

RESEARCH REPORT SERIES  
*(Survey Methodology #2007-8)*

**Changes to the SIPP 2004 Panel Wave 1  
Reciprocity History Topical Module and Their Impacts  
on Survey Estimates and Data Quality**

Jeffrey C. Moore

U.S. Census Bureau  
Washington, DC 20233

Report Issued: January 30, 2007

*Disclaimer:* This report is released to inform interested parties of research and to encourage discussion. The views expressed are those of the author and not necessarily those of the U.S. Census Bureau.

## **Changes to the SIPP 2004 Panel Wave 1 Reciprocity History Topical Module and Their Impacts on Survey Estimates and Data Quality**

Jeffrey C. Moore

**Keywords:** CIIG, data quality, item nonresponse, measurement error, questionnaire design, spells

### **Abstract:**

The SIPP Continuous Instrument Improvement Group (CIIG) recommended a set of changes to SIPP's wave 1 Reciprocity History topical module (RHTM), several of which were intended to address known or suspected data quality issues. Most of the recommended changes were approved, and, following testing, implemented in the redesigned SIPP 2004 panel questionnaire. This report summarizes those changes, and evaluates their impact on RHTM data through a comparison with data from the most recent pre-redesign questionnaire, administered in wave 1 of the 2001 SIPP panel. The limitations of the research – in particular, the lack of an experimental design, and the absence of objective validity criteria – prevent definitive conclusions. Nevertheless, the results are strongly suggestive of important improvements in the quality of key RHTM estimates, including the presence/absence of spells of receipt prior to the wave 1 reference period and the total number of spells experienced.

# Changes to the SIPP 2004 Panel Wave 1 Reciprocity History Topical Module and Their Impacts on Survey Estimates and Data Quality

Jeffrey C. Moore

## 1. Introduction and Background

In preparation for the 2004 Survey of Income and Program Participation (SIPP) panel, the U.S. Census Bureau's SIPP Executive Committee chartered the Continuous Instrument Improvement Group (CIIG), an interdivisional committee whose initial task was to review the SIPP core instrument and recommend improvements to address both burden/efficiency and data quality issues. CIIG's recommendations for the core questionnaire were tested and refined in a series of field experiments known as the SIPP Methods Panel (Moore, et al., 2004), and many were implemented in the revised core instrument implemented in the 2004 SIPP panel.

CIIG also reviewed and recommended improvements to several of the survey's topical modules – sets of questions that focus on more specialized topics, and that are administered less frequently (e.g., generally either annually or once per panel) than the core. The Reciprocity History topical module (RHTM) was one such module reviewed by CIIG. The RHTM is designed to capture detailed information about respondents' lifetime history of participation in and receipt of benefits from three major assistance programs: "cash" welfare (hereinafter referred to by its most common acronyms, AFDC/TANF), Food Stamp program benefits (hereinafter simply "Food Stamps"), and Supplemental Security Income (SSI). For each program the module asks questions to ascertain the start date of any left-censored spell of participation (i.e., a spell identified in the core interview that is in progress at the beginning of the wave 1 reference period), whether respondents have ever participated in the program before the core survey's reference period, and, if so, the dates of their first and last receipt, and the total number of receipt "spells" experienced. The product of CIIG's review was a recommended set of improvements to the RHTM (Moore, 2002). Following a cognitive interview pretest (Hughes and Hunter, 2003), which suggested a few refinements to CIIG's recommendations, a redesigned version of the module was prepared and implemented in wave 1 of the 2004 SIPP panel.

In addition to many other more minor modifications, the primary changes to the RHTM were intended to address specific data quality issues, mainly having to do with the complete identification of the existence of spells of receipt prior to the wave 1 reference period. These included the following: (1) a repair to the universe of eligibility for entry into the entire module, removing restrictions based on an odd and inscrutable combination of core-reported receipt of the three target programs and parent/guardian status; (2) a repair to the universe for the module's AFDC/TANF questions, expanding the universe to include not just current parents/guardians, but anyone who ever was a parent or guardian; (3) the correction of an instrument error which prevented the left-censored spell questions from being asked of those who reported SSI receipt in the core; (4) a re-sequencing of the RHTM's sub-sections (putting Food Stamps last) to permit

special probes about Food Stamps receipt based on reports about the other programs; and (5) rewording to clarify the intent of key RHTM questions about the existence of prior spells. In addition, the redesigned RHTM included (6) improved questions and procedures for capturing the total number of lifetime spells.

This paper examines the impact of these changes to the SIPP 2004 RHTM module on key module estimates, through a comparison with estimates derived from the pre-redesign version of the module used in the 2001 SIPP panel. Although this evaluation lacks objective criteria for assessing data quality differences, and is also limited by the absence of a true experimental design, with the assistance of a few reasonable assumptions a good case can be made that the differences in the estimates indicate improved quality for the data from the 2004 RHTM questionnaire.

The remainder of this report is organized as follows: Section 2 presents a more detailed description of the specific changes implemented in the RHTM for the 2004 SIPP panel. Section 3 describes the data used in the analyses, as well as some of the relevant technical details concerning the analyses themselves. Section 4, the results section, comprises the bulk of the report, and includes several comparisons of 2004 panel RHTM results with corresponding data from the 2001 panel. Finally, section 5 provides a summary of major findings, offers some conclusions as to their meaning and import, and outlines possibilities for additional research.

