Journal of Official Statistics
Vol. 10. No. 4, 1994, pp. 359–380

The Relative Empirical Validity of Dependent and Independent Data Collection in a Panel Survey

Daniel H. Hill[1]

Abstract: Data from the 1985 and 1986 SIPP panels are used to assess the relative empirical validity of occupational and industrial change measures obtained with independent and dependent data collection methods. These assessments are made in the context of simple descriptive and more elaborate event history models using two distinct definitions of change. The results suggest that for unreliable ("noisy") measures, like occupation and industry of employment, dependent measurement methods in panel surveys result in a net improvement in data quality. While it may be true that dependent measurement methods miss some true change, they also provide vastly lower amounts of spurious change. As a result, the dependent data collection method appears to improve the empirical validity of noisy measures in panel surveys.

Key words: Panel surveys; change measures; measurement error.

1. Introduction

Partly because of computer assisted interviewing methods, the use of information obtained in one wave of a panel study to condition interviewing in subsequent waves is rapidly increasing. In the newly redesigned Current Population Survey (CPS) in the United States, for instance, occupation and industry of employment measures are brought forward from prior interviews if the respondent says there was no change in employment. While this can result in important reductions in interviewing time and respondent burden, there is concern that these savings may come at the cost of increased measurement error.[2] Since respondents are prone to forget change in retrospective reporting, screening dependent measurement, like that soon to be used in the CPS, is apt to miss more true changes than would independent measurement in which all respondents were asked the full

[1] Survey Research Institute, University of Toledo, 2801 West Bancroft, Toledo, OH 43606, U.S.A.
Acknowledgement: This paper is based on research conducted under a Joint Statistical Agreement between the U.S. Bureau of the Census and the University of Toledo. The author would like to thank Dan Kasprzyk and two anonymous referees for their many helpful comments on the initial draft of this paper. I would also like to thank Rajendra Singh and Tom Scopp of the Census Bureau for their comments on earlier versions of this paper and Jim Lepkowski, Steve Pennell and David Miller of the University of Michigan for their assistance in obtaining the data. Finally, I would like to thank Mary Lou Mikula of the Survey Research Institute for her careful and extensive compilation of occupational characteristics indices. I alone am responsible for any errors.

[2] Computer aided interviewing using prior information can result in substantial reductions in error if it is used to identify inconsistencies as they are reported. In this case the interviewer can be alerted to resolve the problems while the respondent is still available to help.

set of questions each wave. Fear of such missed true changes (referred to as "false negative" measures of change) is the major argument against dependent data collection methods. To the extent that the measurement process is inherently unreliable, however, apparent change when no true change occurred (i.e., false positives) may be a more important problem in panel surveys than false negatives, and using prior measures for "non-changers" may result in a net improvement in data quality. While the ideal solution to such false-positive problems is to improve the reliability of the measurement process, using the prior measures may be a second best solution.

The 1985 and 1986 panels of the Survey of Income and Program Participation (SIPP) present a rare opportunity to study the effects of data collection method in panel surveys. The reason is that the 1986 panel used a question to determine whether occupation and industry questions should be asked in the next wave (i.e., dependent data collection), whereas the 1985 panel did not (i.e., independent data collection). Since these panels overlap in historical time and each is a probability sample of the same population and since there is no evidence of panel conditioning (Lepkowski, Pennell, Miller, Luis, and Kalton 1992), the combined data represent a randomized experiment with the treatment being the use of the screening question. The outcome measures consist of the month to month gross change in occupation and industry of employment.

In this paper, I use these data to assess the relative empirical validity of occupational (and to a lesser extent industrial) change measures using the dependent and independent collection methods. I will do so in the context of both simple cross-tabulations and more elaborate event-history models. Lacking reliable independent measures of true change, I am forced to judge the relative validity of the methods in terms of the strength of association of the observed changes with changes in other measures with which they "should" be associated. This imparts a certain amount of subjectivity to the assessment. Fortunately, the SIPP is designed in such a manner as to provide one measure (whether the transition month is a seam month) which, in the absence of measurement error, would not be associated with observed change.

Section 1 presents the results of earlier research into the dependent-independent measurement method and describes the 1985 and 1986 SIPP design and instrumentation. Section 2 presents, by method of data collection, the cross tabulations of occupation and industry change with changes in work hours, wages and employer. Section 3 develops the definitions of occupational spells and presents descriptive statistics of occupational spells by method of data collection. Finally, in Section 4 an occupational event history model is developed, estimated and used to assess the relative empirical validity of the two methods in a dynamic context.

The potential savings of screening-dependent measurement methods in panel studies have been recognized for some time. Indeed, the U.S. Census Bureau gave serious consideration to switching from independent to dependent methods for occupation and industry measures in the CPS in the mid 1970s. Based largely on the large number of false-negative change reports identified in a special job mobility study conducted in 1973, they decided to retain the independent measurement methodology. Since that time other data and analytic methods have become available which allow a fresh look at the issue.

Table 1. Job mobility rates by measurement method, CPS job mobility study (monthly change at three-digit level)

Measurement method	Industry (in percent)	Occupation (in percent)
M_i (independently asked and coded)	17.0	31.6
M_r (respondent reported change)	2.3	2.5
M_e (expert coder "true" change)	4.6	9.9

U.S. Department of Commerce (1975).

1.1. The Job Mobility Study

From April 1973 to September 1973, the CPS reinterview program included a special "Job Mobility Supplement" comprised of questions about whether the respondent's employer or job had changed since the prior month's interview. Data from these questions, in conjunction with prior month's industry and occupation from the CPS interview and the corresponding current-month data from the CPS reinterview, were used to construct three measures of job mobility. The first mobility measure (M_i) was based on a simple comparison of the two monthly codes. If the independently asked and coded occupation and industry codes were not identical, then a change was defined as having occurred. The second mobility measure (M_r) was the respondent's report of change. The third and final mobility measure (M_e) was based on clerical recoding of the occupation and industry for both months by expert "referral" coders for all cases in which there was an apparent change in the independently coded data. These referral coders were required to determine whether the apparent job change was due to (1) coding error, (2) "real" job change, or (3) due to "ambiguous" wording in the responses being coded.

Table 1 presents monthly three-digit occupation and industry mobility rates by measurement method. The most striking aspect of these results is that the independent measurement method yields mobility rates some 7–12 times greater than those reported by respondents. If we take the expert coders' mobility estimates to represent the truth (as the Census Bureau did in 1975), then we can attribute the overall difference between independent and respondent reports into that due to false-positives of the former and false-negative measures of the latter. Accordingly, approximately two-thirds of the apparent occupational mobility from the independent method is due to false positives. Although the independent method can also yield false-negative measures of change, the Joint Mobility Study (JMS) was not designed to detect them. With respect to respondent reports of change, nearly 40% of the occupation changes were judged by the expert coders not to be changes at all (i.e., false positives).[3] The remaining respondent reported changes accounted for only 13% of the "true" monthly changes identified by the expert coders.

