Estimating Employment Transitions in the Presence of a Seam Effect

Robert B. Nielsen
University of Georgia

Alfred O. Gottschalck
U.S. Census Bureau

In this paper we demonstrate several methods used to account for seam effect in the Survey of Income and Program Participation (SIPP). We also examine the degree to which a seam effect is present in labor force data in the 1993, 1996, and 2001 panels. Strategies for analyzing data affected by the seam effect are presented to assess their utility when estimating employment transitions in panel data. Empirical results suggest that gross employment transition rates should be estimated using bounded estimates. When predicting the probability of a transition, the seam effect remains pronounced in each panel. To account for this error a seam-month indicator variable should be incorporated into multivariate prediction models.

I. INTRODUCTION

The term "seam effect" is often used to describe a form of measurement error associated with many longitudinal surveys, including the Survey of Income and Program Participation (SIPP). The tendency for longitudinal survey data to exhibit higher rates of change between interview reference periods than within reference periods may lead to biased estimates of transitions (U.S. Census Bureau, 1998). In the case of SIPP, the seam effect is evidenced by a disproportionate number of transitions in the status for a variable being reported as having occurred between the fourth month of one interview reference period (a wave) and the first month of the next wave. To the extent that respondents misreport transitions as occurring at the seam rather than within the reference period, estimates of change within the wave will be understated. Similarly, misreporting the timing of transitions as occurring at the seam overstates estimates of change between waves (Marquis & Moore, 1989). Because panel data are collected with the goal of providing analysts with an accurate description of respondents' characteristics over time, this possible misrepresentation of the timing of transition events may reduce one's confidence that the data accurately reflect the respondents' characteristics over time (Czajka, 1999).

To better understand the extent to which a seam effect is present in the SIPP and, if present, the degree to which any error is associated with labor force participation estimates, we will proceed as follows. In the next section we review existing research into the causes and consequences of measurement error at the seam in longitudinal surveys. We also outline several methods often used to remedy this form of error. In Section III we offer a description of the data we use in our analyses and outline our empirical demonstration of the seam effect and alternative methods used to account for the effect. The results of our

Robert B. Nielsen (rnielsen@uga.edu) is an Assistant Professor of Consumer Economics, University of Georgia. Alfred O. Gottschalck (alfred.o.gottschalck@census.gov) is Chief of the Labor Force Statistics Branch and Economist in the Housing and Household Economic Statistics Division of the U.S. Census Bureau.

analyses are reported in Section IV, followed by a discussion of the implications that can be drawn from this research. Finally, based on the empirical demonstration we offer guidance to users of longitudinal data who desire to account for the seam effect in longitudinal estimates of dynamic characteristics.

II. BACKGROUND

Over the past two decades researchers in numerous countries have investigated the impact of seam effects on estimates of labor force participation using longitudinal data from the Canadian Labor Market Survey (Cotton & Giles, 1998; Lemaitre, 1992), the European Community Household Panel (Jackle & Lynn, 2004), the German Socio-Economic Panel (Kraus & Steiner, 1998), the Italian Labor Force Survey (Trivellato, 1999), and the U.S. Survey of Income and Program Participation (Bassi, 1998; Moore, Marquis, & Bogen, 1996; Rips, Conrad, & Fricker, 2003; Ryscavage, 1993; Young, 1989). Despite these efforts, both the cause and ultimate remedy remain elusive (Callegaro, 2008; Czajka, 1999; Lynn, Buck, Burton, Jäckle, & Laurie, 2005; U.S. Census Bureau, 1998). As a result, researchers must make decisions about how best to account for this effect when making longitudinal estimates of dynamic characteristics.

Plausible causes typically emphasize one of two broad categories of sources of error that affect the quality of the data. The first broad category emphasizes the respondent's ability or willingness to provide accurate information. In particular, researchers have focused on the respondent's role in the data collection process. This research suggests that poor memory or confusion on the part of the respondent produces the over-reporting of transitions at the seam (Marquis & Moore, 1990; Rips et al., 2003). A tendency for respondents to employ any number of coping/efficiency strategies to reduce one's response burden has also been suggested (Bassi, 1998; Czajka, 1999). One such efficiency strategy, constant wave responding, describes a tendency for a respondent to simply provide the same value for all months within a wave in order to ease response burden. For example, if a change in status occurs at the end of the first month of the four-month reference period, the respondent simply reports the new status across all four months of the reference period. This possibility is cited as a plausible source of response errors at the seam (Czajka, 1999; Marquis & Moore, 1990; Rips et al., 2003). A third plausible respondent-related cause of bias at the seam is respondent conditioning. That is, over time a respondent learns that certain responses to questions result in more probes and, therefore, a more onerous survey. This implies that seam bias should change over the course of a panel as respondents become more familiar with the interview procedures. However, Martini (1989) has shown that there is notable transition clustering even at the first seam of SIPP panels.

The second broad category of plausible causes of transitions clustering at the seam focuses on methodological artifacts. In particular, methodological explanations focus on the design, implementation, and post-collection processing of the survey. For example, seminal work by Burkhead and Coder (1985) suggests that instrument construction, especially the clarity of individual questionnaire items, may affect the extent to which any seam effect is present. Researchers have since posited that the seam effect may be reduced by using more precise questionnaire wording to better guide respondents toward quality responses (Czajka, 1999; Hill, 1987), though little published work is available to support this assertion.

Still another potential methodological cause may be the result of interviewing procedures that accommodate changes in the person who actually provides information about the household for a given wave. The Census Bureau's survey protocols allow for proxy respondents to provide information about other members of the household. When the data for a particular person are provided by different respondents from one wave to the next, one might expect an increase in the number of false transitions reported at the seam (Czajka, 1999; Hill, 1987; Marquis & Moore, 1990). As a result, a change from the person of interest to a proxy respondent, which would occur at the seam because of the interview protocol, may produce transition histories that reflect greater transitions at the seam. It is possible that survey processing procedures or data user inferences introduce clustering of transitions at the seam (Lynn et al., 2005). For example, post-survey data processing procedures for filling in missing values often fill responses for one full reference period at a time, which may contribute to an unexpectedly large number of transitions at the seam. Alternatively, to ease analyses data users may infer that transitions occur at the seam if reports in successive waves differ (Martini, 1989; Young, 1989).