## **2. Detailed Description of Primary RHTM Instrument Changes**

As noted above, the full set of CIIG's proposed improvements to the RHTM are outlined in Moore (2002).<sup>1</sup> For the most part, those improvements were implemented in the 2004 panel instrument, with some minor modifications based on a cognitive interview pretest (Hughes and Hunter, 2003). This section summarizes the major changes to the module, which were intended to improve the quality of key data concerning the existence of spells of receipt prior to the wave 1 reference period, the total number of lifetime spells of receipt, and the dating of the start of pre-wave-1 receipt spells.

### **2.1. Changes to Questions and Procedures Concerning the Existence of Pre-Wave-1 Spells**

#### **2.1.1. RHTM eligibility criteria**

In addition to a minimum age criterion (18 or older), the 2001 RHTM, for reasons that are unclear, also imposed other eligibility criteria which screened additional people out of the module entirely. Specifically, no module questions were asked of those who were otherwise

---

<sup>1</sup> A few additional RHTM items are embedded in the core questionnaire – see section 2.1.3. CIIG's recommended changes for these items are documented along with those for the rest of the "general income II" section of the core questionnaire in SIPP Continuous Instrument Improvement Group (1998).

eligible (i.e., were at least 18 years of age) but who

(a) did not report Food Stamps receipt in the core interview

AND

(b) did report receipt of SSI (either federal or state)

AND EITHER

(c) did report AFDC/TANF receipt

OR

(d) were not a guardian.

All inquiries into the origin of these mysterious eligibility criteria, or their purpose, went unresolved. Nor were any reasonable hypotheses offered. By consensus among all interested parties, the redesigned RHTM in 2004 eliminated all restrictions on eligibility for the module, with the sole exception of the 18+ age criterion.

### 2.1.2. AFDC/TANF question series universe

The 2001 version of the RHTM (this was also true of prior panels) only asked the questions about pre-wave-1 spells of AFDC/TANF participation of those who were currently the official guardian – generally, the mother – of a child member of the SIPP sample household. This restriction is clearly inappropriate for an inquiry into AFDC/TANF receipt history over the whole span of the person’s adult lifetime. The 2004 panel RHTM admitted into the AFDC/TANF question series all those who were the parent of another household member (regardless of the age of the “child” at the time of the interview), and added a question, EVERGARD, to identify and direct into the AFDC/TANF question series those who were not currently a parent/guardian of any household member, but who had been at some point in their lives:

>EVERGARD (2004)<

ASK OR VERIFY: Some benefit programs are designed to help needy children.

[Have/Has] [NAME/you] ever had any children, or served as a child's legal guardian?

Thus in the 2004 panel RHTM the only people who were not asked any questions about prior participation in AFDC/TANF were those who were not currently, and had never been, a parent or guardian.

### 2.1.3. SSI left-censored spell questions

As noted, the RHTM’s “left-censored spell” questions establish a start date for spells found to be in progress at the beginning of the wave 1 reference period. Due to an error in the 2001 questionnaire, however, these data were not collected for SSI spells reported in the core interview. Through a small change in background logic this error was corrected in the 2004 panel questionnaire.

#### 2.1.4. RHTM sub-section reordering

In general, according to Census Bureau subject-matter specialists, receipt of AFDC/TANF or SSI benefits automatically confers eligibility for Food Stamps. The 2004 RHTM questionnaire took advantage of this known relationship by re-ordering the sequence of questions, asking about Food Stamps last, after the AFDC/TANF and SSI questions, rather than first, as had been the case in 2001 and prior panels. This enabled the addition of special probes in 2004 for those who, for example, had reported a pre-wave-1 spell of AFDC/TANF or SSI, but who initially reported that they had never before applied for (or if they had applied, had never before received) Food Stamps. For example, those with the appropriate characteristics, and who responded “no” or “don’t know” to an initial question asking whether they had ever applied for Food Stamps, received the following probe:

>APLFS2 (2004)<

What about when [NAME/you] received [AFDC/TANF [and] / SSI] benefits before this past [MONTH1]? Did [he/she/you] apply for Food Stamps then?

A similar probe (RECVFS2) was administered to pre-wave-1 SSI and/or AFDC/TANF recipients who reported that they had applied for Food Stamps but had never received them. The probes were intended to stimulate more complete recall and reporting of prior Food Stamps spells.

#### 2.1.5. Rewording of questions about the existence of prior spells

At the beginning of the Food Stamps sub-section, the 2001 RHTM included the following question about the existence of prior (i.e., pre-wave-1) spells of receipt, asked of those who had already reported receipt of Food Stamps in the core questionnaire:

>CURFS (2001)<

Besides the Food Stamps [NAME/you] received during the last four months, have there been any other times when [he/she/you] [was/were] authorized to receive Food Stamps?

(Similar questions opened the AFDC/TANF and SSI question series – CURADC and CURSSI, respectively.) The 2004 instrument implemented changes to the design and wording of these questions in order to ensure that they were clear to respondents, and properly understood. The key design change was to automatically mark the question “yes,” without asking, if the respondent had already reported pre-wave-1 receipt in response to the left-censored spell questions embedded in the core interview. Otherwise, for receipt during the wave 1 reference period that started in wave 1, the primary change was to carefully list the reported months of receipt, and then to ask about receipt before the reference period:

>CURFS< (2004)

Earlier I recorded that [NAME/you] received Food Stamps in [list months]. Have there been any other times before this past [MONTH1] when [he/she/you] [was/were]

authorized to receive Food Stamps?

