E_2, \gamma^E_3)\) and \((\gamma^U_1, \gamma^U_2, \gamma^U_3)\) respectively. With seam bias the contribution to the likelihood function for the reported history in Figure 2 is no longer given by (4.5). Instead it is given by
Derivation of (5.13) is shown in our Online Appendix 2.

We also consider extensions of our model. First we allow for a constant probability $P_A$ that an individual never misreports. If the individual never misreports she will have the standard multi-state, multi-spell likelihood function $L^A(\cdot)$, see e.g. Flinn and Heckman (1982, 1983). If she misreports at least some of the time, her contribution to the likelihood would be the appropriate seam bias likelihood function $L^B(\cdot)$, such as (5.13) above. Since we do not know whether an individual misreports, a representative individual’s contribution to the likelihood function is

$$L(\cdot) = P_A L^A(\cdot) + (1 - P_A) L^B(\cdot).$$

Maximizing the log likelihood based on (5.14) involves estimating only one additional parameter, $P_A$.26

Second, for the models described by (5.13) and (5.14) we consider two alternatives: i) the misreporting probabilities ($\gamma$’s and $P_A$) are constant for the whole sample and ii) the misreporting probabilities vary by individual demographics. For example, suppose the employment spell misreporting probabilities differ among Whites and non-Whites (African Americans and Hispanics) as follows

$$\gamma^E_m = \frac{1}{1 + \exp\left(-\alpha^E_{0m} - \alpha^E_{1m} NW\right)} \quad m = 1, 2, 3. \quad (5.15)$$

where $\alpha^E_{1m}$ allows ethnicity to affect the employment misreporting for month $m$. In terms of identification, recall that the model is overidentified once we restrict the misreporting parameters to be the same in fresh and left-censored spells of the same type, and this overidentification allows us to extend the model by letting ethnicity affect the misreporting parameters.

---

26 Care must be exercised if one wants to test the null hypothesis $P_A = 1$, both because it is on the boundary of the parameter space and because the misreporting parameters, the $\gamma$’s, are not identified under this null hypothesis (Davies 1987). Fortunately, the estimate of $P_A$ is quite far from 1 in our case.
5.4 Alternative Misclassification Schemes

Of course, there is the possibility that the transitions are misclassified in a way that differs from our approach described above. Here we consider several other possibilities, which we argue can be rejected based on aggregates of the micro data. The first possibility we consider is that some of the month 1 transitions are pushed into month 2, some of the month 2 transitions are pushed into month 3, and some of the month 3 transitions are pushed into month 4, but none of the month 4 transitions are pushed into the next reference period (because it is the last month in the reference period). If 50% of the transitions in months 1, 2 and 3 are pushed to the next month, then we would see 12.5% of the transitions in month 1, 25% in month 2, 25% in month 3, and 37.5% in month 4. Alternatively, suppose 75% of the transitions get pushed out of months 1, 2, and 3. Then we would see 6.25% of the transitions in month 1, 25% in month 2, 25% in month 3, and 43.75% in month 4. In either case, the observed pattern should be: about 25% of the observed transitions reported in months 2 and 3, a smaller proportion of transactions reported in month 1, and a larger proportion in month 4. However, in our data, months 1, 2, 3, and 4 have 16.57%, 19.08%, 18.49%, and 45.86% of the employment/non-employment transitions respectively, which is clearly inconsistent with this alternative model (note that our proposed model is consistent with this pattern.)

Next, we consider the possibility that some of the transitions in month 1 of the reference period $l+1$ are pushed back into month 4 of reference period $l$. (Recall that month 1 of the reference period $l+1$ is the interview month for reference period $l$.) If this is the only source of misclassification, then the pattern should again be: about 25% of the observed transitions reported in months 2 and 3, a smaller proportion of transactions reported in month 1, and a larger proportion in month 4. This model would also be rejected by the summary statistics presented in the previous paragraph.

Third, we explore a scenario that individuals may forget about a fraction of very short spells starting in reference period months 1, 2 and 3. In other words, the number of transitions in month 4 is accurately reported, but the numbers in months 1, 2 and 3 are under-reported. We cannot investigate this possibility with the aggregate pattern of employment transitions, but we can examine this explanation in another way. We know from administrative data that short spells are much more frequent in employment durations than in welfare durations for the disadvantaged single mothers that we study. Thus, if we compare the transitions in months 1, 2, 3 and 4 for employment duration data and welfare duration data, we would expect to see a larger fraction of transitions in month 4 for the employment data if this explanation is correct. However, we find that 52.7% of all transitions out of employment were reported to have occurred in month 4, while 62.7% of all transitions out of welfare were reported to have occurred in month 4, casting doubt on this last explanation.