[3] It is not clear from the report whether the expert coders actually recoded these occupations. At one point in the report the authors say that only those cases with a different three-digit code in the interview for month m and the reinterview for month m-1 were recoded. If so, then the cases with respondent reported changes without a change in three-digit codes are merely assumed to be false positives – it could equally well be that the independent method yielded false negatives for them.

Given this latter result, it is not surprising that screening-dependent methods which relied on respondent reports of change to trigger remeasurement of occupation and industry were rejected. There are, however, a number of aspects of the study which should reduce our confidence in this conclusion. Perhaps most important of these is that the determination of what constituted a true job change was made on the basis of a series of coding rules in isolation from any behavioral significance of the outcome. For example, the criterion used to identify a "true change" in occupation is "the respondent did not describe the activities, duties, and kind of work in the same way *and in the same order* from month to month" (U.S. Department of Commerce 1975, pp. 5–6, emphasis added). This means that a true change would be attributed to an individual who listed his/her duties as "*a*, *b* and *c*" in month *t*-1 and "*b*, *a*, and *c*" in month *t*. The respondent, however, is never told that the *order* of the activities and duties reported is of any importance, or, for that matter, what order (by time, importance, etc.) is preferred. Whether such a change in the order of listing of activities reflects more than just temporary variability in the salience of various job activities is irrelevant to the coding scheme. Furthermore, no attempt was made in the JMS to associate measured job changes with the levels or concurrent changes in other characteristics of the respondent or of the respondent's job. To be fair, very little could be done along these lines with the JMS – the CPS reinterview provides very few covariates of job change with which to assess the empirical validity of the job mobility measures and statistical methods then available for dynamic modeling were also limited.

Another aspect of the JMS which reduces its external validity is that the interviewers for the two monthly observations were drawn from different pools. The initial month responses were recorded by CPS production interviewers while the second month responses were recorded by reinterviewers who are mostly regional office supervisors with much more experience. Some portion of the "true" job change identified by the referral coders is likely due to systematic differences in the completeness of the responses recorded in the two months.

As I will explain below, new data and methods provide a way around these two limitations. First, the richness of the SIPP means we need not even attempt to identify "real" or "true" job changes but rather can judge the quality of data from alternative methods in terms of "meaningful" job change in the context of behavior models which include related changes and characteristics of jobs and individuals as covariates. Second, the essential survey conditions for the two measurement methods are much closer in the new data than in the JMS and there is less potential for confounding effects.

1.2. *Design and instrumentation of the 1985 and 1986 SIPP panels*

The Survey of Income and Program Participation (SIPP) is an ongoing longitudinal survey of the noninstitutional population of the United States consisting of a set of overlapping panels (see Jabine, King, and Petroni 1990 for a detailed description and evaluation of the SIPP). Each year a new panel consisting of a fresh sample from the population is added and an old panel retired. The basic design used in both the 1985 and 1986 SIPP panels consists of interviewing each sampled individual once every four months for a thirty-month period. Each interview collects monthly information on employment, income and government transfer program participation for each of the last four months. Additionally, a rich set of demographic measures is also collected for each panel individual and family.

To even out interviewing loads, sample units are randomly assigned to one of four independent "rotation groups." All members of a particular rotation group are interviewed in the same month while members of different rotation groups are interviewed in successive months. The result is a rotating reference period which insures an even distribution of recall lengths for each calendar month. Interviewing for the first rotation group for each panel begins in February of the reference year and, consequently, there is perfect overlap in the coverage of the 1985 and 1986 panels from January 1986 to May 1987.

Thus, the substantive measures are obtained from both panels for the same historic time period. Furthermore, each panel is designed to represent the same population (ignoring immigration). With respect to occupation and industry of employment, the only difference between the two panels is that the entire set of occupation and industry questions was asked and coded each wave in the 1985 panel, whereas these questions were only asked (after wave 1) of the 1986 panel if the respondent reported a new employer or answered the following screening question in the affirmative:

"2b. Have your (...'s) main activities or duties with this employer changed during the past 8 months?"

If no change were reported, the prior data were brought forward as they will be in the redesigned CPS.

As noted in the introduction, this design results in an almost ideal quasi-experiment in the sense that (nearly) identical samples of respondents are assigned at random to one of two measurement methods for the same substantive measures during the same period of time. The only factors confounding the "experiment" are the result of the 1985 panel being one year "older" than the 1986 panel. The older the panel the more it may be affected by attrition or panel conditioning. I will discuss these potential confounding effects in detail below. First, however, it is useful to point out one further feature of the SIPP design which facilitates judgement of the validity of observed changes under the two measurement methods. Specifically, a by-product of the SIPP design discussed above is that the 28-month data record is composed of seven 4-month wave records which are matched or "stitched" together. Thus, because of the random assignment of cases to rotation groups, for a rotating random quarter of the sample data for month to month change measurement come from two consecutive interviews – these observations are referred to as seam cases. The data for the remaining three-fourths of the sample come from a single interview (nonseam cases). One of the best documented manifestations of measurement errors in panels with reference periods longer than the time-unit of measurement is the tendency for between-interview change to dominate within-interview change. The SIPP design assures that whether an occupation change is associated with a seam is uncorrelated with other factors which might legitimately be associated with change. Thus, the relative association of seams and measured change is a good inverse indicator of data quality from the two data collection methods.

1.2.1. Panel attrition and conditioning

The basic concern with attrition is that at the beginning of the panel overlap period (i.e., January 1986) nonresponse in the 1985 panel (16.3% @ wave 4) was more than twice that of the 1986 Panel (7.3% @ wave 1). Since it is well documented that attrition is not random (McArthur 1987; Short and McArthur 1986; Kasprzyk, Duncan, Kalton, and Singh 1989) the populations described by the two panels for the overlap period are not the same and we cannot be entirely sure to what extent differences in job mobility rates are due to differences in measurement method alone. To alleviate this problem, the Census Bureau has constructed a series of sampling weights which include adjustments for differential nonresponse (see, e.g., Jabine, King, and Petroni 1990, chs. 5 and 8 for a description of nonresponse and weight construction). All the estimates presented in this paper are based on the 1986 final panel weights taken from the 1985 and 1986 Full Panel Longitudinal Research Files. In addition to the nonresponse adjustments these weights include post-stratification adjustments to bring the two SIPP panels into agreement with respect to age, sex, ethnicity and race as of January 1, 1986.