The new wording avoided the use of ambiguous phrasing (“...during the last four months”) that would cause confusion if it were perceived as indicating receipt *in* all four months, and made the main intent of the inquiry – receipt *before* the current wave’s reference period – very explicit.

## 2.2. Changes to Questions and Procedures Concerning Reports of Total Spells of Receipt

An important goal of the RHTM is to capture the total number of AFDC/TANF and Food Stamps receipt spells a person has experienced in his or her lifetime;<sup>2</sup> for this purpose the 2001 panel questionnaire used a single question of the following form:

>TMFSTIME< (2001)

How many times in all have there been when [NAME/you] received Food Stamps?

(AFDCTIME was the parallel question in the AFDC/TANF question series.) The concept of a “spell” – an unbroken period of time during which some condition or circumstance applies – is a difficult one to communicate to respondents, and research of various types has suggested that the above approach has substantial problems. Miller, Bogen, and Moore (1996), for example, used behavior coding techniques to highlight the serious difficulties experienced by both interviewers and respondents in administering and responding to TMFSTIME and AFDCTIME. In addition, Gordon Lester (HHES), for many years the Census Bureau’s primary RHTM subject-matter specialist, has remarked on the obvious anomalies in the data elicited by pre-redesign versions of the RHTM which used the above version of the number-of-spells question – in particular, the frequency of implausibly high total spells reports, numbering 50 or more in some cases.

CIIG’s proposed solution was to avoid asking the difficult “number-of-spells” question to the maximum extent possible, and instead to break up the task into smaller, clearer, and easier-to-answer components from which a total number of spells could be logically inferred. The new procedures were pretested successfully in a cognitive interview study (Hughes and Hunter, 2003), and eventually incorporated with minor revisions into the 2004 questionnaire. The two major new features of these new procedures are summarized in the following two sections.

### 2.2.1. Automatic number-of-spells entries

To simplify the reporting of total spells, a new question was added to the RHTM for people who, in the core interview, reported a left-censored spell, and a start date for that spell prior to the wave 1 reference period. The new question determines whether that left-censored spell is the person’s first-ever experience with the program in question, or whether they also participated at some even earlier time, prior to the start of the spell that has continued into the wave 1 reference period, e.g.:

---

<sup>2</sup> The RHTM does not attempt to capture the number of lifetime spells of SSI receipt.

>FSWHEN1< (2004)

Earlier I recorded that the most recent time [NAME/you] started receiving Food Stamps was [month and year]. Was that the first time [he/she/you] had EVER received Food Stamps, or had [he/she/you] been on Food Stamps before?

(ADCWHEN1 is the counterpart question in the AFDC/TANF series.) If the left-censored spell is reported to be the person's first such spell, then the number of lifetime spells is known to be 1; in such circumstances a "1" was automatically entered in the 2004 panel total spells variable without having to ask a direct question.

Under slightly different circumstances, a total of one or two lifetime spells could be logically inferred without having to ask a number-of-spells question directly. Those for whom a left-censored spell was identified in the core interview (as above), but who reported in FSWHEN1 (ADCWHEN1) having had an even earlier spell of Food Stamps receipt ("on before"), were asked to report their earliest-ever receipt date and the last receipt date before the already-identified left-censored spell began. Another question was added to the RHTM to determine whether there were any breaks in receipt between those dates:

>FS1TIME< (2004)

And [was/were] [he/she/you] on Food Stamps continuously between [first receipt date] and [last receipt date]?

(ADC1TIME is the counterpart question in the AFDC/TANF series. Note also that this question mirrors the new question included in the core questionnaire's "left-censored spells" series to help ensure that respondents report the correct spell start date – see section 2.1.) A "yes" report, indicating a continuous spell of receipt, meant that the person had a total of 2 spells – the one just reported plus the left-censored spell that extended into the wave 1 reference period.<sup>3</sup>

Those without a wave 1 core interview spell of receipt (or whose core receipt spell started in a wave 1 reference period month), but who, in the RHTM, reported some receipt prior to the reference period, were also asked to report the dates of their first and last receipt, as above, and also then directed to FS1TIME (ADC1TIME) to determine whether those dates defined a continuous spell of receipt. Again, if receipt was continuous, then the number of lifetime spells before the wave 1 reference period was known to be 1, and the total number spells could be automatically entered in the questionnaire.<sup>4</sup>

---

<sup>3</sup> In fact, in these circumstances the 2004 RHTM instrument incorrectly entered a "1" automatically as the total number of spells, ignoring the already-identified left-censored spell. This error has been accounted for in the current analyses. The point remains, however, that this was another set of circumstances under which the total number of spells was known based on the responses to other items, without having to ask the question directly.

<sup>4</sup> Here again the 2004 instrument logic was incomplete – a "1" was always entered, even if the person had reported receipt during the wave 1 reference period, meaning that the true number of spells was 2. This error has also been accounted for in the current analyses



### 2.2.2. Rewording of the number-of-spells questions (AFDCTIME and TMFSTIME)

Only the remaining group of respondents – those whose responses (or nonresponses) left their number of pre-wave-1 receipt spells uncertain – were asked a direct number-of-spells question, comparable to the one asked in 2001. However, the 2004 version of the question implemented a number of wording changes to make the intent of the question more clear – most notably, new wording to convey the concept of “going on” a program, as opposed to the number of times benefits were “received”:

>TMFSTIME< (2004)

How many separate times did [NAME/you] go on Food Stamps between [first receipt date] and [last receipt date/...] [the most recent time [he/she/you] went back on, [left-censored spell start date/...]]?