One can also consider a mixture of our misreporting model and one of the models suggested above; for example we considered a scenario where some of the transitions are misreported according to the first alternative scheme above and others according to our scheme. If the fraction following the alternative scheme...
is small, it will of course be hard to reject this mixed model, since the data will look very much like that which would be generated by our assumed misreporting scheme. However, if the fraction of those following the alternative scheme is substantial, with this mixed scheme we would expect fewer transitions observed in months 1, 2, and 3 than month 4, with month 1 having the least number of observed transitions observed. Interestingly, our data aggregated across all panels show this pattern. However, when we take the distribution of the observed transitions by interview month for each of the 1986-1993 panels of the SIPP, we only find this pattern in 2 out of the 7 individual panels.27

6. Empirical Results

As noted above, we estimated multi-state, multi-spell duration models when: i) the misreporting probabilities are constant across individuals (constant correction model hereafter); ii) the misreporting probabilities vary across individuals by demographic characteristics (variable correction model hereafter); iii) allowing for a constant probability that an individual never misreports when the misreporting probabilities are constant across individuals and iv) letting the misreporting probabilities and the probability that an individual never misreports depend on her characteristics. All models are estimated with duration dependence, time-changing explanatory variables, and unobserved heterogeneity. For each model estimated we choose the best fitting polynomials (in logarithms) for duration dependence, and the number of support points for the unobserved heterogeneity (as specified in Section 4.1), using the Schwartz criterion. In addition to the minimum wage, we use a relatively standard mix of policy, demographic and demand variables (a subset of these variables are time-changing).28 Since there are a very large number of estimated hazard coefficients for the models described above, we present them in our Online Appendix.

In this section we discuss our estimated misreporting parameters for all of the models described above. We then discuss the effect of changing an explanatory variable on the expected durations. Last we report results from simulations of the effects of changing an explanatory variable on the fraction of time spent employed.

6.1 Estimating the Misreporting Probabilities and Other Parameters

Table 2 presents the estimates from our constant and variable correction models. All of the misreporting probabilities are statistically and economically significant. In general, the misreporting probabilities are larger for the employment spells than for the non-employment spells. Another interesting pattern is that probabilities of misreporting are descending from month 1 through month 3. This pattern indicates that, among all transitions occurring in months 1, 2 and 3, the longer the time distance between the

27 See the Table in Online Appendix 3.
28 When researchers use variables like state unemployment rates, they often fix the standard errors by clustering within a state. Unfortunately their parameter estimates will not be consistent if spells are correlated across individuals.
transition and the interview, the more likely it is that the respondent heaps that transition into the previous month 4. (Recall that interviews are conducted in month 1 of the following reference period; thus, month 1 transitions occurred furthest from the interview time). According to the constant correction model, 56.8%, 45.6% and 44.1% of month 1, 2, and 3 transitions out of non-employment spells, respectively, have been shifted to month 4, while 72.9%, 60.1%, and 58.5% of month 1, 2, and 3 transitions out of employment spells, respectively, have been shifted to month 4. In the variable correction model, the misreporting probabilities are specified in the form of equation (5.15). We find that race is the only demographic variable that significantly affects misreporting behavior. African Americans and Hispanics are more likely to misreport by about 7 to 8 percentage points for non-employment spells and by about 7 to 10 percentage points for employment spells.

Next we consider the model where a fraction $P_A$ of the sample never misreport. Our estimate of $P_A$ is 15.5% with a 95% confidence interval of approximately (9.6%, 21.4%). When we allow both $P_A$ and the six misreporting parameters to vary with education and race, neither education nor race significantly affects $P_A$ or the misreporting parameters.

We also find that the estimated hazard coefficients and standard errors are remarkably similar across our misreporting approaches, so below we focus on results based on the estimates from the constant correction model. Thus, for the sole purpose of correcting for seam bias, our results suggest that a constant correction model is sufficient. However, if a researcher is also interested in misreporting behavior per se, increasing the model complexity provides valuable information about misreporting behavior.

### 6.2 The Effects of Changing an Explanatory Variable on Expected Durations

This section presents the estimated expected durations and the effects of changing an explanatory variable on expected durations from our constant correction model. In addition, for comparison we show the results from the two ad hoc approaches to dealing with seam bias discussed above: i) adding a month 4 dummy variable to each of the hazard functions (the last month dummy model or LDUM hereafter) and ii) adopting the GHS approach of using the month 4 data only (the last month data model or LDAT hereafter). When we calculate the expected duration for the LDAT model, we have taken into consideration that this model uses a four-month hazard (compared with the other models’ monthly hazards), while when we calculate the expected duration for LDUM model, we add one-quarter of the last month dummy coefficient to each support point of unobserved heterogeneity.

---

29 When we calculate the expected duration for the LDAT model, we have taken into consideration that this model uses a four-month hazard (compared with the other models’ monthly hazards), while when we calculate the expected duration for LDUM model, we add one-quarter of the last month dummy coefficient to each support point of unobserved heterogeneity.
expected durations and their respective standard errors. Fresh spells have expected durations of roughly a year, and left-censored spells have longer expected durations than fresh spells, with left-censored employment spells being somewhat longer (at 42.2 months) than left-censored non-employment spells (39.3 months). The expected durations are quite precisely estimated, with confidence intervals roughly two months in size.

Our results for the welfare policy variables indicate that 10 percent higher maximum welfare benefits significantly lengthen left-censored non-employment spells by half a month and shorten fresh employment spells by a fifth of a month. The “carrot” welfare reform policies (increased earnings disregards) had a significant effect on the length of fresh non-employment spells in the direction that theory would predict, shortening them by roughly three months. Interestingly, we find no statistically significant effect of these earnings disregards on the expected duration of left-censored non-employment spells, nor were there statistically significant effects on either type of employment spell. Similarly, the “stick” type of welfare reform policies (time limits, work requirements, family cap, and a reduction in exemptions for the JOBS program or an increase in JOBS sanctions) had no statistically significant effect on any of the four types of spells—the coefficients are fairly sizeable but the standard errors are large as well.