How effective these weights are in eliminating differential attrition effects is an open question. A great deal of recent research suggests that the weights are quite effective – at least with respect to measured levels of variables. Of particular relevance to my analysis are the reports by Lepkowski et al. (1992) and McCormick, Butler, and Singh (1992). Both of these studies examine weighted estimates from the 1985 and 1986 SIPP panels to see whether there are any residual differences. Any such differences could reflect either inadequacies in the weights or differential panel conditioning (i.e., changes in respondent reporting or actual behavior due to repeated interviewing). After extensive analysis of a wide variety of SIPP measures and, in the McCormick et al. study, after comparisons with external estimates, both studies found little if any evidence of effects.

Weighting is not the only method available for dealing with differential attrition between SIPP panels, and Lepkowski et al. compared a variety of approaches. One alternative approach would be to base comparisons of the 1985 and 1986 panels on those individuals in each panel who responded in all seven waves. To the extent that the nonresponse process is stable over time, these individuals will be similar across panels in terms of both the measured and unmeasured factors affecting nonresponse. Lepkowski et al. found the same general pattern of (non) effects when comparisons of the 1985 and 1986 SIPP panels were based on these "panel respondents" as they did with the attrition adjusted sampling weights.

Although these findings are comforting, they should not be taken to mean that the effects of differential attrition and panel conditioning are totally removed by weighting. These analyses, as well as the weights themselves, are based on the levels of variables at particular points in time. The analyses in this paper are of month to month *change*. Surprisingly little is known about the effects of attrition and conditioning on measured change in panels. There is anecdotal evidence, however, that both attrition and conditioning tend to reduce measured change as panels age – attrition because less stable individuals are harder to find, and interview and conditioning because respondents learn to be more consistent in their reports from one interview to the next. If this is so, then estimated change from the 1985 SIPP panel may be suppressed relative to that from the 1986 panel by attrition and conditioning. The net effect would be to bias my estimates of the differences in measured mobility from independent and dependent measurement methods toward zero.

2. Main-Job Occupation and Industry Changes and Their Association with Changes in Hours, Wages and Employers

I begin the empirical analysis by assessing the extent to which the overall amount of observed occupational and industrial change is affected by collection method, and how these changes relate to contemporaneous changes in other job characteristics. The simplest way of doing this is to concentrate on the main-job during the January 1986 through April 1987 period and create a record for each occurrence of a change from one month to the next in the three-digit occupation or industry code. The distribution of changes in hours, wages and employer identifiers between the corresponding months can then be calculated. In doing so, I will confine my attention to intra-marginal changes by excluding occupation and industry changes for those entering or leaving the labor force. Table 2 presents the distribution of changes in hours for all person-month changes in occupation by method of collection. The bottom row of the table presents the total number of occupational changes estimated for the population during the 15-month period. This total is nearly five times as large for the independent method data (approximately 123 million changes) as for the dependent data (approximately 26 million changes). The table also reveals that the majority of occupation changes in the independent method data had no corresponding change in work hours. With the dependent method data, on the other hand, more than half of the occupation changes were

Table 2. Distribution of changes in work hours: all month to month main job occupation changes (All workers: February 1986–April 1987)

Percent change in hours	Independent method 1985 panel	Dependent method 1986 panel
< –24% (decrease)	10,746 8.7%	5,713 21.8%
–24% to –10%	10,466 8.5%	2,514 9.6%
–9% to –1%	4,940 4.1%	1,133 4.3%
0% (no change)	72,175 58.8%	10,302 39.3%
1% to 9% (increase)	5,235 4.2%	816 3.1%
10% to 24%	10,344 8.4%	2,048 7.8%
>25%	8,784 7.1%	3,667 14.0%
Total	122,663 100.0%	26,188 100.0%

Amounts in thousands.

Table 3. Distribution of wage changes: all month to month main job occupation changes (Hourly workers: February 1986–April 1987)

Percent change in wages	Independent method 1985 panel	Dependent method 1986 panel
< −24% (decrease)	5,052 7.2%	2,670 17.1%
−24% to −10%	5,447 7.8%	2,043 13.0%
−9% to −1%	13,571 19.2%	2,184 13.9%
0% (no change)	34,239 48.6%	4,092 26.1%
1% to 9% (increase)	6,597 9.3%	1,710 10.9%
10% to 24%	3,129 4.5%	1,643 10.5%
>25%	2,454 3.5%	1,326 8.5%
Total	70,491 100%	15,671 100%

Amounts in thousands.

accompanied by an increase or decrease of 10% or more in work hours and an additional 7% had changes ranging from 1% to 10% in absolute value. Table 5 shows a very similar pattern of work hour changes for those with industry changes.

Since this same general pattern appears with respect to wage change of hourly workers (Tables 3 and 6) and employer changes (Tables 4 and 7) for both occupation and industry, it seems safe to conclude that, as in the CPS Job Mobility Study, most of the observed "change" with independent data collection methods is a result of variability in the response/coding process (i.e., noise). The distribution of changes in employer for industry changes (Table 7) provides particularly strong support for this conclusion. With independent method data,

Table 4. Distribution of employer changes: all month to month main job occupation changes (All workers: February 1986–April 1987)

	Independent method 1985 panel	Dependent method 1986 panel
No change in employer	104,606 85.3%	9,121 34.8%
Change in employer	18,057 14.7%	17,066 65.2%
Total	122,663 100%	26,188 100%

Amounts in thousands.

Table 5. Distribution of changes in work hours: all month to month main job industry changes (All workers: February 1986–April 1987)

Percent change in hours	Independent method 1985 panel	Dependent method 1986 panel
< −24% (decrease)	7,194 10.4%	5,385 26.2%
−24% to −10%	5,641 8.2%	2,000 9.7%
−9% to −1%	2,704 3.9%	708 3.4%
0% (no change)	39,427 57.0%	6,921 33.6%
1% to 9% (increase)	2,621 3.8%	636 3.0%
10% to 24%	5,437 7.9%	1,551 7.5%
>25%	6,150 8.9%	3,377 16.4%
Total	69,177 100%	20,580 100%

Amounts in thousands.

Table 6. Distribution of wage changes: all month to month main job industry changes (Hourly workers: February 1986–April 1987)

Percent change in wages	Independent method 1985 panel	Dependent method 1986 panel
< −24% (decrease)	3,986 10.4%	2,399 19.5%
−24% to −10%	3,367 8.8%	1,653 13.4%
−9% to −1%	6,378 16.7%	1,507 12.2%
0% (no change)	16,709 43.8%	2,692 21.8%
1% to 9% (increase)	3,668 9.6%	1,400 11.3%
10% to 24%	2,172 5.7%	1,472 12.0%
>25%	1,919 5.0%	1,206 9.7%
Total	38,191 100%	12,332 100%

Amounts in thousands.