As illustrated by the above discussion, the exact causes of seam bias are not yet fully understood. Empirical evidence that respondents' transitions cluster around the seams, however, is readily available. For example, Marquis and Moore (1989) used administrative records to estimate the presence of a seam effect. Their inspection of survey data revealed that relative to administrative records, changes between waves were overstated and changes within waves were understated. Given this knowledge, researchers have attempted to account for the clustering around the seam. Gottschalk and Moffitt (1999), for example, specify predictive models that incorporate a dummy indicator that equals one when transitions are reported as having occurred at the seam when estimating the timing of job exits to identify the presence and the effect of seam bias in multivariate estimates of labor force transitions. In their work, seam dummy indicators are invariably significant predictors of labor force transitions. As further evidence of a pronounced seam effect, Nelson (2003) reports that in the 1996 SIPP panel, spells without health insurance are overwhelmingly reported as having occurred in 4-month intervals, the same number of months between SIPP interviews. As a result, estimates of spells without insurance coverage that do not account for clustering at the seams are overwhelmingly estimated to be multiples of 4 months—a statistically unlikely proposition.

Given that a seam effect remains present in SIPP data despite efforts to reduce it (U.S. Census Bureau, 2005), researchers who use the data to study transitions must seek alternative ways to use the rich longitudinal data while accounting for the demonstrated error. In the remainder of this paper we discuss descriptive data from the 1993, 1996, and 2001 SIPP panels that illustrate the presence of a seam effect in each of the panels, despite efforts to eliminate it. Then, we outline our demonstration of several methods used to account for the presence of a seam effect by considering estimation methods that may be used when estimating labor force transitions over time.

To provide a baseline from which to assess the presence of a seam effect we take interviews at face value and treat respondent reports as truth. This approach assumes that no seam effect exists. Next, we consider an estimation strategy that explicitly acknowledges the seam effect by constructing bounded estimates, similar to Martini (1989). The second estimation strategy that we consider accounts for the possibility that respondents simply report their current status across all months in the reference period. A creative analytical approach employed by Fitzgerald (1999) illustrates this method. This approach, which we refer to as a "lagged check" method, counts only transitions that are verified to have remained in effect four months later in the subsequent wave. These first two strategies are useful when estimating gross labor force transition flows. The final estimation strategy, which we refer to as the "seam indicator" method, is based on work by Gottschalk and Moffitt (1999). This strategy, which is appropriate when specifying predictive models rather than estimating gross flows, incorporates a dummy indicator that equals one when an observation occurs in a seam month. Based on the empirical results of these methods for accounting for the presence and impact of the seam effect, we offer recommendations for employment-related panel data users.

III. DATA AND METHODS

Data used in this research are drawn from the 1993, 1996, and 2001 panels of the Survey of Income and Program Participation (SIPP), a nationally-representative sample of households conducted by the U.S. Census Bureau (U.S. Census Bureau, 1998). The survey's population universe is the civilian non-institutionalized population of the United States. Each household included in SIPP is selected via a two-stage sample and interviewed every four months, a period called a "wave."

The 1993 and 2001 SIPP panels consist of 36-months of data whereas the 1996 panel was fielded for 48 months. To simplify comparisons among the three panels, we use only the first 36 months of data from the 1996 panel. All data used in these analyses are drawn from the core sections of the first nine waves (36 months) of the 1993, 1996, and 2001 SIPP panels. Longitudinal full panel weights that adjust for differential non-response and attrition rates are used (for a complete discussion of SIPP sample selection and weighting procedures for the three panels, see U.S. Census Bureau, 1995, 2003, 2005).

Labor force status transition rates are estimated via the construction of a standard labor force transition matrix, as shown in Figure 1. For this analysis, people are considered employed (E) if they were with a job all or part of the month, unemployed (U) if they had no job all month and had at least one week on layoff or of looking for work, and not-in-labor force (N) if they had no job all month and spent no time on layoff or looking for work. Of importance here are the off-diagonal elements that identify a transition from one status to another status (in bold in Figure 1). Six such transitions are possible: 1) employed to unemployed (EU); 2) employed to not-in-labor force (EN); 3) unemployed to employed (UE); 4) unemployed to not-in-labor force (UN); 5) not-in-labor force to employed (NE); and 6) not-in-labor force to unemployed (NU). These off-diagonal elements are estimated using the two strategies described in Section II.

Figure 1. Labor Force Status Transition Matrix

	Time (T+1)		
Time (T)	**Employed (E)**	**Unemployed (U)**	**Not-in-labor-force (N)**
Employed (E)	EE	**EU**	**EN**
Unemployed (U)	**UE**	UU	**UN**
Not-in-labor-force (N)	**NE**	**NU**	NN

We also estimate the probability of transitioning from employed to unemployed and unemployed to employed (i.e., the EU and UE cells of the labor force transition matrix shown in Figure 1). These transitions are estimated via a standard dichotomous dependent variable model using two different specifications. As shown in Equation 1, the first model estimates transition probabilities by calculating month-to-month transitions without accounting for the presence of any seam effect. Specifically, our base model takes the form

$$T_i = \beta_0 + \beta_1 \, AGE_i + \beta_2 \, AGE_i^2 + \beta_3 \, TIME_i + \varepsilon_i \,, \quad (1)$$

where T_i is a dichotomous indicator variable taking the value of one for a transition occurring and a value of zero for a transition not occurring; *AGE* is the age of the individual at the time of a transition occurring or not occurring; AGE^2 is the individual's age squared; *TIME* is a trend term representing the panel-month of the observation; and ε_i is an error term.

As shown in Equation 2, the second model also estimates transition probabilities based on month-to-month transitions. However, Equation 2 differs in that it incorporates a seam indicator variable into the prediction equation to estimate the unique contribution of any seam bias associated with the estimate. Specifically, our second model specification takes the form

$$T_i = \beta_0 + \beta_1 \, AGE_i + \beta_2 \, AGE_i^2 + \beta_3 \, TIME_i + \beta_4 \, SEAM_i + \varepsilon_i \,, \quad (2)$$

where T_i is a dichotomous indicator variable taking the value of one for a transition occurring and a value of zero for a transition not occurring; *AGE* is the age of the individual at the time of a transition occurring or not occurring; AGE^2 is the individual's age squared; *TIME* is a trend term representing the panel-month of the observation; *SEAM* is equal to 1 when the observation occurs at a seam (i.e., the observation occurs between month 4 of one wave and month 1 of the subsequent wave) and zero when the observation occurs between non-seam months; and ε_i is an error term.