By making work more attractive relative to welfare, increasing minimum wages would be predicted to induce workers to shorten non-employment spells and lengthen employment spells. On the other hand, in response to an increase in the minimum wage firms may lay off workers if their productivity is below the new minimum wage, shortening employment duration, and may set higher standards for hiring, lengthening unemployment spells. We find statistically significant evidence of the first (supply) effect, but only for fresh employment spells, which are predicted to be a month and a half longer when the minimum wage is increased ten percent.

Finally, our results show the importance of macroeconomic conditions to employment and non-employment for less-educated single mothers. Three of the four types of spells show significant effects of a 25 percent increase in the unemployment rate (e.g. from 6% to 7.5%) in the direction that theory would suggest, with left-censored non-employment spells predicted to be a month and a half longer, fresh non-employment spells predicted to be almost a month longer, and fresh employment spells predicted to be two-thirds of a month shorter. Surprisingly, none of the welfare policy or macroeconomic variables has a statistically significant estimated effect for left-censored employment spells (top right panel).

Comparing these results to those from the LDUM and LDAT models (second and third columns of each panel in Table 3A), it is clear that the ad hoc approaches differ most for fresh spells. While the estimated expected durations for the left-censored spells are similar across the three models, the LDUM and LDAT models predict much longer expected durations for the fresh spells. The fresh spell estimates from the LDAT are the longest, which is consistent with the problem noted earlier that the LDAT approach would exclude many short fresh spells. Moreover, the standard errors on the expected durations for fresh spells
calculated for the LDUM and LDAT models are considerably larger than the standard errors from our seam bias-corrected model.

We have indicated in italics where the estimated effects of changes differ in magnitude or statistical significance between our seam-bias-corrected model and the ad hoc approaches. For left-censored non-employment spells the LDAT approach differs in the statistical significance of the maximum welfare benefit effect, and that is true for the fresh non-employment spells for carrot waiver policy and the unemployment rate. The conclusions one would draw from the LDAT (and to a lesser extent the LDUM) model differ most strikingly for fresh employment spells, where the maximum welfare benefit, the minimum wage, and the unemployment rate results all differ from our misreporting model.

6.2.2 Effects for the Demographic Variables

Table 3B presents the effects of changing demographic characteristics on the expected durations of the different spell types, and again we start with the effects from our constant misreporting model. Demographic variables are clearly important in affecting the duration of non-employment and employment. Being older (age 35 versus age 25) significantly increases the expected duration by 7.5 months for left-censored non-employment spells and by 5.1 months for left-censored employment spells. It also increases the expected duration of fresh employment spells (by 1 month), although not fresh non-employment spells. Unlike age, having more education (12 years instead of fewer than 12 years) leads to shorter non-employment spells and longer employment spells, with larger estimated effects for the left-censored spells in both cases (5.3 months shorter for non-employment spells and 7 months longer for employment spells). Race/ethnicity have significant effects on both types of left-censored spells and fresh non-employment spells, while the number of children younger than 6 years old significantly affects all spell types except fresh employment spells. Going from white to nonwhite or adding a child under 6 all produce effects of the opposite sign compared to increasing schooling, as one would expect.

Compared with the effects presented in Table 3A, in general there are larger discrepancies between the constant correction and the LDUM models in terms of both magnitude and statistical significance of the counterfactual effects, while the LDAT model continues to differ from the seam-bias-corrected model most often. Again we have used italics to indicate when the differences are substantial in either sign or statistical significance or both, or when the estimated effects fall outside the confidence interval of the seam bias correction model estimates.

In summary, our results in Table 3A and 3B suggest that changing welfare policy parameters can have substantial effects on the duration of employment and non-employment among less-educated women. While these effects are concentrated on non-employment spells (both left-censored and fresh), there are effects on fresh employment spells, but little evidence for an impact on left-censored employment spells. To the extent that women in such spells are the ones most solidly connected to the labor market, the lack of significant effects of welfare policy may not be particularly surprising. The estimated impacts of
demographics are similarly consistent with theory and previous research, although the magnitudes of these effects have been generally unknown previously. Completing high school has strikingly large effects (around half a year for the left-censored spells), and while race and ethnic impacts are smaller, they remain even though we are accounting for individual heterogeneity. Importantly, the effects of the counterfactual changes we consider are much larger in absolute value in the left-censored spells than in the fresh spells, indicating that it is crucial for policy purposes to analyze left-censored spells. Moreover, we demonstrate that there are important differences between our seam bias estimates and estimates from the last month dummy variable model in terms of the counterfactual effects for the demographic variables but not for the business cycle and policy variables, while there are very important differences between our seam bias estimates and those obtained using the last month data for all of the counterfactual effects. In terms of effects on expected duration, there are more substantial discrepancies between the seam-bias-corrected model and these ad hoc models in the estimates of the effects of changing demographic characteristic than in the estimates of the effects of changing public policy or macro conditions.