Table 7. Distribution of employer changes: all month to month main job industry changes (All workers: February 1986–April 1987)

	Independent method 1985 panel	Dependent method 1986 panel
No change in employer	50,822 73.5%	3,578 17.4%
Change in employer	18,355 26.5%	17,003 82.6%
Total	69,177 100%	20,580 100%

Amounts in thousands.

nearly three quarters (73.5%) of the month to month industry "changes" were for individuals working for the same employer in both months. Whilst it is possible for individual firms to change industry, such changes are rare. Somewhat less rare are transfers from one subsidiary of a conglomerate to another which could entail an industry change without an employer change. However, these changes are also likely to be uncommon. Thus, the distribution for the dependent method data, in which more than 80% of the industry changes are accompanied by an employer change, is far more plausible than the independent method distribution.

These distributions of employer changes for industry "changers" also provide some evidence that the dependent method data capture most of the meaningful industry change in the population. Specifically, from Table 4 it is clear that the dependent method estimate of the total number of concurrent changes in employers and industries (17.066 million) is only somewhat less than 8% lower than the corresponding total (18.057 million) observed using the independent collection method. If all "real" industry changes involve an employer change, and if the independent method data capture all real changes, then these data suggest that the dependent collection methodology also captures most (more than 92%) of them.

3. Definitions of Occupational Spells

The event history models we examine here attempt to understand the timing of individual exits from occupations as a function of the characteristics of the individuals and their occupations, as well as of the amount of time the individual has been in the occupation. This last fact means that for an observation to be informative about the dynamic process, we must know when the individual entered the occupation. When we do not observe the beginning of the spell, the spell is termed "*left-censored*." When the end of a spell is not observed, the spell is termed "*right-censored*." Unlike left-censored spells, right-censored spells do provide some information about the dynamic process – although not as much information as do "*completed*" spells (those in which both the beginning and end are observed).

In order to decide when an occupational spell has ended, begun or is censored we must make some assumptions about how long an individual must be away from the occupation in order for the absence to constitute an end to the current spell. Many jobs are not full-time/full-year in nature.

Table 8. Occupational event history statistics by method of collection

	Independent method 1985 panel	Dependent method 1986 panel	SRS *t*-ratio of difference
Sample size (weight-sum in thousands)	11,042 (107,766)	10,204 (96,115)	–
Average number of occupational spells January 1986–April 1987	2.39	1.47	−60.04
Percent of spells left-censored	56.0	66.7	16.03
Percent of spells right-censored	43.9	65.1	31.06
Percent left *and* right-censored	21.3	43.8	35.17
Percent of non-left-censored spells beginning at seam	82.0	52.5	−28.81
Percent of non-right-censored spells ending at seam	87.1	64.5	−26.37
Percent of completed uncensored spells beginning and ending at seam	67.0	23.6	−24.75

Construction workers and teachers, among others, often can expect to have periodic unpaid absences from the job of several months at a time. Such absences do not signify that the individual's employment situation has changed at all. To allow for this sort of periodic non-work in some jobs, I will require a minimum absence from the (three-digit) occupation or industry of three months before it is considered as signifying an end of an employment spell. Similarly, I will require that we observe three months of non-employment prior to (after) employment in an occupation or industry in order for the subsequent spell to be considered non-left (non-right) censored.

3.1. *Effects of collection method on occupational spells*

Table 8 presents a variety of descriptive statistics for occupational spells by method of data collection. The average number of (three-digit) occupational spells per individual observed in the 16-month observation period is some 70% higher for the independent data collection method (2.39) than dependent (1.47). The difference between these estimates is highly significant (as are all of the statistics in the table) with a *t*-ratio for the null hypothesis of no difference of −60.04.[4] This is consistent with the results described in the preceding section. Similarly, the amount of censoring is substantially and significantly higher with the dependent collection method than with the independent. More than two-thirds (66.8%) of the occupational spells observed in the 1986 panel (dependent collection) were left-censored, as opposed to less than three-fifths (56%) in the 1985 panel (independent collection). Even more dramatic is the increase in right-censoring and dual-censoring (i.e., both left- and

[4] The *t*-ratios of column three of Table 8 are based on variances which do not adjust for clustering, stratification and other departures from simple random sampling. Adjustment of these variances for design effects would raise them and lower the resulting *t*-ratios – but not enough to alter the conclusions.

right-censoring) brought about by the dependent data collection method. Right-censored spells are nearly 50% (65.1% vs. 43.9%) more prevalent with dependent data collection than with independent, while dual-censored spells are more than twice as prevalent (43.8% vs. 21.3%). Thus, one effect of data collection method on event history analysis is to reduce the number of apparently informative (i.e., non-left-censored) spells available for analysis – from 4,867 for the 1985 panel to 3,578 for the 1986.

Whether or not the reduced number of non-left-censored spells brought about by dependent data collection represents a loss of information is not clear since we do not know what the true pattern of occupational change is. We do know, however, that in the absence of measurement error entrances and exits from occupations should be evenly spread over the months of the reference period. Furthermore, given the SIPP design, we would expect approximately one-quarter of all occupational spells to begin (end) at a seam and only about a sixteenth to both begin and end at seam months. The extent of clustering of transitions at the seam months is, therefore, an indication of the extent of respondent errors in properly placing events in time. Table 8 indicates that, in this respect, the dependent data collection method provides cleaner data. Roughly 85% of all non-right-censored spells observed with the independent collection method ended at a seam month. This compares with about 64% with the dependent collection method. Even more revealing is the fact that roughly two-thirds of the completed spells observed with the independent collection method both began and ended at the seam. This compares with approximately a quarter of those observed with the dependent collection method.

Thus, while the number of seam-coincident occupation changes is higher than it should be for either method of data collection, it is substantially closer to what it should be with the dependent method.

4. Effects of Collection Method on Event History Model Estimates

Although the above descriptive statistics suggest that the dependent data collection method results in less but cleaner data regarding the timing of entrances and exits from occupations, it does not necessarily follow that parameter estimates for event history models will be significantly affected by collection method. To investigate this issue, we must first develop an explicit event history model and then compare the estimates obtained from the two collection methods. Because the appropriate treatment of left-censored spells depends crucially on whether the hazard rate is a function of time in occupation or industry, we will include time explicitly in our formulation. Specifically, we will investigate models based on the following hazard function

$$h(t_i) = \Lambda(\alpha + \beta' X_i + \gamma_1 t_i + \gamma_2 t_i^2 + \gamma_3 S(t_i)) \qquad (1)$$

where Λ is the logistic function ($\Lambda(z) = \exp(z)/(1+\exp(z))$); $\mathbf{X}_i$ is a vector of characteristics of individual i; $\boldsymbol{\beta}$ is a vector of effects of these characteristics on the hazard of exiting an occupation; t_i is the time the individual has been in the occupational spell; and $S(t_i)$ is a dummy variable equaling 1 if month t_i is a "seam" month. Allowing for right-censoring, the likelihood function for our model can be expressed as

$$L = \prod_{t_i < t_{\max}} f^i(t_i) \prod_{t_i \geq t_{\max}} (1 - F^i(t_i)) \qquad (2)$$