(AFDCTIME, the counterpart question in the AFDC/TANF series, was similarly reworded.) Note that, in the interests of making sure the question was clear to respondents, the new wording focuses explicitly on receipt spells *other than* those already reported in the core interview – that is, other than those that involved receipt in some or all wave 1 reference period months. For respondents who reported a spell in the wave 1 reference period, the true total number of spells is 1 plus the number reported in TMFSTIME (AFDCTIME).<sup>5</sup>

### 2.2.3. Blank number-of-spells reports due to missing start or end date information

A third change in the 2004 RHTM questionnaire which affected the number of spells reports was inadvertent. New logic was introduced in the 2004 questionnaire which caused those who said “don’t know” or refused to answer the RHTM’s “first received” or “last received” questions to skip the remaining questions for that program, including the question which captured total lifetime receipt spells. In the 2001 questionnaire such date non-responders still continued on to the number-of-spells item. These procedural differences could have important effects on differences in missing data rates to the RHTM’s number-of-spells questions.

## 2.3. Changes to Questions and Procedures Concerning Start Dates

### 2.3.1. Changes to the “left-censored spell” start date questions

As part of the 1996 conversion from paper-and-pencil to computer-assisted interviewing, the RHTM questions to capture start dates of spells found to be in progress in the first month of wave 1 – i.e., “left-censored” spells – were separated from the rest of the RHTM and embedded in the core questionnaire. The purpose of this shift in location was simply to place them in close proximity to the questions which trigger the need to ask them – the questions which reveal

---

<sup>5</sup> As above, the instrument logic was again incomplete, and did not take into consideration a spell that had already been reported in the core interview. The appropriate adjustment has been made in the analyses reported here.

receipt in the first reference period month. Despite their location, the “left-censored spell” questions are still considered to be RHTM items. In the 2001 panel, these questions were of the following form (e.g., for Food Stamps):

>FBEG120< (2001)

When did [NAME/you] apply for the Food Stamps you received?

CIIG’s review (SIPP Continuous Instrument Improvement Group, 1998) found the 2001 approach questionable on two main grounds. First, it did not really focus on or attempt to make clear to respondents the central concept of interest – the start of the spell of receipt that continued into the wave 1 reference period. And second, it offered no hint that its intent was to capture the respondent’s past receipt as an adult, and not as a child beneficiary, or as a child covered by benefits received by a parent. Thus, CIIG recommended major changes to these items, the intent of which was to clarify – and, in the case of the “as an adult” criterion, make explicit – these key concepts. This was accomplished in a series of up to four questions, as follows:

>FBEG120< (2004)

When did [NAME/you] start receiving Food Stamps?

*(if the reported date is before the person’s 18<sup>th</sup> birthday, ask FBEG120A;  
otherwise, skip to FBEG120B)*

>FBEG120A< (2004)

When did [NAME/you] start receiving Food Stamps on [his/her/your] own, or in [his/her/your] own name?

>FBEG120B< (2004)

And [has/have] [NAME/you] received Food Stamps every month since then?<sup>6</sup>

*(if “yes,” then end the series, as the spell start date has been established;  
otherwise, continue with FBEG120C)*

>FBEG120C< (2004)

When did [NAME/you] start receiving Food Stamps CONTINUOUSLY, every month [.../through [month]]?

### 2.3.2. Changes to the RHTM’s first-ever receipt date questions

The issue noted above – lack of clarity that the intended focus of the RHTM questions is on the person’s adult life – was also the motivation for new procedures in the module’s questions about respondents’ “first ever” receipt. These questions incorporated the same new logic to assess the date information initially offered by respondents when answering the “when did you first start

---

<sup>6</sup> This question, in fact, took several different forms, depending on the pattern of reported receipt during the reference period. The variant shown here was used if receipt had been reported for all reference period months.

receiving [X]” (ADCWHEN2, SSIWHEN2, FSWHEN2) question. Specifically, the 2001 and 2004 questionnaires had the same starting point, e.g.:

>FSWHEN< (2001)

When did [NAME/you] first start receiving Food Stamps?

In 2001 that was the end of the matter – the response, as long as it was not logically impossible, was simply accepted. In 2004, however, new logic was added to the questionnaire to evaluate the reported date relative to the respondent’s birthday. If the initially-reported “first receipt” date preceded the person’s 18<sup>th</sup> birthday, then a new question was administered, e.g.:

>FSWHEN<

WE WANT TO KNOW WHEN THE PERSON FIRST RECEIVED FOOD STAMPS ON HIS/HER OWN, OR IN HIS/HER OWN NAME – E.G., AS THE “HEAD” OF HIS/HER FAMILY – NOT AS A CHILD WHOSE PARENT(S) RECEIVED THEM.

When did [NAME/you] first start receiving Food Stamps in [his/her/your] own name?

(The comparable items in the AFDC/TANF and SSI question series are AFDCWHEN and SSIWHEN.) As was the case with the left-censored spell questions, the main purpose of the new procedures here was to ensure the most accurate date information possible. On occasion, however, they identified “false positives” – people who, had they responded similarly in the 2001 RHTM, would have been recorded as having had a history of receipt, but who could in 2004 be identified as “pre-adult” recipients, without any relevant (i.e., as an adult) receipt history.