6.3 Simulation Results for the Effects of Changing an Explanatory Variable on the Fraction of Time Spent Employed

The effect of changing an explanatory variable on the fraction of time that an individual is employed is another important parameter for policy makers. Note that the simulated employment histories and fractions depend only on the estimates of the parameters for all four hazard functions \( \lambda_U(\bullet), \lambda_{\hat{U}}(\bullet), \lambda_{\hat{E}}(\bullet) \) and \( \lambda_{\hat{E}}(\bullet) \) and the estimates of the support points \( \theta_1, \theta_2, ..., \theta_J \), and the associated probabilities \( p_1, p_2, ..., p_J \); they do not depend on the parameters describing misreporting, i.e. the \( \gamma \)'s and \( P_A \), since the latter parameters refer solely to the mismeasurement of spell lengths, which is not an issue in calculating the counterfactual effects here nor in Section 6.2. Table 4 presents our simulation results, with estimated 95% confidence intervals presented in brackets. Our reported estimates are the first in the literature for a nationally representative US sample as opposed to the experimental training data sets used previously, and our confidence interval estimates from the CI-bootstrap approach are the first consistent estimates in the literature for any counterfactual effects when the outcome variable is discrete.

The first row contains the estimated average fraction of time employed for 3, 6 and 10 years after the sample begins. Our estimates [95% confidence intervals] are 0.431 [0.396, 0.471], 0.439 [0.399, 0.480] and 0.449 [0.408, 0.490]. Panel A shows estimates of how the fraction of time employed changes with welfare policies and general macroeconomic conditions. Increasing the state maximum monthly welfare benefits by 10% reduces the predicted fractions of time employed by 0.2 to 0.3 percentage points. Note that while these

---

30 Strictly speaking our data is a representative sample for the states that we can identify in these SIPP data, i.e. the US excluding Maine, Vermont, Iowa, North Dakota, South Dakota, Alaska, Idaho, Montana, and Wyoming
effects are small and statistically insignificant, all three estimates have narrow 95% confidence intervals. For example, in elasticity terms, the 95% confidence interval for the estimated 3-year period effect of increasing welfare benefits is $[-0.14, 0.05]$. Further, the estimated effects for the 6-year and 10-year horizons are only slightly larger in absolute value, suggesting that cumulative dynamic effects are not playing much of a role here. Implementing a carrot waiver policy increases the predicted fraction of time employed by 3.6 to 4.2 percentage points, with the effects for the 6-year and 10-year horizons being slightly larger than that for the 3-year horizon. Although these effects are large and of the expected sign, none of them are statistically significant at the 5% level, given the estimated 95% confidence intervals. Interestingly, if we had employed the traditional AD-bootstrap approach that is inconsistent for this setting, we would have found that the short run, medium run, and long run effects are statistically significant at the 10% level for both changing welfare benefits and introducing a carrot policy. The estimated effects of implementing a stick type of welfare policy or increasing the minimum wage are very small and statistically insignificant, with larger confidence intervals. Finally, the 3-year horizon effect of a 25% increase in the unemployment rate (e.g. from 6 percentage points to 7.5 percentage points) has the expected sign and just misses being statistically significant, since the 95% confidence interval is $[-0.031, 0.003]$. Again, it is interesting to note that the AD-bootstrap approach would have suggested that this counterfactual effect is statistically significant at the 5% level.

Panel B contains the effects of changing the demographic variables. Having a high school degree versus having less education is estimated to increase the time employed fraction by a substantial amount, 9 to 10 percentage points, and the effects for all time horizons are statistically significant at the 5% level. The estimated fractions of time employed over the different horizons for African Americans are 3 to 4 percentage points lower than the fractions for Whites with other characteristics being held constant, but these effects are not quite statistically significant since the 95% confidence intervals for the 3-, 6- and 10-year horizons are $[-0.067, 0.013]$, $[-0.073, 0.012]$ and $[-0.076, 0.012]$ respectively. The next row shows that the effects of being Hispanic relative to being White are similar. Finally, having one child less than 6 years old versus none reduces the time employed fraction by about 3 to 3.5 percentage points, and these effects are on the margin of being significant at the 5% level since the 95% confidence intervals are $[-0.069, 0.004]$, $[-0.074, 0.003]$, $[-0.077, 0.004]$ for the 3-year, 6-year and 10-year horizons. It is perhaps worthwhile noting that again all of these effects would be judged statistically significant at the 5% level if we had used the AD-bootstrap approach.

7. **Summary and Conclusion**
Understanding transitions into and out of employment for disadvantaged single mothers is of crucial importance to policy makers, as these transitions determine unemployment rates, poverty rates and the overall well-being of many low-income families. Such employment dynamics have been relatively
understudied in the literature, given the substantial policy interest paid to less-educated single mothers and their children, and the emphasis that policy makers have placed on increasing employment durations as a means of increasing on-the-job human capital for disadvantaged women. We use the SIPP to estimate monthly transition rates into and out of employment for these women, and then use these to calculate a host of important counterfactual policy effects. Contrary to much of the literature, we consider both left-censored and fresh spells, since the former dominate the employment histories of disadvantaged mothers. Moreover, we find that the most important policy effects occur in left-censored spells. We focus on a time period during which there was a fundamental shift in US policy towards single mothers, moving away from cash-based assistance unconditional on work and towards work-oriented policies such as enhanced earnings disregards (“carrots” to leave welfare) and time limits, work requirements and family caps (“sticks” to leave welfare).