where $t_{\max}$ is the right limit of the observation period; $f^i(t_i)$ is the probability density of individual i exiting at time t_i given that he/she has not exited prior to that time; and $F^i(t_i)$ is the corresponding cumulative density function. The first product in equation (2) represents the contribution to the likelihood function of the non-right censored spells, while the second portion represents that of the right-censored spells. The cumulative density function is related to the hazard function by

$$F(t_i) = \prod_{t=1}^{t_i-1}(1 - h(t)) \tag{3}$$

where $h(t)$ is the hazard of exiting at time t. Similarly, the probability density function is related to the hazard function via

$$f(t_i) = h(t_i)\prod_{t=1}^{t_i-1}(1 - h(t)). \tag{4}$$

Substituting equations (3) and (4) into equation (2) yields the following likelihood function for the discrete time model

$$\begin{aligned} L &= \prod_{t_i < t_{\max}} h(t_i)\prod_{t=1}^{t_i-1}(1 - h(t)) \\ &\quad \times \prod_{t_i \geq t_{\max}} \prod_{t=1}^{t_i-1}(1 - h(t)) \\ &= \prod_{t_i < t_{\max}} h(t_i)\prod_{i} \prod_{t=1}^{t_i^*-1}(1 - h(t)) \end{aligned} \tag{5}$$

where $t^* = t_{\max}$ for right-censored cases.

When dealing with a non-EPSEM sample such as the SIPP, one can apply sampling weights via the following weighted likelihood function

$$L_w = \prod_{t_i < t_{\max}} h(t_i)^{w_i}\prod_{i=1}^{n} \prod_{t=1}^{t_{\max}^*-1}(1 - h(t))^{w_i} \tag{6}$$

where w_i is the individual's sampling weight scaled by the average weight of the sample.

Finally, equation (6) is made a function of α, β and γ by substituting equation (1) for $h(t_i)$, and consistent estimates of these parameters can be obtained by maximizing the natural logarithm of the result with respect to them. The estimates are fully efficient only under the assumption of simple random sampling. The estimated sampling errors from maximizing equation (6) do not reflect the effects of departures from simple random sampling in the SIPP design and will tend to understate the true sample variability. Since the 1985 and 1986 SIPP designs are quite similar, however, comparisons of the estimated standard errors are still indicative of the relative precision of the estimates. While we shall maximize the logarithm of equation (6) directly using an algorithm written by the author, we should note that it is also possible to use packaged logit programs by creating t_i pseudo-observations for each of the i individuals in the sample (see e.g., Allison 1984). Also, we should note that as the number of time periods in the observation period increases, the probability of an individual exiting in any one period decreases and that, in the limit, our model reduces to Cox's proportional hazards model.

Table 9 presents the estimates obtained by maximizing equation (6) with respect to the parameters of equation (1) for the combined 1985–1986 SIPP panels as well as for each panel separately. The sample consists of the first non-left-censored occupational spell observed for each individual. We discard left-censored spells because we expect the hazard rate to be a function of time in occupation and these spells are uninformative. Non-left-censored spells subsequent to the first are potentially informative but require one to assume independence between spells – a very strong assumption. If the independence assumption is violated, the use of multiple spells per individual will result in biased parameter estimates.

Table 9. Discrete-time event history analysis: exit from first non-left-censored occupation (SRS t-ratios in parentheses)

	Combined sample	Independent method 1985 panel	Dependent method 1986 panel
Constant	−3.10**	−3.38**	−1.37**
	(−19.05)	(−15.84)	(−4.76)
Time in occupation	0.63**	0.63**	0.35**
	(20.55)	(14.42)	(6.93)
Time-squared	−0.08**	−0.08**	−0.06**
	(−26.12)	(−18.57)	(−10.39)
Whether seam month	3.00**	3.58**	1.94**
	(83.52)	(74.43)	(33.25)
Age (@ start)	−0.58**	−0.52**	−1.04**
	(−8.76)	(−6.12)	(−8.45)
Age-squared	0.07**	0.06**	0.12**
	(8.16)	(5.46)	(7.35)
Wage	−0.02**	−0.03**	−0.01**
	(−3.25)	(−3.07)	(−6.24)
Education	0.04	0.13	0.01
	(0.64)	(1.55)	(0.73)
Whether black	0.02	0.05	−0.01
	(0.44)	(0.78)	(−0.11)
Whether female	−0.08*	−0.11*	0.02
	(−2.27)	(−2.50)	(0.28)
Specific vocational preparation	−0.05**	−0.03*	−0.11**
	(5.22)	(−2.34)	(−5.98)
Occupational inconsistency	0.05*	−0.01	0.14**
	(2.20)	(−0.39)	(3.60)
Log likelihood	−12,541.64	−7,682.58	−4,270.21
(base $\log L$)	(−18,160.76)	(−12,744.32)	(−5,160.02)
Adj. likelihood-ratio index	30.88%	39.63%	17.04%
(χ^2) d.f. = 11	(11,238**)	(10,123**)	(1,178**)
Number of cases	10,372	6,798	3,574

*Significant at the 95% level. **Significant at the 99% level.

The question of whether the method of data collection has a significant effect on event history analyses is a very straightforward one with a clear formal test procedure. Under the null hypothesis of no structural difference, the likelihood-ratio statistic, $-2(\ln(L_c) - \ln(L_i) - \ln(L_d))$, where subscripts c, i and d represent combined, independent and dependent data collection, respectively, is distributed χ^2, with degrees of freedom equal to the number of parameters in the model. In our case, this statistic is 1,178 with 12 degrees of freedom and the null hypothesis is clearly and soundly rejected. Thus, method of data collection makes a big and very highly significant difference in the estimates attained for event history analysis of occupational exits in the SIPP data. Which method yields the "better" estimates, however, is a question which is not so easily addressed. It requires examination of the individual estimated effects and some judgements regarding their "reasonableness."

The independent variables included in our model can be divided into three groups according to how firm our *a priori* knowledge is about their true effects on occupational exit hazards. The first group consists of whether the month in question is a seam month, and the occupational inconsistency index. The inconsistency index is taken from Jabine and Tepping (1973) and is the proportion of variation attributable to response error. The rationale for including the inconsistency index in the model is that in other surveys certain occupations have been found to have inherently higher measurement errors than others. To the extent that this pattern of measurement errors persists in the SIPP, the exit hazards for high measurement error occupations will be biased upwards. The model presented below controls for this by including a version of the index constructed from the single digit occupation code and the coefficients given by Jabine and Tepping (Table 2, column 3). *In the absence of measurement errors in the SIPP, however, neither the inconsistency index nor the seam variable should have any effect on the exit hazards.*

The second set of independent variables consists of time and its square in the occupation. While time in occupation should affect the exit hazard rate, our priors about the precise pattern of this effect are not very strong. The final set of independent variables are the substantive measures of characteristics of the individual (age, wage, education, gender) and the occupation (specific vocational preparation).[5] These variables should affect exit hazards with higher hazards for younger low wage individuals in occupations with little specific human capital.