### **3. Data Sources and Analysis Procedures**

The SIPP data used in the analyses described in this report derive from the preliminary “TransCASES” files produced and maintained by the Demographic Surveys Division (DSD). The TransCASES files are SAS datasets which are produced directly from the instrument output (see Finke, Downs, and Forsythe, 1999). TransCASES data are minimally edited and recoded, primarily only insofar as those manipulations are executed in the course of administering the instrument. For example, the TransCASES files contain no edited or imputed data to correct for nonresponse. In fact, unedited (or minimally edited) data are arguably the most appropriate to use for this analysis, which is focused on how people *responded* to the revised RHTM questionnaire, as opposed to the estimates of population parameters one might make from these data.

Sample design differences between the 2001 and 2004 panels, and the potential for differential impacts of nonresponse, render weighted analyses somewhat more appropriate than unweighted analyses for most of this evaluation’s purposes. The weights used in the analyses are the final weights for each panel, which include adjustments for selection probability, wave 1

noninterviews, and a second-stage adjustment to population controls (see Boies (2003) for a description of SIPP 2004 weighting procedures). For these analyses the actual weighting factor used was a “small” weight, calculated for each case as the actual weight divided by the average non-zero weight across all interviewed adults. Using the small weight in place of the actual, full weight offers the advantage of producing a weighted n that is approximately equal to the unweighted n, while leaving percent estimates unaffected. As will be demonstrated in the next section, in the end the weighted/unweighted issue is of no real consequence. Both types of results are presented in the tables, and in no case is there any noticeable difference between the weighted and unweighted results themselves or the conclusions to be derived from them.

The primary tool used to assess statistical significance is a simple t-test of the difference between two proportions. Following guidance from Steve Mack (DSMD), these tests incorporate a design effect adjustment to account for the clustering of the SIPP sample and other departures from a simple random sample design. In the analyses presented below, the preliminary SIPP 2004 design effect,  $DEFF = 1.7454$ , is used to adjust the critical value of the t-statistic in tests of statistical significance; specifically, the calculated value of t is evaluated against the standard critical value adjusted by a factor of the square root of the estimated design effect. Under these procedures, adjusted t-test critical values for various levels of statistical significance are as follows:

for $p < .10$ ,	$t \text{ (adjusted)} \geq \pm 2.173$
$p < .05$ ,	$t \text{ (adjusted)} \geq \pm 2.589$
$p < .01$ ,	$t \text{ (adjusted)} \geq \pm 3.403$
$p < .001$ ,	$t \text{ (adjusted)} \geq \pm 4.347$
$p < .0001$ ,	$t \text{ (adjusted)} \geq \pm 4.913$ .

Note that the estimates presented in this report are preliminary, and based only on actual responses to the survey. They are presented here only for the purpose of this methodological investigation, and no claims are made as to their exact concordance with more rigorously developed “official” estimates of these characteristics in the U.S. population.

#### 4. Results

This section of the report presents research evidence concerning the impacts of the RHTM questionnaire changes, outlined in section 2 above, on the data produced by the module. It focuses separately on data regarding the existence of pre-wave-1 spells of receipt, the total number of receipt spells, and on the accurate dating of the start of left-censored and first-ever receipt spells.

##### 4.1. The Effects of 2004 RHTM Instrument Changes on Pre-Wave-1 Receipt Reports

Table 1 summarizes the 2001 panel and 2004 panel results with regard to respondents’ reports of participation in the three programs of interest – AFDC/TANF, SSI, and Food Stamps – prior to

the wave 1 reference period. In addition to a total, the table provides separate estimates of those who first reported their pre-wave-1 receipt in the RHTM questions embedded in the core questionnaire, in the questions which capture the start date of left-censored spells found to be in progress in the first month of the wave 1 reference period, and those who first reported pre-wave-1 receipt in the RHTM proper. (Respondents could have reported different pre-wave-1 receipt episodes in both the core and the module; if so, they are only tallied in the “core” row of the table, which shows where pre-wave-1 receipt was *first* identified.) For all three programs the 2004 estimates of pre-wave-1 receipt are significantly higher than those obtained in 2001. While there is no definitive evidence that *more* reporting of such events is *better* reporting, it is certainly an intuitively appealing conclusion (under the reasonable assumption that forgetting is a major source of error in the reporting of past events), and one that is in line with the expected impact of most of the changes implemented in the 2004 redesign.

Below I focus briefly on a few of the most obvious noteworthy features of the results for each program separately; subsequent sections will address related results and the presumed causes of these differences in more detail.

AFDC/TANF The 2004 estimate of the proportion of adults age 18+ with pre-wave-1 receipt of AFDC/TANF, 4.0%, is almost 75 percent higher than the 2001 estimate, 2.3%. It is quite clear that this difference is driven solely by the difference in the reporting of past receipt episodes elicited by the “main” RHTM questions, since the 2001 and 2004 proportions first identified in the left-censored spell questions embedded in the core questionnaire are identical. This set of results makes sense given the absence of any important changes for the items embedded in the core, and the major changes implemented in the RHTM proper – in particular, the opening up of the module to all persons who had ever had a child or been a child’s guardian, as opposed to limiting it to those who were currently the parent/guardian of a young child.

SSI The estimated proportion of adult recipients of SSI before the wave 1 reference period is almost three times higher in 2004 (3.1%) than it was in 2001 (1.1%). Here the cause of the increase clearly resides primarily in the core-embedded left-censored spell questions, which, as noted, were completely missed in 2001 due to instrumentation errors, but which were captured as intended in 2004.