Seam bias is an important problem that faces any researcher estimating transition rates using panel data given that both the end date and start date of spells may be mismeasured. In this study we propose an approach that is consistent with the telescoping phenomenon widely documented in the statistics literature. Since our model (in the absence of seam bias) is much richer than those for which misclassification has been previously considered, we encounter identification issues unique to our study. We show that our models are identified without restricting the form of the duration dependence as long as we assume that the misreporting probabilities are the same for left-censored and fresh employment (non-employment) spells, and the data are reasonably consistent with this assumption. We find that allowing for individual heterogeneity in misreporting is important for understanding misreporting behavior but not for estimating the hazard functions.

Results from our seam bias correction model indicate that changing welfare benefits significantly increases the length of left-censored non-employment spells only and decreases the duration of fresh employment spells only. The carrot policy of increasing the amount of earnings a single mother can keep before losing welfare benefits has a sizeable effect on fresh non-employment spells, reducing their expected duration by three months, although we are unable to detect significant effects on the other types of spells. An increase in the minimum wage is predicted to increase the length of fresh employment spells by almost a month and a half. Finally, an increase in the unemployment rate significantly affects all but left-censored employment spells, with the absolute size of the effect being largest in left-censored non-employment spells.

Not surprisingly, we find that demographic variables play an important role in determining the length of spells of employment and non-employment. Our estimates imply that growing older increases the duration of left-censored non-employment spells, left-censored employment spells and fresh employment spells, with the effects on the left-censored spells of both types being around half a year. Perhaps most importantly, completing 12 years of school increases the duration of each type of employment spell and decreases the duration of each type of non-employment spell. These effects are between two and three months for the fresh spells and between five and seven months for left-censored spells. African Americans and Hispanics
generally have longer non-employment spells and shorter employment spells, while having a child less than 6 years old increases the duration of non-employment spells and decreases the duration of left-censored employment spells. Finally, we find significant effects on the fraction of time employed of having more schooling, while the effects for increasing the unemployment rate and having a child are right on the edge of statistical significance. While the signs of the effects that we find are not surprising, to the best of our knowledge, the magnitudes of the effects were not currently known in the literature.

We also estimate the hazard functions using two ad-hoc approaches used previously to address seam bias: using only the last month data and putting a dummy variable in the hazard for the last month. We find that the last month dummy approach does a good job of estimating the effects of changing policy and macro variables, but does a poor job with regard to changing the demographic variables. We find that the last month data model does a poor job of estimating all the effects we consider.

References


Table 1. Characteristics of Employment and Non-employment Spells
Single Mothers with a Maximum of Twelve Years of Education

<table>
<thead>
<tr>
<th>Panel A: Non-employment spells</th>
<th>Left-censored spells</th>
<th>Fresh spells</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std Dev</td>
</tr>
<tr>
<td>Right censored (%)</td>
<td>64.5%</td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>0.34</td>
<td>0.48</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.23</td>
<td>0.42</td>
</tr>
<tr>
<td>12 years of schooling</td>
<td>0.44</td>
<td>0.50</td>
</tr>
<tr>
<td>Age</td>
<td>30.08</td>
<td>9.01</td>
</tr>
<tr>
<td>Never married</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td># of children &lt; 18</td>
<td>1.98</td>
<td>1.19</td>
</tr>
<tr>
<td>Age of youngest child</td>
<td>5.00</td>
<td>5.09</td>
</tr>
<tr>
<td># of children &lt; 6</td>
<td>0.93</td>
<td>0.91</td>
</tr>
<tr>
<td>Disability (adult or child)</td>
<td>0.24</td>
<td>0.42</td>
</tr>
<tr>
<td>Disability variable missing</td>
<td>0.17</td>
<td>0.37</td>
</tr>
</tbody>
</table>

| Number of spells               | 3,528    |         | 2,578    |         |
| Number of individuals          | 3,528    |         | 1,889    |         |
| Number of observations: year*individual | 63,384 |         | 18,811   |         |

<table>
<thead>
<tr>
<th>Panel B: Employment spells</th>
<th>Left-censored spells</th>
<th>Fresh spells</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std Dev</td>
</tr>
<tr>
<td>Right censored (%)</td>
<td>69.9%</td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>0.26</td>
<td>0.44</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.14</td>
<td>0.35</td>
</tr>
<tr>
<td>12 years of schooling</td>
<td>0.74</td>
<td>0.44</td>
</tr>
<tr>
<td>Age</td>
<td>33.99</td>
<td>8.49</td>
</tr>
<tr>
<td>Never married</td>
<td>0.27</td>
<td>0.45</td>
</tr>
<tr>
<td># of children &lt; 18</td>
<td>1.55</td>
<td>0.86</td>
</tr>
<tr>
<td>Age of youngest child</td>
<td>8.14</td>
<td>5.44</td>
</tr>
<tr>
<td># of children &lt; 6</td>
<td>0.46</td>
<td>0.66</td>
</tr>
<tr>
<td>Disability (adult or child)</td>
<td>0.12</td>
<td>0.33</td>
</tr>
<tr>
<td>Disability variable missing</td>
<td>0.13</td>
<td>0.34</td>
</tr>
</tbody>
</table>

| Number of spells               | 3,826    |         | 2,732    |         |
| Number of individuals          | 3,826    |         | 2,000    |         |
| Number of observations: year*individual | 71,613 |         | 21,376   |         |