Equation 1 is estimated as a PROBIT model for eight race, sex, and education sub-groups using all reported transitions. That is, we estimate transitions for White males who did not attend college, White males who attended college, Non-White males who did not attend college, Non-White males who attended college, White females who did not attend college, White females who attended college, Non-White females who did not attend college, and Non-White females who attended college. Equation 2 is disaggregated into these same eight race, sex, and education sub-groups. The predictors in our employment transition models are selected simply out of convenience rather than to illustrate any effects of individual variables.

IV. RESULTS

For each of the three SIPP panels our analysis sample is comprised of persons aged 15 years or older who were interviewed each wave of the panel (i.e., only individuals who have a positive longitudinal panel weight). Sample means for select demographic variables by SIPP panel are shown in Table 1. In each of the three panels, females account for approximately 52 percent of the sample, whereas males account for 48 percent. In addition, White individuals comprise 83 percent of our sample, whereas Blacks and Asians comprise 12 and 4 percent of our sample, respectively. The average age for each panel is roughly 43 years old.

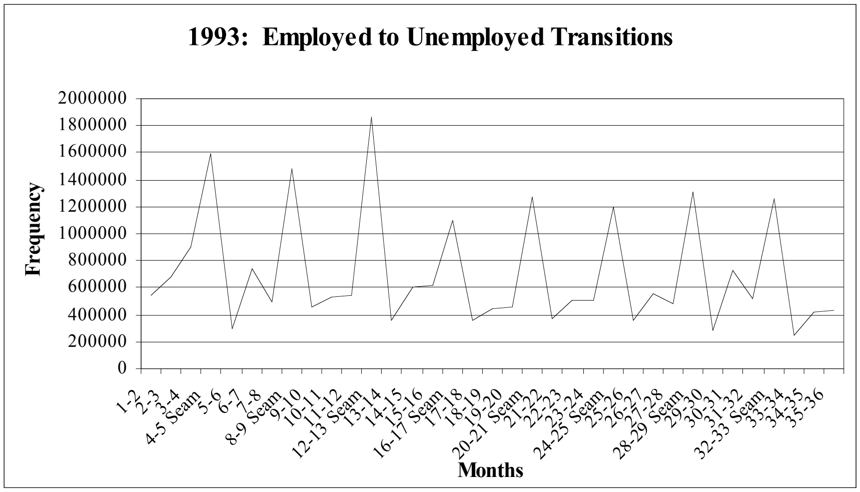
Table 1. Sample Characteristics by Select Demographic Variables by SIPP Panel (percent)

	SIPP Panel		
	1993	**1996**	**2001**
Male	48.1%	48.2%	48.2%
	(0.27)	(0.29)	(0.26)
Female	51.9%	51.8%	51.8%
	(0.27)	(0.29)	(0.26)
White	84.3%	83.4%	82.7%
	(0.2)	(0.22)	(0.2)
Black	11.6%	11.8%	11.7%
	(0.2)	(0.22)	(0.19)
Asian	3.4%	3.8%	4.3%
	(0.1)	(0.11)	(0.11)
Mean Age (years)	42.4	42.7	43.4

Source: 1993, 1996, and 2001 SIPP Panels, U.S. Census Bureau.
Note: Standard errors (in parentheses) are estimated using generalized variance parameters. All estimates are weighted.

Month-to-month gross flows from employed to unemployed – the EU cell of the labor force transition matrix – for the 1993, 1996, and 2001 panels of SIPP are shown in Figures 2-4. Of particular importance here is the confirmation that transitions from employed to unemployed spike at the seams of each wave (month pairs 4-5, 8-9, and so forth). Recall from our earlier discussion of the SIPP survey design that data from each wave are collected from a different interview for each individual. Figures 2-4 graphically illustrate how the majority of labor force transitions reported by respondents, or inferred or allocated in

Figure 2. Gross Flows of Employed to Unemployed: 1993 SIPP Panel



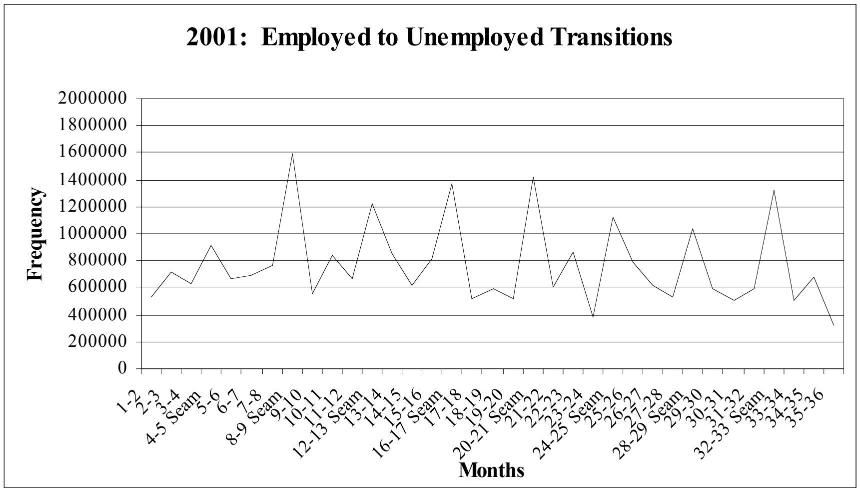
Source: Authors' tabulations from the 1993 SIPP Panel, U.S. Census Bureau. Estimates are weighted.

the production process, occur between waves and not within waves. Though not displayed here, this tendency for transitions to be reported at the seams is evident across all three panels for each and every one of the off-diagonal labor force transition cells in Figure 1.

Figure 3. Gross Flows of Employed to Unemployed: 1996 SIPP Panel

Source: Authors' tabulations from the 1996 SIPP Panel, U.S. Census Bureau. Estimates are weighted.

Figure 4. Gross Flows of Employed to Unemployed: 2001 SIPP Panel



Source: Authors' tabulations from the 2001 SIPP Panel, U.S. Census Bureau. Estimates are weighted.