Perhaps the most important thing to note about the estimates of Table 9 is that overall goodness of fit of the model, as measured by the adjusted likelihood ratio index, is substantially higher for the independent method (39.6%) than for the dependent method (17.0%). This is due almost entirely, however, to the gigantic effect of the seam on the independent method data. The coefficient is 3.58 for whether the month in question is a seam month implies that odds of exiting an occupation are some 35 times ($= \exp(3.58)$) as high in seam months than in non-seam months. The corresponding effect for the dependent method is just under 7 ($\exp(1.94)$).

4.1. *Seam and inconsistency index*

The effect of this difference in the seam variable is that it dominates the model for the independent but not the dependent method data. Figures 1 through 3 present the marginal adjusted likelihood-ratio indices[6] for the three sets of predictors by method of collection. Figure 1 shows these measures of explanatory power in absolute terms for both methods, whereas Figures 2

[5] Specific vocational preparation is a measure of the amount of time required to become proficient in an occupation. It was merged via a three-digit occupation code matched with information published in National Academy of Sciences (1981).

[6] The marginal adjusted likelihood-ratio index (or marginal adjusted pseudo-R^2) is obtained via $\rho^2 = (L_u - L_r + k)/L_r$ where L_r is the log-likelihood value obtained when the "k" coefficients relating to the variables under examination are restricted to 0 and L_u is the corresponding unrestricted log-likelihood value.

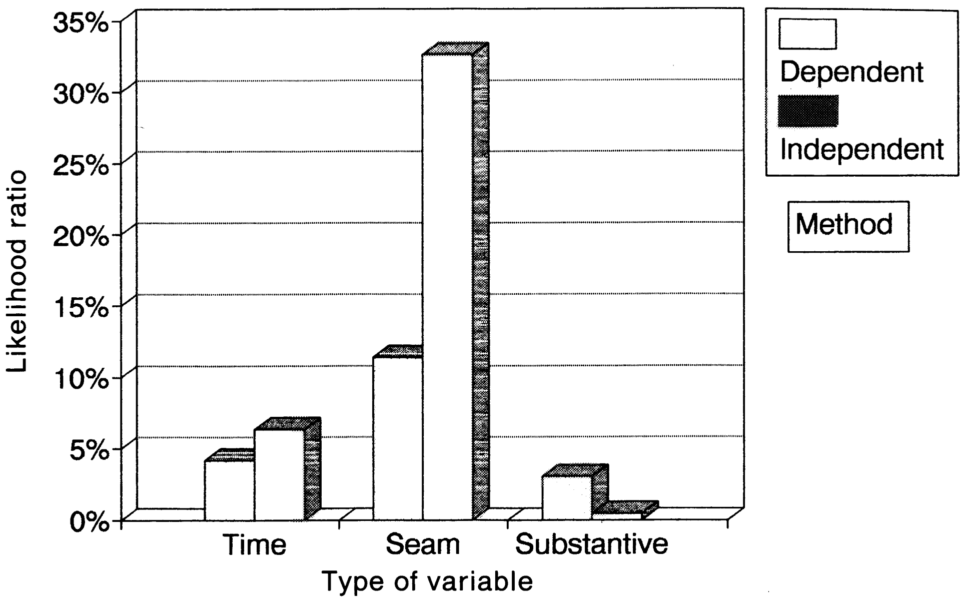

Fig. 1. Occupation event history analysis explanatory power by source and method. Adjusted likelihood ratio indices

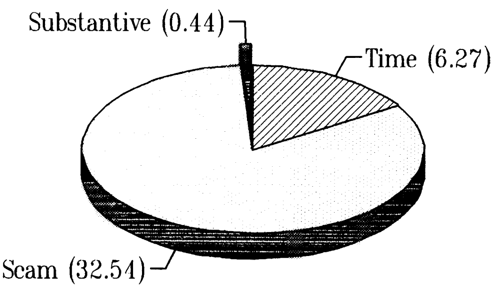

Fig. 2. Variance explained occupation hazard: independent method. Adjusted likelihood ratio indices

and 3 show them relative to the total for the independent and dependent methods, respectively. Figure 1 clearly shows that the seam measure (and the inconsistency index) in the independent method model is three times as important than in the dependent method model, and is at least six times as important as any of the other sets of predictors.

The dominance of correlates of measurement error and the comparatively puny effects of substantive variables for

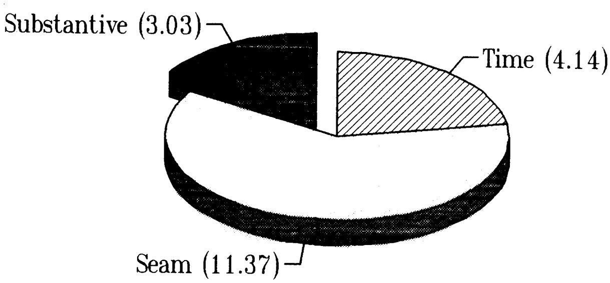

Fig. 3. Variance explained occupation hazard: dependent method. Adjusted likelihood ratio indices

occupational event history analysis using independent method data are even more dramatically seen in Figure 2. Furthermore, the improvement in the signal-to-noise ratio brought about by moving to the dependent method of collection can be easily appreciated by comparing Figures 2 and 3. While the seam and inconsistency index still account for the majority of the overall fit with the dependent method data, there is a clear improvement in the power of the substantive variables.

4.2. *Time in occupation*

The positive and significant coefficient for time in occupation, combined with the negative (and significant) coefficient for time squared for all three samples, indicates that the hazard of exiting an occupation increases at a decreasing rate with time for the first three or four months (i.e., the function $\gamma_1 t + \gamma_2 t^2$ attains a maximum at $t = \gamma_1/(2\gamma_2) = 0.63/(2 * 0.08) = 3.9$) and declines thereafter. The combined effects of time and whether the month of exit is a seam month on occupational exit hazards and the occupational survival function are illustrated for members of the third rotation group in Figures 4 and 5. The effects of the seam month are apparent from the hazard functions in that they cause large spikes at the fifth and ninth months. Especially for the independent method data, these seam spikes are so dominant as to make it difficult to discern the downward trend in the base hazard – much less its acceleration with time. The corresponding estimated survival function reflects the seam effects in their step-like shape. The survival function for the independent method data starts out at a higher level but drops well below the dependent method survival function after the first seam is experienced. The shape of the survival curve for the independent method is,