Food Stamps In 2004, 10.0% of the interviewed adult (18+) sample reported having experienced at least one episode of Food Stamps receipt prior to the wave 1 reference period, a small but statistically significant increase over the 9.1% estimate from the 2001 panel. Although the differences are not nearly as stark as with the other two programs, both components of the reporting of prior spells appear to have played a role in the overall difference between the 2001 and 2004 results.

#### 4.1.1. The impact of changes to RHTM eligibility criteria

As noted above, the 2001 RHTM imposed a set of eligibility criteria for the module questions

which were eliminated in the 2004 redesign. Specifically, in 2001 no module questions were asked of those who were otherwise eligible (i.e., were at least 18 years of age) but who

(a) did not report Food Stamps receipt in the core interview

AND

(b) did report receipt of SSI (either federal or state)

AND EITHER

(c) did report AFDC/TANF receipt

OR

(d) were not a guardian.

Table 2 summarizes the impact of these screening criteria on the estimates derived from the RHTM – see the row labeled “Adjusted Data for 2004,” in which the 2001 eligibility criteria have been imposed on the 2004 data after the fact. Had those criteria been retained in 2004, 1,145 cases would have been excluded from the analysis. The adjusted results make it quite clear that for all three programs the 2001 eligibility criteria artificially suppressed the rate of reported pre-wave-1 receipt. In each case the rate of reported pre-wave-1 receipt among the would-have-been-screened-out cases significantly exceeds that of the remaining group of respondents; the impact is especially large for SSI. The elimination of the eligibility criteria in 2004 does not tell the whole story about the 2001-2004 differences in pre-wave-1 receipt, however, since in all three cases the adjusted results in 2004 – presumably from a set of respondents with parallel characteristics to those who responded to the RHTM in 2001 – still show a significantly higher rate than was observed in the prior panel. Below I summarize some highlights of the results for each program individually.

AFDC/TANF Of those who would have been screened out of the 2004 RHTM questions had the 2001 eligibility criteria been applied, about 9% reported pre-wave-1 receipt of AFDC/TANF benefits. This is a significantly higher rate of receipt than that exhibited by the remaining set of 2004 respondents, who would not have been excluded from the module under the 2001 rules. However, because the number of affected respondents is rather small, the impact of the screening procedures on the overall 2004 estimate is trivial: the adjusted estimate for 2004, 3.9%, does not differ from the 4.0% estimate derived from the full 2004 sample, and there is no change in the strength of the relationship between the 2001 and 2004 results – the adjusted estimate for 2004 remains significantly greater than the observed figure (2.3%) for 2001. Perhaps the main conclusion to draw from these results is that the removal of the 2001 RHTM screening procedures in the 2004 redesign is not a major factor in the increased (and presumably improved) reporting of pre-wave-1 spells of AFDC/TANF receipt in 2004.

SSI Almost three-quarters (73.8%) of the respondents who would have been screened out of the RHTM in 2001, but who were eligible for the module in 2004, reported pre-wave-1 SSI receipt – enormously higher than the 2.2% rate among those 2004 respondents whose characteristics would have made them eligible for the module in 2001. As a result, the 2004 adjusted estimate is significantly lower than the observed estimate for the full 2004 sample (3.1%), although even

that lower figure still significantly exceeds the rate observed in 2001. Thus, the evidence points strongly toward the conclusion that the 2001 eligibility criteria artificially suppressed the reporting of pre-wave-1 SSI receipt episodes, and that the 2004 data are substantially improved as a result of the elimination of those criteria.

Food Stamps The Food Stamps results most resemble those for AFDC/TANF. The reported rate of pre-wave-1 Food Stamps receipt among those 2004 respondents who would have been ineligible in 2001 (29.9%) is three times higher than the rate among the rest of the 2004 sample (9.7%). Even excluding these artificially ineligible cases from the 2001-2004 comparison, however, is not enough to erase the increase in reporting in 2004. This fact again points to the conclusion that the elimination of the eligibility restrictions, while an obviously important factor in explaining the difference between the 2001 and 2004 panel results, does not account for all of the difference.

#### 4.1.2. The impact of changes to the AFDC/TANF question series universe

As noted above, prior to the 2004 panel the RHTM only asked questions about pre-wave-1 spells of AFDC/TANF participation of those who were currently the designated guardian<sup>7</sup> of a young child in the household. This restriction – clearly inappropriate if the measurement goal is to capture – was lifted in the 2004 redesign, resulting in the opening up of the AFDC/TANF question series to anyone who had ever been a parent or guardian.

The change was implemented in part by a simple change in logic – instead of admitting only current guardians into the AFDC/TANF question series, the redesigned RHTM admitted anyone who was the parent of another household member (regardless of the age of the “child” at the time of the interview). This change resulted in the administration of the AFDC/TANF questions to an additional 17,956 respondents in 2004 who would have been screened out of the series in 2001 – namely, those who were not the designated current guardian of anyone in the household, but who were the parent of a household member. Among this group, the RHTM questions elicited 580 new reports of pre-wave-1 AFDC/TANF receipt.

The more complex change to the module involved adding a new question, EVERGARD (see section 2.2.), to identify people who, although not currently a parent/guardian of any household member, had been a parent/guardian at some point in their lives. Among the 45,045 respondents to whom EVERGARD was administered, 17,948 (about 40%) said “yes,” they had been a parent or guardian, of whom 1,055 eventually reported a spell of AFDC/TANF receipt prior to the wave 1 reference period.