The percentage of the total month-to-month transitions, shown in Table 2, confirm that the majority of transitions are reported between waves, not within waves. Notably, each of the labor force transition types shows a decline between the 1993 and 1996 panels. This decline is not too surprising given that the 1996 survey underwent a major redesign that included the adoption of computer aided interviewing techniques that allowed interviewers to access responses from prior waves (Doyle, Martin, & Moore, 2000). The data

Table 2. Total Month-to-Month Transitions that Occur at a Seam (percent)

	SIPP Panel		
	1993	**1996**	**2001**
EU	45.17	33.92	37.10
	(0.77)	(0.88)	(0.72)
EN	55.11	47.08	51.76
	(0.55)	(0.59)	(0.51)
UE	33.68	27.53	33.73
	(0.64)	(0.71)	(0.66)
NE	48.45	45.00	52.42
	(0.58)	(0.61)	(0.53)
UN	74.38	72.55	70.95
	(0.74)	(0.96)	(0.79)
NU	62.68	59.22	57.81
	(0.75)	(0.92)	(0.78)

Source: 1993, 1996, and 2001 SIPP Panels, U.S. Census Bureau.
Note: Standard errors (in parentheses) are estimated using generalized variance parameters. All estimates are weighted.

in Table 2 provide preliminary evidence that the 1996 panel redesign resulted in improved enumeration of transitions that occurred within interview reference periods. However, seam transition rates increase for nearly all transition types in the 2001 panel even though the 2001 panel retained the 1996 panel seam effect-reduction changes.

Average labor force transition rates of the off-diagonal elements of the labor force transition matrix for the 1993, 1996, and 2001 panels are shown in Table 3. Of the more pertinent transition types (i.e., EU and UE transitions), the transitions in Table 3 mirror the macroeconomic trends that were present in the United States over the 1993–2003 period (U.S. Government Printing Office, 2005). For example, as the growth in real Gross Domestic Product (GDP) increased and the unemployment rate decreased during the decade of the 1990s, the percentage of people who experienced an employed-to-unemployed transition decreased whereas the percentage of people who experienced an unemployed-to-employed transition increased. As the economy began to slow at the beginning of the new century (evidenced by the reduced growth in real GDP and rising unemployment rates), the percentage of people who experienced an employed-to-unemployed transition increased whereas the percentage of people who experienced an unemployed-to-employed transition decreased.

Despite mirroring macroeconomic trends on the whole, average transition rates at the seam are significantly higher than the average rates between non-seam months regardless of the type of transition. Turning again to Table 3, the average transition rate from employed to unemployed at the seam in the 1993 panel is 1.13 percent whereas the non-seam transition rate is just 0.30 to 0.48 percent. This same pattern is evident in the 1996 and 2001 panels: EU transition rates are 0.67 at the seam compared to 0.39 percent between non-seam months in the 1996 panel and 0.91 percent at a seam compared to 0.45 percent between non-seam months in the 2001 panel. This transition rate pattern remains relatively constant in all three panels across the non-seam month pairs regardless of the type of labor force transition. This empirical verification of higher transition reports at the seam provides evidence that estimates of average transition rates over the life of a panel, a particular year, or even within a single reference period, will be biased if one fails to account for the seam effect present in each of the three panels.

As noted earlier, one plausible source of error at the seam may be a tendency for respondents to fail to report spells as the recall period becomes longer. This transition reporting pattern is evident in each of the panels. However, the more recent 1996 and 2001 panels which were re-designed to, among other things, reduce seam bias, show a less pronounced effect (see Table 3). This error of omission biases downward the average transition rate.

Bounded Estimate Strategy

Despite the potential errors generated by the seam effect when calculating simple average transition rates, one can produce a bounded estimate of average gross transition rates to assess the precision of the estimate by exploiting variation in the on-seam and off-seam transitions. Following a method proposed by

Table 3. Average Monthly Transition Rates by Transition Type and SIPP Panel (percent)

Transition Type by Panel	Seam Gross Flows	Non-Seam Gross Flows			Average Gross Flows
	Mo. 4 => Mo. 1	Mo. 1 => Mo. 2	Mo. 2 => Mo. 3	Mo. 3 => Mo. 4	
EU					
1993	1.13	0.30	0.48	0.45	0.59
	(0.07)	(0.04)	(0.05)	(0.05)	(0.05)
1996	0.67	0.39	0.40	0.38	0.46
	(0.06)	(0.05)	(0.05)	(0.05)	(0.05)
2001	0.91	0.45	0.49	0.42	0.57
	(0.06)	(0.04)	(0.05)	(0.04)	(0.05)
EN					
1993	2.70	0.45	0.72	0.79	1.17
	(0.11)	(0.05)	(0.06)	(0.06)	(0.07)
1996	2.37	0.78	0.80	0.80	1.19
	(0.11)	(0.07)	(0.07)	(0.07)	(0.08)
2001	2.66	0.78	0.75	0.68	1.22
	(0.11)	(0.06)	(0.06)	(0.05)	(0.07)
UE					
1993	20.40	11.62	11.79	11.64	14.25
	(1.21)	(0.96)	(0.96)	(0.96)	(1.05)
1996	25.08	18.24	17.75	20.13	20.25
	(1.86)	(1.59)	(1.61)	(1.70)	(1.69)
2001	23.51	14.23	13.33	14.82	16.40
	(1.40)	(1.17)	(1.15)	(1.18)	(1.23)
UN					
1993	29.15	2.32	3.11	3.36	9.69
	(1.36)	(0.45)	(0.52)	(0.54)	(0.89)
1996	35.84	3.50	3.95	4.09	11.58
	(2.06)	(0.76)	(0.82)	(0.84)	(1.35)
2001	31.50	4.07	3.56	4.20	10.92
	(1.53)	(0.66)	(0.62)	(0.67)	(1.03)
NE					
1993	4.11	1.43	1.38	1.08	2.01
	(0.19)	(0.11)	(0.11)	(0.10)	(0.13)
1996	4.60	1.76	1.61	1.61	2.39
	(0.23)	(0.14)	(0.14)	(0.14)	(0.17)
2001	4.99	1.39	1.29	1.32	2.24
	(0.20)	(0.11)	(0.10)	(0.11)	(0.14)
NU					
1993	2.97	0.47	0.57	0.53	1.14
	(0.16)	(0.07)	(0.07)	(0.07)	(0.10)
1996	2.57	0.40	0.50	0.67	1.03
	(0.17)	(0.07)	(0.08)	(0.09)	(0.11)
2001	2.47	0.34	0.51	0.74	1.01
	(0.14)	(0.05)	(0.07)	(0.08)	(0.09)

Source: Authors' tabulations from the 1993, 1996, and 2001 SIPP Panels, U.S. Census Bureau.
Note: Standard errors (in parentheses) are estimated using generalized variance parameters. All estimates are weighted.