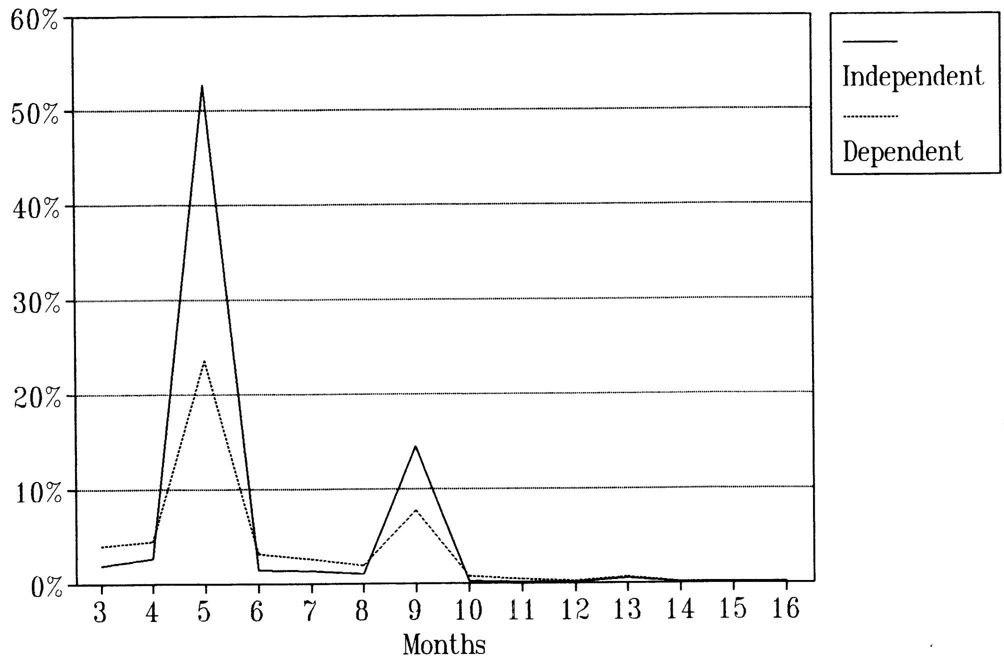

Fig. 4. Occupation exit hazard function by method of collection

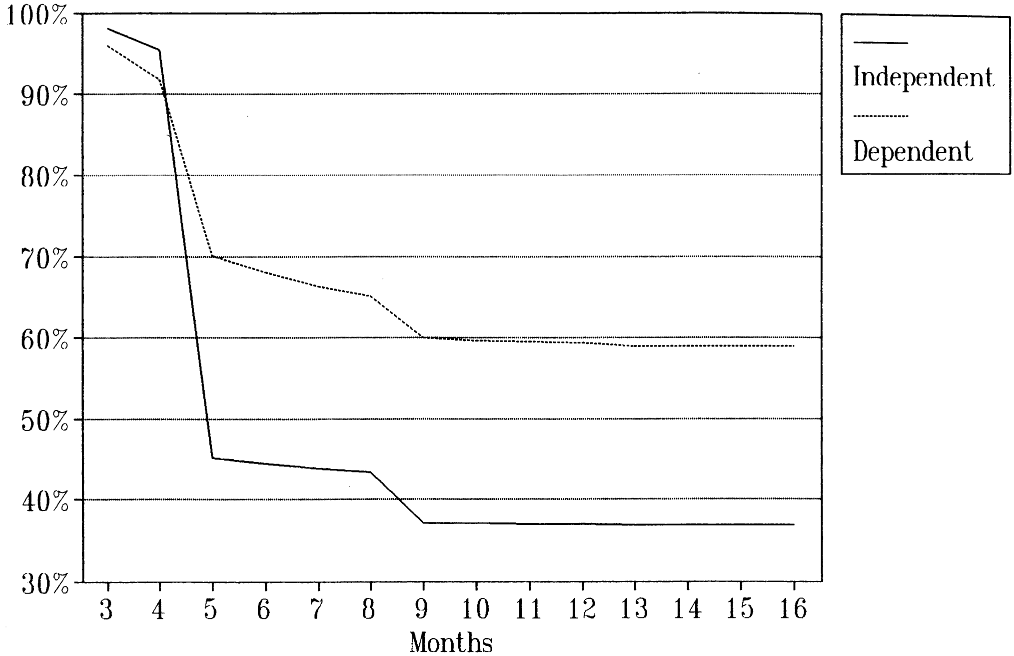

Fig. 5. Occupation survival function by method of collection

by the way, virtually identical to that obtained by Hill and Hill (1986) for unemployment exits using a Cox proportional hazards model on the 1984 SIPP panel.

4.3. *Substantive variables*

With respect to the substantive predictors, method of collection has strong effects for estimated effects of some, but not all, measures. While the *direction* of the effects of age (at the beginning of the observation period), wage and specific vocational preparation are the same for both data collection methods, their size and significance are much stronger for the dependent method data. For both collection methods, occupational exit hazards decrease with age at a decreasing rate until approximately age 45 at which point they begin rising again. This pattern is quite reasonable in that it reflects younger workers' higher risks of unemployment and older workers' higher risks of leaving the labor force. The strength of these age effects for the dependent method data is twice that of the independent. The estimated wage and specific vocational preparation effects are also consistent with our *a priori* knowledge and are much more highly significant, and in the case of the specific preparation, more powerful with the dependent method data.

The only substantive variable which is more powerful for the independent data than for the dependent is gender. There is evidence that females have lower occupational exit hazards than men in the independent data. This result seems somewhat suspect. Historically, women workers had higher exit hazards than men – a reflection of more intermittent labor force participation. By 1986, work patterns had certainly changed, with women being more likely to

Table 10. Estimated effects of occupational characteristics on occupational exit hazards by method of data collection (SRS t-ratios in parentheses)

Occupational characteristic	Independent method 1985 panel	Dependent method 1986 panel
Data complexity	−0.064** (−5.35)	−0.218** (−11.81)
People complexity	−0.057** (−4.54)	−0.171** (−8.95)
Things complexity	0.007 (0.71)	0.006 (0.41)
General educational development	−0.048** (−4.25)	−0.168** (−10.13)
Specific vocational preparation	−0.085** (−4.07)	−0.327** (−11.32)
Strength	0.120** (4.89)	0.332** (9.45)
Physical demands	0.089** (4.11)	0.245** (7.55)
Environment	0.122** (4.12)	0.298** (7.16)
Prestige	−0.046** (−5.86)	−0.135** (−11.48)
Sample size	6798	3574

*Significant at the 95% level. **Significant at the 99% level.

work even when family demands peaked, but whether these changes are enough to explain a reversal in exit hazards is doubtful.

4.4. *Effects of occupational characteristics on survival*

In addition to the specific vocational training measure, the Dictionary of Occupational Titles provides a number of other characteristics of occupations which may affect the hazard of leaving the occupation. These consist of three measures of job complexity ("Data Complexity," "People Complexity" and "Things Complexity"), a measure of general educational development (GED), two measures of physical requirements (Strength and Physical Demands) and a measure of adverse environmental exposure (Environment). The most complex occupations with respect to data are those requiring the use of calculus and statistics and the least complex require only the comparison or copying of numbers.[7] The most complex "people" occupations involve either mentoring or negotiating while the simplest jobs require taking instructions. Setting up precision machine tools is the most complex "things" activity, while feeding stock to a machine is the least.