Total number of individuals: 7,354
Total number of observations: 175,184

Notes:
1. Sample means are from the first month of spells.
2. Summary statistics across spells are not independent in the sense that some individuals show up in multiple left-censored and fresh spells.
3. The numbers of spells reported in this table include both completed spells and right-censored spells.
4. The total number of spells reported in this table is not the sum of the number of individuals in the four types of spells because some individuals have multiple spells of different types, e.g. a left-censored non-employment spell followed by a fresh employment spell.
Table 2. Misreporting Probabilities Due to Seam Bias
Constant Probabilities vs. Probabilities Varying by Race

<table>
<thead>
<tr>
<th></th>
<th>Non-employment Spells</th>
<th>Employment Spells</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant Probabilities</td>
<td>Probabilities Varying with Race</td>
</tr>
<tr>
<td>Month 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.568**</td>
<td>White 0.525**</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.042)</td>
</tr>
<tr>
<td></td>
<td>Minorities 0.598**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.039)</td>
<td></td>
</tr>
<tr>
<td>Month 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.456**</td>
<td>White 0.413**</td>
</tr>
<tr>
<td></td>
<td>(0.043)</td>
<td>(0.048)</td>
</tr>
<tr>
<td></td>
<td>Minorities 0.487**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td></td>
</tr>
<tr>
<td>Month 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.441**</td>
<td>White 0.407**</td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
<td>(0.047)</td>
</tr>
<tr>
<td></td>
<td>Minorities 0.481**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Standard errors are in parentheses.
Table 3A. Expected Durations and the Effects of Changes in Macro and Public Policy Variables
Employment and Non-employment Spells

<table>
<thead>
<tr>
<th></th>
<th>Left-censored non-employment spells</th>
<th>Left-censored employment spells</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant Misreporting Probabilities</td>
<td>Last-Month Dummy</td>
</tr>
<tr>
<td>Average Expected Duration (in months)</td>
<td>39.305** (0.731)</td>
<td>35.478** (0.574)</td>
</tr>
</tbody>
</table>

Changes with respect to:

- **Maximum welfare benefits increasing by 10%**
  - Carrot waiver (implemented - not implemented): -0.639 (2.892) - 0.552** (0.143) - 0.585 (0.495)
  - Stick waiver (implemented - not implemented): 2.427 (3.458) - 3.228 (3.416) - 1.005 (3.195)
- **Minimum wage increasing by 10%**
- **Unemployment rate increasing by 25%**

<table>
<thead>
<tr>
<th></th>
<th>Left-censored non-employment spells</th>
<th>Left-censored employment spells</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant Misreporting Probabilities</td>
<td>Last-Month Dummy</td>
</tr>
<tr>
<td>Average Expected Duration (in months)</td>
<td>11.821** (0.516)</td>
<td>16.458** (1.312)</td>
</tr>
</tbody>
</table>

Changes with respect to:

- **Maximum welfare benefits increasing by 10%**
  - Carrot waiver (implemented - not implemented): -3.099** (1.009) - 4.531** (1.109) - 3.139 (1.270)
  - Stick waiver (implemented - not implemented): 3.138 (2.100) - 4.388 (3.212) - 2.950 (4.323)
- **Minimum wage increasing by 10%**
- **Unemployment rate increasing by 25%**

Note: Standard errors are calculated using the delta method. ** indicates estimate is statistically significant at the 95% level; * indicates estimate is statistically significant at the 90% level. Italics indicates cases where there are differences between results from our seam bias corrected approach and the last month approaches in statistical significance or magnitude outside our estimated confidence interval.
Table 3B. The Effects of Changes in Demographic Variables
Employment and Non-employment Spells

<table>
<thead>
<tr>
<th>Changes with respect to:</th>
<th>Left-censored non-employment spells</th>
<th>Left-censored employment spells</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant Misreporting Probabilities</td>
<td>Last-Month Dummy</td>
</tr>
<tr>
<td>Age</td>
<td>7.471** (1.069)</td>
<td>6.253** (0.875)</td>
</tr>
<tr>
<td>(age=35) - (age=25)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 years of schooling</td>
<td>-5.293** (1.051)</td>
<td>-6.398** (0.977)</td>
</tr>
<tr>
<td>(s = 12) - (s &lt; 12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race/Ethnicity</td>
<td>2.524* (1.299)</td>
<td>1.802 (1.206)</td>
</tr>
<tr>
<td>(Black - White)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race/Ethnicity</td>
<td>2.708* (1.408)</td>
<td>2.139* (1.292)</td>
</tr>
<tr>
<td>(Hispanic - White)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of children less than 6 years old</td>
<td>4.225** (0.906)</td>
<td>4.512** (0.895)</td>
</tr>
<tr>
<td>(one - zero)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fresh non-employment spells</th>
<th>Fresh employment spells</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant Misreporting Probabilities</td>
</tr>
<tr>
<td>Age</td>
<td>-0.070 (0.532)</td>
</tr>
<tr>
<td>(age=35) - (age=25)</td>
<td></td>
</tr>
<tr>
<td>12 years of schooling</td>
<td>-1.940** (0.595)</td>
</tr>
<tr>
<td>(s = 12) - (s &lt; 12)</td>
<td></td>
</tr>
<tr>
<td>Race/Ethnicity</td>
<td>1.842** (0.718)</td>
</tr>
<tr>
<td>(Black - White)</td>
<td></td>
</tr>
<tr>
<td>Race/Ethnicity</td>
<td>0.435 (0.771)</td>
</tr>
<tr>
<td>(Hispanic - White)</td>
<td></td>
</tr>
<tr>
<td>Number of children less than 6 years old</td>
<td>1.151** (0.581)</td>
</tr>
<tr>
<td>(one - zero)</td>
<td></td>
</tr>
</tbody>
</table>