Martini (1989) and using the labor force transition estimates in Table 3, a lower bound estimate of employment status transitions can be determined by examining the off-seam transition rates. Because relatively few transitions are reported within the reference period, one may plausibly use the average transition rate between months three and four as a lower bound, especially since it is the closest of the non-seam month pairs to the interview month. In cases where the non-seam transition rate closest to the seam is not the highest rate, the highest of the remaining non-seam transition rates may be chosen. To estimate the upper bound of the transition estimate, one can use the average transition rate across both seam and non-seam months. Though any month could conceivably be used, this is a defensible and conservative upper bound because it is likely biased upwards due to the classification error component overwhelming the omission error component. Average transition rate estimates using this strategy are reported in Table 4.

Given the measurement errors described above, the bounded estimates from Table 4 show that the "true" average transition rate for a particular time period (e.g., a reference period, year, or panel) likely falls into a range between the upper and lower bounds, depending on the degree of seam effect present in the data. For both the EU and UE transitions, the range estimated to contain the "true" transition rate decreases or increases across panels (Table 4) as the prevalence of transitions occurring at a seam changes across panels (Table 2). For example, as the percentage of EU transitions occurring at a seam decreases from 45.17 percent for the 1993 panel to 33.92 percent for the 1996 panel (Table 2) the magnitude of the range of the bounded EU transition rate estimate also decreases, from 0.11 percent for the 1993 panel to 0.06 percent for the 1996 panel. The reverse scenario is also evident. That is, as the percentage of UE transitions occurring at a seam increases from 27.53 percent for the 1996 panel to 33.73 percent for the 2001 panel (Table 2), the range of the bounded UE transition rate estimate also increases, from 0.12 percent for the 1996 panel to 1.58 percent for the 2001 panel.

Using bounded estimates is useful when a researcher seeks a conservative estimate of the average transition rate that explicitly takes the seam effect into consideration. Bounded estimates do not require the researcher to choose a particular monthly pair on which to base one's estimates. This frees the researcher from the unique bias that may be inherent in any one particular monthly pair over the course of a panel.

Lagged-Check Strategy

Another possible method for accounting for the seam effect compares respondents' answers from two separate reference periods or waves. As a reminder, this "lagged check" method records any actual transitions if, and only if, that transition is verified four months later in the subsequent reference period. For example, if a transition is reported between months 2 and 3, then the transition will be considered "true" if and only if the transition remained in effect in month 7. Of course, using this method will result in the loss of information about spells shorter than four months in duration. But, it does offer empirical verification that a transition actually occurred. Using this lagged-check method we estimate labor force transition rates for each of the three panels and report the estimates in Table 5.

Table 4. Bounded Average Monthly Transition Rates by Transition Type (percent)

Transition Type by Panel		Lower Bound	Upper Bound
EU	1993	0.48	0.59
		(0.05)	(0.05)
	1996	0.40	0.46
		(0.05)	(0.05)
	2001	0.49	0.57
		(0.05)	(0.05)
EN	1993	0.79	1.17
		(0.06)	(0.07)
	1996	0.80	1.19
		(0.07)	(0.08)
	2001	0.78	1.22
		(0.06)	(0.07)
UE	1993	11.79	14.25
		(0.96)	(1.05)
	1996	20.13	20.25
		(1.70)	(1.69)
	2001	14.82	16.40
		(1.18)	(1.23)
UN	1993	3.36	9.69
		(0.54)	(0.89)
	1996	4.09	11.58
		(0.84)	(1.35)
	2001	4.20	10.92
		(0.67)	(1.03)
NE	1993	1.43	2.01
		(0.11)	(0.13)
	1996	1.76	2.39
		(0.14)	(0.17)
	2001	1.39	2.24
		(0.11)	(0.14)
NU	1993	0.57	1.14
		(0.07)	(0.10)
	1996	0.67	1.03
		(0.09)	(0.11)
	2001	0.74	1.01
		(0.08)	(0.09)

Source: Authors' tabulations from the 1993, 1996, and 2001 SIPP Panels, U.S. Census Bureau.
Note: Standard errors (in parentheses) are estimated using generalized variance parameters. All estimates are weighted.

Table 5. Average Four-Month Transition Rates by Type: Lagged-Check Strategy (percent)

Transition Type by Panel	Seam Gross Flows	Non-Seam Gross Flows			Average Gross Flows
	Mo. 4 => Mo. 8	Mo. 1 => Mo. 5	Mo. 2 => Mo. 6	Mo. 3 => Mo. 7	
EU					
1993	1.52	1.38	1.34	1.40	1.41
	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)
1996	1.15	1.01	1.01	1.05	1.06
	(0.08)	(0.07)	(0.07)	(0.08)	(0.08)
2001	1.69	1.40	1.40	1.46	1.49
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
EN					
1993	3.57	3.10	3.04	3.08	3.20
	(0.13)	(0.12)	(0.12)	(0.12)	(0.12)
1996	3.85	3.24	3.30	3.35	3.43
	(0.14)	(0.13)	(0.13)	(0.13)	(0.14)
2001	4.03	3.44	3.54	3.58	3.65
	(0.13)	(0.12)	(0.12)	(0.12)	(0.12)
UE					
1993	33.84	31.23	31.07	30.85	32.72
	(1.42)	(1.39)	(1.38)	(1.38)	(1.42)
1996	44.66	45.16	43.41	41.45	43.68
	(2.13)	(2.05)	(2.09)	(2.09)	(2.09)
2001	40.91	39.24	38.17	36.97	38.51
	(1.62)	(1.63)	(1.64)	(1.60)	(1.61)
UN					
1993	26.14	22.83	23.25	23.01	24.53
	(1.31)	(1.26)	(1.26)	(1.26)	(1.30)
1996	30.45	22.60	23.73	25.31	25.46
	(1.97)	(1.72)	(1.79)	(1.85)	(1.84)
2001	27.34	21.38	22.22	23.60	23.48
	(1.47)	(1.37)	(1.40)	(1.41)	(1.41)
NE					
1993	6.68	5.50	5.61	5.79	5.94
	(0.24)	(0.22)	(0.22)	(0.22)	(0.23)
1996	8.09	6.80	6.96	7.08	7.23
	(0.30)	(0.28)	(0.28)	(0.28)	(0.28)
2001	7.66	6.58	6.67	6.71	6.89
	(0.25)	(0.23)	(0.23)	(0.23)	(0.23)
NU					
1993	2.54	2.66	2.46	2.32	2.51
	(0.15)	(0.15)	(0.15)	(0.14)	(0.15)
1996	1.96	2.11	1.95	1.81	1.96
	(0.15)	(0.16)	(0.15)	(0.15)	(0.15)
2001	2.10	2.02	1.91	1.91	1.98
	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)

Source: Authors' tabulations from the 1993, 1996, and 2001 SIPP Panels, U.S. Census Bureau.
Note: Standard errors (in parentheses) are estimated using generalized variance parameters. All estimates are weighted.