As with specific vocational preparation,

[7] As published, the complexity scales range from low values for the *most* complex to high values for the *least*. To avoid the obvious confusion of this reverse scaling, we reverse it in our analysis so that high numbers represent high degrees of complexity.

we would expect job complexity to lead to reduced hazards of exiting. Again, the reason is that the level of complexity is positively associated with the occupation-specific human capital, and both employers and employees have productivity incentives to maintain and utilize it. Similar reasoning applies to general educational development necessary to perform adequately in an occupation.

Physical demands and adverse environmental working conditions, on the other hand, should be associated with increased hazards of exiting an occupation since they are associated with manual labor and low levels of job-specific human capital.

Finally, occupations can be rated according to their socio-economic prestige (Duncan score). Individuals, in what the designers of the scheme consider the most esteemed occupations (e.g., Sociologists), are assigned a score of 100 while those in least esteemed occupations receive a score of 1. For a variety of reasons we would expect turnover (hence exit hazards) to be lower in high than in low prestige occupations.

Not surprisingly, there is considerable collinearity in these various scales – so much so as to preclude their inclusion as a group in our event history models. We can, however, enter them singly in a specification which includes only the time-varying covariates (time, time-squared and whether a seam month) and compare their predictive power under the independent and dependent data collection methods. Table 10 presents the resulting estimates.

For all occupational characteristics other than the "Things" complexity measure (which is insignificant with either collection method), the dependent data collection method produces significantly stronger associations. The point estimates of the effects of occupational characteristics on exit hazards for the dependent method are several times as large as for the independent method. This is consistent with there being more error in occupational classification with independent collection, and this results in greater measurement error in the occupational characteristics data, with the result being attenuation of the classical sort.

In this section, we have attempted to assess the relative quality of occupational data obtained from the independent and dependent methods by examining the association of measured occupational change with exogenous variables. This "empirical validity" was found to be significantly higher for the dependent (1986 SIPP panel) method data than for the independent (1985 SIPP panel) method data. Things which, in the absence of measurement error, should not affect occupational change (e.g., whether a seam month) had smaller effects with the dependent data, while things which should be associated with change had larger effects. These differences were very highly significant. Our analysis suggests that, at least within the context of event history models, the analytic potential of SIPP occupation data was increased substantially by the move to dependent data collection methods.

5. Conclusions

In this paper, we have examined the effects of data collection method on the observed occupational and industrial change in the SIPP. We have done so using two distinct definitions of change and a combination of univariate, bivariate and multivariate methods. For all definitions of change and all methods of analysis, we find evidence consistent with the following conclusion:

1. The amount of gross change in occupations and industries is several times greater when the questions are asked

and coded each wave (independent method collection) as when they are only asked if the respondent reports a change in duties (dependent method collection); but
2. Most of the change "missed" by the dependent method collection methodology is noise; and
3. Enough of the real change is captured by the dependent method collection methods as to substantially improve the signal-to-noise ratio as indicated by the higher empirical validity of the dependent method data.

The first two parts of this conclusion are consistent with earlier research on collection method for occupation in the Current Population Survey (i.e., U.S. Department of Commerce 1975). The final part of our conclusion, however, is not consistent with the finding of the 1975 Job Mobility Study that most (of what expert occupation coders considered) "real" change in occupation is missed by dependent method collection methods. Both our conclusions here and those of the Job Mobility Study are judgmental. Our conclusion is based on the observation that occupational and industrial changes from the dependent-method data relate more closely to factors they *should* relate to (e.g., changes in hours, wages and employers, and levels of age, education and occupational characteristics) and less strongly to factors that, in the absence of measurement error, they *should not* relate to (e.g., whether the change occurred at a seam month). The judgements in the Job Mobility Study are based on whether the recorded descriptions of duties differed sufficiently in content or order of presentation to consider the two reports to be of different occupations. While the coding perspective on real change may be most appropriate for some descriptive purposes, ours is more appropriate to judging the utility of occupational and industrial change data for structural analysis.

Given these findings, we would predict that the increasingly popular practice of using dependent collection methods and prior wave information in panel surveys will result in net improvements in the empirical validity of noisy measures such as occupation and industry change. Having said this, we feel compelled to say that these improvements represent only a second best solution for the tendency of independent collection methods to overstate true change. The root cause of the problem is the lack of reliability in the questions and coding procedures used to collect the point-in-time measures and real improvements in data quality require addressing this directly.

6. References

Allison, P.D. (1984). Event History Analysis: Regression for Longitudinal Event Data. Beverly Hills, CA: Sage.

Hill, M. and Hill, D. (1986). Labor Force Transitions: A Comparison of Unemployment Estimates from Two Longitudinal Surveys. Proceedings of the Section on Survey Research Methods, American Statistical Association, 220–225.

Jabine, T.B., King, K.E., and Petroni, R.J. (1990). SIPP Quality Profile. Washington, DC: U.S. Department of Commerce.

Jabine, T.B. and Tepping, B.J. (1973). Controlling the Quality of Occupation and Industry Data. The Bulletin of the International Statistical Institute, 45, 360–392.

Kasprzyk, D., Duncan, G., Kalton, G., and Singh, M. (1989). Panel Surveys. New York, NY: John Wiley & Sons.

Lepkowski, J.M., Pennell, S.G., Miller, D.P., Luis, E., and Kalton, G. (1992). Time in Panel Effects in the SIPP. Ann Arbor, MI: Institute for Social Research, memo.

McArthur, E. (1987). Measurement of Attrition Through the Completed SIPP 1984 Panel: Preliminary Results. Internal Census Bureau Memorandum to the SIPP Nonresponse Working Group (Nov. 20).

McCormick, M., Butler, D., and Singh, R. (1992). Investigating Time in Sample Effects for the Survey of Income and Program Participation. Proceedings of the Section of Survey Research Methods, American Statistical Association, 554–559.

National Academy of Sciences (1981). Jobs, Work, and Occupations. The Fourth Edition Dictionary of Occupational Titles: Structure and Content. Washington, DC: Cain, P.S.

Short, K. and McArthur, E. (1986). Life Events and Sample Attrition in the Survey of Income and Program Participation. Proceedings of the Section on Survey Research Methods, American Statistical Association, 15–20

U.S. Department of Commerce, Social and Economic Statistics Administration (1975). Comparison of Month to Month Changes in Industry and Occupation Codes with Respondent's Report of Change: CPS Job Mobility Study. Response Research Staff Report No. 75-5.

Received November 1992
Revised August 1994