---

<sup>7</sup> “Guardian” status is assigned in SIPP primarily for the purpose of deciding who will answer the survey’s occasional questions about children under age 15 (regarding health insurance coverage, for example). If a child has a mother in the household, she is designated as guardian automatically; absent a mother, the father will be designated as the guardian; absent both parents, the guardian is identified by a question which asks the name of the person in the household “who is responsible for” the child.

Thus, in 2004, an additional 1,635 (1,055 + 580) reports of pre-wave-1 AFDC/TANF receipt were captured from respondents who otherwise did not report any such spell in the core questions, and who would have been skipped out of the RHTM's AFDC/TANF question series in 2001. This represents close to half of the 3,604 people identified in 2004 as having had a pre-wave 1 spell of AFDC/TANF receipt.

The impacts of the 2001 screening procedures are shown in the "Projected Data for 2004" row of Table 3, which recodes as "skipped" all 2004 cases which would not have met the 2001 criteria for inclusion in the AFDC/TANF question series. (For example, it assumes that the additional 1,635 "yes" reports would not have been captured, and so subtracts them from the original total of 3,604, leaving only 1,969.) The essential outcome of this exercise is that under the 2001 screening rules the 2004 estimate of pre-wave-1 receipt (2.1%) would have been quite close to that obtained in 2001 (2.3%). (In fact, statistical analysis suggests that the 2001 estimate significantly exceeds the "projected" 2004 estimate.) Clearly, the 2001 requirement that only current SIPP-defined "guardians" be asked the AFDC/TANF reciprocity history series had a large dampening effect on the identification of episodes of AFDC/TANF receipt before the wave 1 reference period.

#### 4.1.3. The impact of changes to the SSI left-censored spell questions

The 2004 redesign repaired a 2001 instrument error which prevented the left-censored spell questions from being asked of people who reported SSI receipt in month 1 of wave 1. Thus, no 2001 respondent had a chance to report pre-wave-1 SSI receipt at that point in the interview, and, as shown in Table 1, none did. Following the instrument repair, 1.8% of 2004 respondents reported pre-wave-1 receipt at this point. This factor accounts for almost all of the 2.0 percentage point difference between the two panels in the overall rate at which respondents reported any pre-wave-1 SSI receipt, including both the embedded-in-core questions and those administered within the RHTM module proper.

#### 4.1.4. The impact of re-ordering the RHTM sub-sections

The 2004 RHTM was modified to exploit the fact that receipt of AFDC/TANF or SSI benefits automatically confers eligibility for Food Stamps. The sequence of questions was reordered, putting Food Stamps last, and new probes were added when those who reported pre-wave-1 receipt of one of the other programs failed to report Food Stamps as well. (The new probes are described in detail in section 2.4.) The new probes were administered in 2,222 interviews, in which they elicited 140 reports of pre-wave-1 Food Stamps receipt that would otherwise have gone undetected. As shown in the "adjusted" results portion of Table 4, the re-positioning of the Food Stamps questions, and the addition of the new probes, did not have a major effect on the 2004 data, resulting only a very small boost to the overall proportion of 2004 cases with a prior spell of Food Stamps receipt. Clearly, this set of changes to the questionnaire does not explain the significant increase in pre-wave-1 Food Stamps spells observed in 2004.



#### 4.1.5. The impact of the rewording of questions about the existence of prior spells

As noted in section 2.5., the RHTM redesign in 2004 included new procedures which automatically entered a “yes” in the initial question about pre-wave-1 receipt if the respondent had already reported such receipt in response to the left-censored spell questions embedded in the core questionnaire.<sup>8</sup> When the question was actually administered, new wording was used in an attempt to avoid ambiguous phrases and clarify the main intent. Unlike the other questionnaire changes having to do with the existence of pre-wave-1 spells, it is not possible to manipulate the data to assess even roughly the independent contribution (if any) of these changes to the increased reporting of such spells in the 2004 panel.

#### 4.2. The Effects of 2004 RHTM Instrument Changes on Total Spells of Receipt Reports

A second major goal of the RHTM is to capture the total number of AFDC/TANF and Food Stamps receipt spells each respondent has experienced in his or her lifetime. The primary changes implemented in the 2004 RHTM to improve reporting of total spells were, first, to avoid asking the direct (and difficult) “number-of-spells” question to the maximum extent possible, and instead to break up the task into smaller, clearer, and easier-to-answer components from which a total number of spells could be logically inferred; and second, if a direct question could not be avoided, to re-word the question for increased clarity. (The new procedures are described in more detail in section 2.2.) The following sections summarize evidence concerning the data impacts of those changes and one other.

The number of spells results, among respondents with any reported receipt in either the core interview or the RHTM, are presented in Table 5. With one exception, the impacts of the two changes noted above on these results are not separable, and will be discussed together in section 4.2.2., below. The exception has to do with the frequency with which respondents’ number of spells reports were entered automatically, which is discussed in section 4.2.1.

##### 4.2.1. The frequency of “automatic” number-of-spells entries

As shown in Table 5, about three-quarters of the number of spells reports in the 2004 panel were recorded automatically; this is in marked contrast to the 2001 results, in which only 3.5% of AFDC/TANF spell totals and 19.2% of Food Stamps spell totals were recorded automatically. In fact, the “automatic” recording of 2001 spell totals only occurred for the purposes of the present analysis – I assigned a value of 1 to respondents who, in the core interview, reported receipt during the wave 1 reference period, and who replied “no” when asked the opening RHTM question about other, pre-wave-1 receipt (and who thus skipped the number of spells question).