As illustrated by the four-month transition rates for each of the six transition types, the four-month transition rates for each panel are higher than their month-to-month counterparts in Table 3. For example, on average, EU transition rates are 1.41 percent versus 0.59 percent, 1.06 percent versus 0.46 percent, and 1.49 percent versus 0.57 percent for the 1993, 1996, and 2001 panels, respectively. This is expected given that most transitions are reported as having occurred between reference periods, and the likelihood of experiencing a transition increases as the time span between observations increases. Furthermore, the transition rates remain stable for the majority of transitions regardless of which four-month pair is examined. For example, the EU transition rates in Table 5 range from 1.34 to 1.52 percent for the 1993 panel, whereas similar transition rates in Table 3 range from 0.30 to 1.13 percent.

It is reasonable to expect in four-month reference periods that, in the absence of any bias, approximately one-quarter of transitions would occur in each adjacent pair of months. However, the transition rate variation in Table 3, particularly when compared to the transition rates using the lagged-check method as reported in Table 5, implies that most of the difference between seam and non-seam transition rates is the result of respondents reporting (or post-processing inferences by data processors) that transitions occurred between reference periods whether the transition actually occurred at the seam or not. As we discuss in further detail in the discussion section, the failure to account for the differential reporting patterns has consequences for longitudinal estimates of dynamic labor force characteristics, including less confidence in the accuracy of the monthly reports.

Dummy Specification Strategy

We now turn our discussion to a method for accounting for bias introduced by measurement error at the seam when estimating multivariate models. As suggested by Gottschalk and Moffitt (1999), multivariate analyses commonly used to investigate labor force dynamics, such as multivariate logistic or PROBIT analyses, may control for bias introduced at the seam by incorporating a seam indicator variable to indicate whether or not the transition of interest occurred at the seam. To assess the utility of this method with each of the three panels, we estimate unemployment transitions for the eight subsamples described earlier: White males who did not attend college, White males who attended college, Non-White males who did not attend college, Non-White males who attended college, White females who did not attend college, White females who attended college, Non-White females who did not attend college, and Non-White females who attended college. Tables 6 and 7 report the estimated probabilities of experiencing an employed to unemployed (EU) or unemployed to employed (UE) transition for each SIPP panel. The predicted probabilities are estimated with and without a seam indicator variable for each model for each of the eight sub-groups.

A comparison of the two models in Tables 6 and 7 confirms that the mean predicted probabilities for each race, sex, and education sub-group are very similar, though some differences are evident in regard to the UE transitions.[1] Non-white females and males with no college education have the highest probability of experiencing a EU transition at any given month, whereas college-educated groups have the lowest probability of experiencing a transition from employed to unemployed. As for UE transitions, the college-educated groups had the highest probabilities, whereas non-whites had the lower probability of moving from unemployed to employed. This pattern is consistent across all three panels, under both model types, and for each transition type. Within each subgroup, the changes in the probability of experiencing an EU or

1 Full PROBIT results are available by request.

UE transition reflect what one would expect to see given the macroeconomic conditions present over the 1993–2003 period (U.S. Government Printing Office, 2005).

However, careful examination of the range of predicted probabilities for each sub-group and transition type suggests that the models with the seam indicator variable consistently produce a much wider range of predicted values than the models estimated without the seam indicator variable. This difference is particularly evident when examining the maximum probabilities of experiencing the employment

Table 6. Average Monthly Probability of Experiencing an Employed to Unemployed Transition

		Seam Indicator Model			Non-Seam Indicator Model		
		Mean Prob.	Minimum Prob.	Maximum Prob.	Mean Prob.	Minimum Prob.	Maximum Prob.
White Males, no college							
	1993	1.15 ***	0.21	6.31	1.16	0.32	4.18
	1996	0.96 ***	0.09	4.32	0.97	0.11	3.31
	2001	1.25 ***	0.44	4.91	1.26	0.54	3.52
White Males, college							
	1993	0.73 ***	0.31	2.74	0.73	0.39	1.89
	1996	0.60 ***	0.19	1.72	0.60	0.21	1.52
	2001	0.87 ***	0.44	2.81	0.88	0.53	2.01
Non-White Males, no college							
	1993	1.73 ***	0.29	9.73	1.76	0.45	6.57
	1996	1.23 ***	0.03	6.14	1.24	0.04	4.46
	2001	1.81 ***	0.05	8.70	1.83	0.07	5.95
Non-White Males, college							
	1993	0.88 ***	0.20	8.36	0.89	0.34	5.38
	1996	0.92 ***	0.06	4.76	0.92	0.07	3.52
	2001	1.05 ***	0.37	4.53	1.06	0.48	3.16
White Females, no college							
	1993	0.87 ***	0.28	4.02	0.88	0.38	2.73
	1996	0.91 ***	0.08	3.47	0.91	0.09	2.71
	2001	1.18 ***	0.08	4.45	1.18	0.10	3.23
White Females, college							
	1993	0.76 ***	0.25	2.68	0.77	0.33	1.84
	1996	0.64 ***	0.19	1.74	0.64	0.21	1.47
	2001	0.80 ***	0.32	2.28	0.80	0.35	1.94
Non-White Females, no college							
	1993	1.36 ***	0.26	10.61	1.37	0.37	7.25
	1996	1.25 ***	0.07	8.60	1.26	0.10	6.80
	2001	1.62 ***	0.22	5.78	1.63	0.27	4.46
Non-White Females, college							
	1993	0.83 ***	0.07	4.85	0.85	0.15	2.94
	1996	0.77 ***	0.22	4.44	0.77	0.30	2.88
	2001	1.16 ***	0.07	3.77	1.17	0.09	2.46

Source: Authors' tabulations from the 1993, 1996, and 2001 SIPP Panels, U.S. Census Bureau.
*** Seam indicator variable at or better than a 1% significance level.
** Seam indicator variable at or better than a 5% significance level.
* Seam indicator variable at or better than a 10% significance level.