Note: Standard errors are calculated using the delta method. ** indicates estimate is statistically significant at the 95% level; * indicates estimate is statistically significant at the 90% level. Italics indicates cases where there are differences between results from our seam bias corrected approach and the last month approaches in statistical significance or magnitude outside our estimated confidence interval.
Table 4. Fraction of Time Spent in Employment and the Effect of Changes in Public Policy, Macro, and Demographic Variables

<table>
<thead>
<tr>
<th></th>
<th>3-year Period</th>
<th>6-year Period</th>
<th>10-year Period</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Average Expected Fraction of Time Employed</strong></td>
<td>0.431** [0.396,0.471]</td>
<td>0.439** [0.399,0.480]</td>
<td>0.449** [0.408,0.490]</td>
</tr>
</tbody>
</table>

**Panel A: Changes With Respect To Policy and Macro Variables**

<table>
<thead>
<tr>
<th>Change Description</th>
<th>3-year Period</th>
<th>6-year Period</th>
<th>10-year Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum welfare benefits increasing by 10%</td>
<td>-0.002 [-0.006,0.002]</td>
<td>-0.003 [-0.007,0.002]</td>
<td>-0.003 [-0.008,0.002]</td>
</tr>
<tr>
<td>Carrot waiver (implemented - not implemented)</td>
<td>0.036 [-0.035,0.111]</td>
<td>0.039 [-0.036,0.124]</td>
<td>0.042 [-0.035,0.132]</td>
</tr>
<tr>
<td>Stick waiver (implemented - not implemented)</td>
<td>-0.002 [-0.118,0.099]</td>
<td>-0.006 [-0.128,0.103]</td>
<td>-0.011 [-0.134,0.105]</td>
</tr>
<tr>
<td>Minimum wage increasing by 10%</td>
<td>0.002 [-0.035,0.046]</td>
<td>0.002 [-0.036,0.050]</td>
<td>0.003 [-0.036,0.052]</td>
</tr>
<tr>
<td>Unemployment rate increasing by 25%</td>
<td>-0.014 [-0.031,0.003]</td>
<td>-0.016 [-0.035,0.002]</td>
<td>-0.018 [-0.037,0.000]</td>
</tr>
</tbody>
</table>

**Panel B: Changes With Respect To Demographic Variables**

<table>
<thead>
<tr>
<th>Change Description</th>
<th>3-year Period</th>
<th>6-year Period</th>
<th>10-year Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>12 years of schooling (s = 12) - (s &lt; 12)</td>
<td>0.090** [0.058,0.123]</td>
<td>0.097** [0.063,0.132]</td>
<td>0.100** [0.066,0.137]</td>
</tr>
<tr>
<td>Race (Black - White)</td>
<td>-0.031 [-0.067,0.013]</td>
<td>-0.034 [-0.073,0.012]</td>
<td>-0.037 [-0.076,0.012]</td>
</tr>
<tr>
<td>Race (Hispanic - White)</td>
<td>-0.026 [-0.069,0.019]</td>
<td>-0.028 [-0.075,0.021]</td>
<td>-0.029 [-0.078,0.025]</td>
</tr>
<tr>
<td>Number of children less than 6 years old (one - zero)</td>
<td>-0.030 [-0.069,0.004]</td>
<td>-0.033 [-0.074,0.003]</td>
<td>-0.035 [-0.077,0.004]</td>
</tr>
</tbody>
</table>

Note: In brackets are the 95% confidence intervals calculated using the CI-bootstrap approach.
Figure 1.1

Monthly Data

U’ spell: 5 months  E spell: 2 months  U spell: 24 months  E spell: 5 months

Last Month Data

U’ spell: 32 months  E spell: 4 months

Figure 1.2

Monthly Data

U’ spell: 5 months  E spell: 5 months  U spell: 21 months  E spell: 5 months

Last Month Data

U’ spell: 8 months  E spell: 4 months  U spell: 20 months  E spell: 4 months

37
Figure 1.3

Monthly Data

U' spell: 3 months
E spell: 8 months
U spell: 25 months

Last Month Data

E' spell: 12 months
U spell: 24 months

Figure 2

observed history

U' spell: 5 months
E spell: 7 months
U spell: 24 months

possible true history A

U' spell: 5 months
E spell: 8 months
U spell: 23 months

possible true history B

U' spell: 5 months
E spell: 9 months
U spell: 22 months

possible true history C

U' spell: 5 months
E spell: 10 months
U spell: 21 months
Appendix 1: Contributions to the Likelihood for a Completed Spell Ending in Months 1, 2, 3, or 4 of Reference Period \( l \)

This appendix provides derivation of equation (5.5) – (5.8) in Section 5.2. Given the assumptions A1 through A3 in Section 5.1, the contribution to the likelihood function for a completed spell of observed length \( K \) that is reported to end in month 1 of reference period \( l \) is

\[
P\left(M^{obs} = (1,l), dur^{obs} = K\right)
\]

\[
= P\left(M^{obs} = (1,l), M^{true} = (1,l), dur^{true} = K\right) + P\left(M^{obs} = (1,l), M^{true} \neq (1,l), dur^{true} \neq K\right).
\]