Table 7. Average Monthly Probability of Experiencing an Employed to Unemployed Transition

		Seam Indicator Model			Non-Seam Indicator Model		
		Mean Prob.	Minimum Prob.	Maximum Prob.	Mean Prob.	Minimum Prob.	Maximum Prob.
White Males, no college							
	1993	14.87 ***	4.64	25.49	15.18	8.09	18.16
	1996	23.93 ***	12.40	31.56	24.14	16.98	26.38
	2001	24.31 ***	8.74	33.76	24.67	13.38	26.41
White Males, college							
	1993	18.19 **	4.81	30.43	18.27	6.47	27.45
	1996	24.59 *	3.31	32.90	24.69	4.91	29.88
	2001	23.22	10.75	33.22	23.26	10.64	30.37
Non-White Males, no college							
	1993	11.69 ***	0.17	29.69	12.23	0.66	18.73
	1996	16.14	0.56	22.88	16.19	0.87	22.61
	2001	18.02 ***	2.31	33.27	18.75	2.86	24.12
Non-White Males, college							
	1993	10.37 **	0.12	32.51	10.76	0.29	24.10
	1996	22.42 *	7.03	37.73	22.94	10.16	30.46
	2001	20.55	7.24	33.09	20.56	7.25	33.01
White Females, no college							
	1993	16.37 **	1.95	21.56	16.45	3.85	19.29
	1996	22.19	6.53	29.72	22.22	7.81	28.51
	2001	20.67 **	3.44	30.61	20.82	6.97	28.08
White Females, college							
	1993	17.90 **	2.31	25.60	18.04	3.46	21.64
	1996	22.23 *	17.54	30.08	22.34	18.97	26.23
	2001	21.86	15.27	30.59	21.91	16.98	27.28
Non-White Females, no college							
	1993	13.05 *	1.99	23.07	13.20	2.63	18.48
	1996	18.74	12.57	32.39	18.78	13.02	30.12
	2001	18.16	10.35	24.01	18.26	12.37	20.36
Non-White Females, college							
	1993	13.96 **	3.63	34.54	14.51	4.52	25.55
	1996	20.61	0.10	33.82	20.69	0.13	30.66
	2001	21.81	9.29	29.28	21.81	9.30	29.28

Source: Authors' tabulations from the 1993, 1996, and 2001 SIPP Panels, U.S. Census Bureau.
*** Seam indicator variable at or better than a 1% significance level.
** Seam indicator variable at or better than a 5% significance level.
* Seam indicator variable at or better than a 10% significance level.

transitions. This pattern holds across all three panels, though it is less pronounced in the 1996 and 2001 panels relative to the 1993 panel. This result is consistent with the decrease in the prevalence of transitions that occur at a seam for these same two panels shown in Table 2.

The dichotomous seam indicator variable is highly significant (at better than 1 percent significance) for all eight race, sex, and education subgroups when predicting EU transitions and better than five-percent significant (except for the non-White females with no college education sub-group, which is better than 10 percent significant) when predicting UE transitions over the 1993 panel. For the 1996 and 2001 panels, the seam indicator variable for the majority of subgroups remains significant for each transition type, though at lower significance levels relative to the 1993 panel, ranging from better than one-percent levels for some of the sub-groups to five-percent, ten-percent, and insignificant levels for other sub-groups. Again, the decrease in significance mirrors the decrease in the prevalence of transitions occurring at a seam shown in Table 2 and further suggests that the 1996 panel redesign reduced the clustering of transitions around the seam.

For researchers, the implication of the differing ranges produced by each model and the statistical significance of the seam indicator variable is important for at least two reasons. First, it suggests that if one does not incorporate a seam indicator variable into multivariate models, a researcher might incorrectly conclude that the probability estimates are more precise than they actually are. In other words, model-based predictions, and the associated conclusions drawn by the researcher, would indicate a higher degree of confidence than is warranted had the estimates been made with explicit acknowledgement of the existing seam effect.

Second, given that a researcher typically does not know *a priori* whether a seam effect will influence the calculated transition estimates for a particular SIPP panel, analysis sub-group, transition type, or time period, he or she should account for the presence of the seam effect regardless of analysis sub-group or SIPP panel examined to control for the demonstrated bias present around the seam.

V. CONCLUSION

In this paper we have demonstrated the extent to which the seam effect is present in labor force data from the 1993, 1996, and 2001 panels of the Survey of Income and Program Participation. Across all three panels, the majority of labor force transitions are reported as having occurred at a seam, though the percentage of transitions occurring at a seam decreased markedly between the 1993 and 1996 panels. We attribute the decline in the prevalence of transitions occurring at a seam to the SIPP redesign efforts that were undertaken for the 1996 panel, a decline that remains in the 2001 panel data. Overall, between the 1993 and 2001 panels, the magnitude of the seam effect has decreased, yet the descriptive and multivariate estimates presented here suggest that the seam effect must still be accounted for when estimating transitions or characteristics over time.

Our estimates of the seam effect's influence on average month-to-month labor force transition rates confirm labor force transition rate estimates are affected by a seam effect in the 1993, 1996, and 2001 panels. The level of incfluence depends on which monthly pair the researcher chooses when calculating an average transition rate (i.e., seam versus non-seam), and whether the average transition rate is calculated over the course of a wave or a panel. Given that these transition rates are demonstrably subject to error and differ depending on what time frame the researcher chooses, we recommend that researchers create lower and upper bounds for the month-to-month transition rates.

Our demonstration also indicates that another reasonable strategy for those who desire to account for the seam effect when estimating gross flow transition rates is to compare data from two separate interviews by calculating four-month average transition rates. Though more long-term in nature than month-to-month rates, the four-month lagged-check method allows the researcher to take advantage of the dynamic nature of SIPP data. In cases where the research question investigates relatively long-term dynamics rather than monthly dynamics, we argue that the use of a four-month rate is a particularly useful technique. Again, if relatively short-term changes are of interest—as is often the case among SIPP users who are interested in monthly dynamics—this may not be the best strategy. Across all three SIPP panels, four-month transition rate estimates remain remarkably stable (i.e., less prone to exhibit the spikes evident in Figures 2-4) regardless of whether a seam or non-seam four-month pair is examined.

Our demonstration also shows that when using multivariate models to predict the probability of experiencing a transition, the presence of a seam effect has an impact on predicted employment transition probabilities for several commonly-used demographic sub-groups. Notably, mean predicted probabilities are minimally affected when we include a seam indicator variable in our models. That is, the estimated probability that a labor force transition occurs remains relatively stable regardless of whether a seam indicator variable is included in the prediction equation. However, the dispersion around the mean is affected by the inclusion of the seam indicator dummy variable. Again, as this result demonstrates, analysts who do not incorporate a seam indicator variable into a multivariate model may conclude that the labor force transition probability estimates are more precise than they actually are. Consistent with Gottschalk and Moffitt (1999), our analyses show that a seam indicator variable is highly significant for all race, sex, and education sub-groups in the 1993 panel and less significant for some race, sex, and education sub-groups in the 1996 and 2001 panels.

Together, these results suggest that researchers should not simply ignore the presence of a seam effect by taking the monthly reports as "truth." Given that the seam effect influences both gross flow and predicted transition probability estimates, researchers should account for this effect by creating bounded estimates when estimating gross flow transition rates, and by incorporating a seam indicator variable when modeling labor force transition dynamics. The use of the simple estimation strategies demonstrated in this paper can increase one's confidence in estimates of dynamic labor force participation characteristics.

REFERENCES

Bassi, F. (1998). Gross flows estimation from the Survey of Income and Program Participation. *Journal of Economic & Social Measurement, 25*(2), 97-111.

Burkhead, D., & Coder, J. (1985). Gross changes in income recipiency from the Survey of Income and Program Participation: Proceedings of the Social Statistics Section, American Statistical Association Annual Meeting, pp. 351-356. .

Callegaro, M. (2008). Seam effects in longitudinal surveys. *Journal of Official Statistics*, *24*(3), 387-409.

Cotton, C., & Giles, P. (1998). The seam effect in the Survey of Labour and Income Dynamics. [Electronic Version]. *Income and Labour Dynamics Working Paper 75f0002m*. Retrieved November 30, 2006 from http://www.statcan.ca/cgi-bin/downpub/listpub.cgi?catno=75F0002MIE1998018

Czajka, J. L. (1999). *The seam effect in SIPP estimates of the number and duration of spells of uninsurance among children*. Paper presented at the 1999 Joint Statistical Meetings.

Doyle, P., Martin, E., & Moore, J. C. (2000). The Survey of Income and Program Participation (SIPP) methods panel: Improving income measurement. Unpublished Manuscript. U.S. Census Bureau.

Fitzgerald, J. (1999). Job instability and earnings and income consequences: Evidence from SIPP 1983-1995. Retrieved June 11, 2009, from http://www.northwestern.edu/ipr/jcpr/workingpapers/wpfiles/fitzgerald.pdf

Gottschalk, P., & Moffitt, R. (1999). Changes in job instability and insecurity using monthly survey data. *Journal of Labor Economics, 17*(4), 91-126.

Hill, D. H. (1987). Response errors around the seam: Analysis of change in a panel with overlapping reference periods (pp. 210-215): Proceedings of the Social Statistics Section, American Statistical Association Annual Meeting.

Jackle, A., & Lynn, P. (2004). Dependent interviewing and seam effects in work history data. [Electronic Version]. *ISER Working Paper 2004-24*. Retrieved December 2, 2006 from http://www.iser.essex.ac.uk/pubs/workpaps/pdf/2004-24.pdf.

Kraus, F., & Steiner, V. (1998). Modeling heaping effects in unemployment duration models with an application to retrospective event data in the German Socio-Economic Panel. *Journal of Economics and Statistics, 217*, 550-573.

Lemaitre, G. (1992). *Dealing with the seam problem for the Survey of Labour and Income Dynamics*. SLID Research Paper Series, No. 92-05. Ottawa: Statistics Canada.

Lynn, P., Buck, N., Burton, J., Jäckle, A., & Laurie, H. (2005). *A review of methodological research pertinent to longitudinal survey design and data collection* (No. Number 2005-29). ISER Working Paper 2005-29. Colchester: University of Essex: Institute for Social and Economic Research.

Marquis, K. H., & Moore, J. C. (1989). Some response errors in SIPP--with thoughts about their effects and remedies (pp. 381-386): Proceedings of the Survey Research Methods Section, American Statistical Association Annual Meeting.

Marquis, K. H., & Moore, J. C. (1990). Measurement errors in the Survey of Income and Program Participation (pp. 721-745): Proceedings of the Annual Research Conference, U.S. Census Bureau.

Martini, A. (1989). Seam effect, recall bias, and the estimation of labor force transition rates from SIPP (pp. 387-392): Proceedings of the Survey Research Methods Section, American Statistical Association Annual Meeting.

Moore, J. C., Marquis, K. H., & Bogen, K. (1996). *The SIPP cognitive research evaluation experiment: Basic results and documentation* (Working Paper No. 212). Washington, DC: U.S. Census Bureau: Survey of Income and Program Participation.

Nelson, L. (2003). *How many people lack health insurance and for how long?* Washington, DC: Congressional Budget Office.

Rips, L. J., Conrad, F. G., & Fricker, S. S. (2003). Straightening the seam effect in panel surveys. *Public Opinion Quarterly, 67*, 522-554.

Ryscavage, P. (1993). *The seam effect in SIPP's labor force data: Did the recession make it worse?* (Working Paper No. 180). Washington, DC: U.S. Census Bureau.

Trivellato, U. (1999). Issues in the design and analysis of panel studies: A cursory review. *Quality and Quantity, 33*, 339-352.

U.S. Census Bureau. (1995). SIPP 93: Source and Accuracy Statement for 1993 Panel Public Use Files. United Stated Department of Commerce, Economics and Statistics Administration: Washington, DC: Memorandum from Preston Jay Waite, Chief, Demographic Statistical Methods Division to Marshall L. Turner, Jr., Chief, Data User Services Division.

U.S. Census Bureau. (1998). *SIPP Quality Profile* (SIPP Working Paper No. 230). Washington, DC: U.S. Department of Commerce.

U.S. Census Bureau. (2003). SIPP 1996: Source and Accuracy Statement for the 12-Wave Longitudinal File. United Stated Department of Commerce, Economics and Statistics Administration: Washington, DC: Memorandum from Alan R. Tupek, Chief, Demographic Statistical Methods Division to Walter C. Odom, Jr., Chief, Administrative and Customer Services Division.

U.S. Census Bureau. (2005). SIPP 2001: Source and Accuracy Statement for the 9-Wave Longitudinal File. United Stated Department of Commerce, Economics and Statistics Administration: Washington, DC: Memorandum from Lawrence S. Cahoon, Acting Chief, Demographic Statistical Methods Division to Walter C. Odom, Jr., Chief, Administrative and Customer Services Division.

U.S. Government Printing Office. (2005). *Economic Report of the President*. Washington, DC.

Young, N. (1989). Wave-seam effects in the SIPP (pp. 393-398): Proceedings of the Survey Research Methods Section, American Statistical Association Annual Meeting.