Since the second term is zero by assumption A1, using assumption A2 we have

\[
P\left(M^{obs} = (1,l), dur^{obs} = K\right) = P\left(M^{obs} = (1,l), M^{true} = (1,l), dur^{true} = K\right)
\]

\[
= P\left(M^{obs} = (1,l) | M^{true} = (1,l), dur^{true} = K\right) \cdot P\left(M^{true} = (1,l) | dur^{true} = K\right) \cdot P\left(dur^{true} = K\right)
\]

\[
= (1-\gamma_1)L(K),
\]

Similarly, if a transition is reported to end in month 2 or month 3 of reference period \( l \), and to have lasted for \( K \) months, we have

\[
P\left(M^{obs} = (2,l), dur^{obs} = K\right) = (1-\gamma_2)L(K) \text{ and}
\]

\[
P\left(M^{obs} = (3,l), dur^{obs} = K\right) = (1-\gamma_3)L(K).
\]

Finally, the contribution to the likelihood function of a completed spell of observed length \( K \) that is observed to end in month 4 of reference period \( l \) is given by

\[
P\left(M^{obs} = (4,l), dur^{obs} = K\right)
\]

\[
= P\left(M^{obs} = (4,l) | M^{true} = (1,l+1), dur^{true} = K+1\right) \cdot P\left(M^{true} = (4,l) | dur^{true} = K+1\right) \cdot P\left(dur^{true} = K+1\right)
\]

\[
+ P\left(M^{obs} = (4,l) | M^{true} = (3,l+1), dur^{true} = K+3\right) \cdot P\left(M^{true} = (4,l) | dur^{true} = K+3\right) \cdot P\left(dur^{true} = K+3\right)
\]

\[
+ P\left(M^{obs} = (4,l) | M^{true} = (4,l), dur^{true} = K\right) \cdot P\left(M^{true} = (4,l) | dur^{true} = K\right) \cdot P\left(dur^{true} = K\right).
\]

Note that \( P\left(M^{obs} = (4,l) | M^{true} = (4,l), dur^{true} = K\right) = 1 \) by assumption A3, and

\[
P\left(M^{obs} = (4,l) | M^{true} = (m,l+1), dur^{true} = K+m\right) = \gamma_m, \ m = 1, 2, 3, \text{ by assumption A2}.
\]

Thus, the contribution of a spell that ends in month 4 of wave \( l \) is

\[
P\left(M^{obs} = (4,l), dur^{obs} = K\right) = \gamma_1L(K+1) + \gamma_2L(K+2) + \gamma_3L(K+3) + L(K).
\]
<table>
<thead>
<tr>
<th>Year</th>
<th>Percentage</th>
<th>Sample Size</th>
<th>Percentage</th>
<th>Sample Size</th>
<th>Percentage</th>
<th>Sample Size</th>
<th>Percentage</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1986</td>
<td>48%</td>
<td>2481</td>
<td>50%</td>
<td>1123</td>
<td>84%</td>
<td>111</td>
<td>70%</td>
<td>83</td>
</tr>
<tr>
<td>1987</td>
<td>54%</td>
<td>308</td>
<td>51%</td>
<td>1105</td>
<td>77%</td>
<td>174</td>
<td>86%</td>
<td>95</td>
</tr>
<tr>
<td>1988</td>
<td>40%</td>
<td>2519</td>
<td>49%</td>
<td>1030</td>
<td>78%</td>
<td>105</td>
<td>77%</td>
<td>97</td>
</tr>
<tr>
<td>1990</td>
<td>49%</td>
<td>4629</td>
<td>55%</td>
<td>2166</td>
<td>85%</td>
<td>331</td>
<td>83%</td>
<td>332</td>
</tr>
<tr>
<td>1991</td>
<td>48%</td>
<td>2663</td>
<td>52%</td>
<td>1301</td>
<td>87%</td>
<td>151</td>
<td>72%</td>
<td>120</td>
</tr>
<tr>
<td>1992</td>
<td>48%</td>
<td>4177</td>
<td>53%</td>
<td>2271</td>
<td>88%</td>
<td>346</td>
<td>76%</td>
<td>238</td>
</tr>
<tr>
<td>1993</td>
<td>50%</td>
<td>3837</td>
<td>56%</td>
<td>1960</td>
<td>83%</td>
<td>326</td>
<td>77%</td>
<td>268</td>
</tr>
<tr>
<td>Mean of Non-Imputed: 50%</td>
<td>Mean of Imputed: 81%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spell Type</td>
<td>Interview Month</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------</td>
<td>-----------------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Month 1</td>
<td>Month 2</td>
<td>Month 3</td>
<td>Month 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left-censored Employment</td>
<td>13.36</td>
<td>18.65</td>
<td>15.78</td>
<td>52.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fresh Employment</td>
<td>9.54</td>
<td>16.35</td>
<td>21.05</td>
<td>53.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left-censored Non-employment</td>
<td>22.6</td>
<td>21.17</td>
<td>20.13</td>
<td>36.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fresh Non-employment</td>
<td>20.74</td>
<td>20.27</td>
<td>16.76</td>
<td>42.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>