---

<sup>8</sup> It was not done automatically in the 2001 questionnaire, but the present analysis assigns pre-wave-1 receipt status to 2001 respondents who reported a left-censored spell start date outside the wave 1 reference period. In some cases this action over-rode a “no” response to the initial RHTM question asking about any “other” receipt, other than that which included one or more reference period months.

In addition to that “automatic” scenario, the 2004 questionnaire automatically assigned spell totals of 1 or 2 to respondents who reported having experienced only one continuous spell wholly outside the wave 1 reference period, and/or one continuous spell that started outside the wave 1 reference period but continued into it.

Not asking the complex and difficult number of spells question clearly represents a reduction in burden. It also seems likely that automatically-entered “reports” that are logically derived from simpler questions are more likely to be accurate than reports generated in direct response to the number of spells question. On both counts the advantage is clearly on the side of the 2004 redesign.

#### 4.2.2. The impact of “automatic” number-of-spells entries and rewording the number-of-spells questions

The data in Table 5 do not in general provide definitive evidence of improved quality in 2004 compared to 2001, but there is much that is suggestive. The consistency of the results across the AFDC/TANF and Food Stamps programs adds weight to the argument that they display real trends, and not anomalies. The differences between the two panels are particularly marked with regard to reports of implausibly high numbers of spells, which have long been assumed to reflect respondents’ fundamental misunderstanding of the response task as one requiring them to report the number of receipt events (payments), rather than the number of spells. Thus, it is suggestive of higher quality that maximum values in 2004 are far lower than those observed in 2001, and similarly that the proportion reporting an unusually large number of spells (10 or more) in 2004 is about one-tenth – or less – of the 2001 estimate. Correspondingly, the proportion of recipients reporting only 1 spell, or 3 spells or fewer, is substantially higher in the 2004 results than it was in 2001. These differences combine to yield number of spells averages in the 2004 panel (1.3 for both AFDC/TANF and Food Stamps) that are about half of the corresponding estimates from the 2001 panel (2.9 and 2.6, respectively). Again, although the evidence is not definitive, there is good reason to believe that in this instance “less is more” – that many reports from the earlier panel severely overstated the true number of spells, and that the lower numbers reported in 2004 represent an improvement in accuracy.

Other indicators also suggest improved quality. “Suspicious” reports (those divisible by 10 or 12) are 10 to 15 times more frequent in 2001 than in 2004, and nonresponse due to “don’t know” or refusal, which often signals a misunderstood response task, or one that is perceived to be too difficult, is 3 to 4 times more frequent in 2001 compared to 2004.

#### 4.2.3. Blank number-of-spells reports due to missing start or end date information

An inadvertent change introduced in the 2004 questionnaire caused those who said “don’t know” or who refused to respond to the RHTM’s “first received date” or “last received date” questions to skip the remaining questions for that program, including the question which captured total lifetime receipt spells. In the 2001 questionnaire such date non-responders still continued on to

the number-of-spells item. As shown in Table 5, the 2004 procedures affected 326 AFDC/TANF recipients, and 841 Food Stamps recipients. It does not appear, however, that this additional nonresponse in 2004 had any noticeable impact on the differences between the 2001 and 2004 results. This is demonstrated by artificially adjusting the 2001 results after-the-fact to make them comparable to the 2004 data – by assigning a missing value to the number of spells report for all 2001 cases in which there was a date question nonresponse.

The same row of Table 5 which shows the 326 and 841 date non-responders in 2004 (who skipped the number of spells question) also shows that in 2001 196 AFDC/TANF and 781 Food Stamps recipients nonresponded to one or both of the “first received” and “last received” date questions (but did not skip the number of spells question).<sup>9</sup> As shown in the subsequent row of the table, removing the 2001 non-responders from the analysis – treating them as if they had been subject to the same skip logic used in 2004 – has virtually no impact on the 2001 estimates. In part, this is a function of the fact that the potential impact of these cases is actually less than appears to be the case on the surface – many of those who non-responded to one of the date questions (for each program, about one-quarter of the date nonresponders) already had a missing number of spells report, since they had nonresponded to the number of spells question, too. In addition, among those who did report a number of spells, the average number of spells reported was only slightly less than for the 2001 reporters as a whole (data not shown). As a result, the “adjusted” 2001 estimates are identical (within rounding) to the observed results for the whole sample. Thus, the reduction in the number of spells reported by 2004 respondents compared to 2001 respondents is not due, to any important extent, to the 2004 instrument bug which caused an increase in missing data.

#### 4.3. The Effects of 2004 RHTM Instrument Changes on Spell Start Date and First-Ever Receipt Date Reports

##### 4.3.1. The impact of changes to the “left-censored spell” start date questions

As noted in section 2.3.1., several changes were implemented in 2004 in the questions designed to obtain start dates for spells found to be in progress in the first month of wave 1 – i.e., “left-censored” spells. These changes involved (a) new, more precise wording of the basic question; (b) the addition of new logic and probes to ensure that the receipt episode being discussed was an in-scope, “as an adult” episode; and (c) new questions to verify that the reported date was truly the beginning of the spell of interest. The basic results with regard to the use of the new probes, and their impact on left-censored spell start date reporting, are shown in Table 6; see the top rows within the separate sections of the table for each program. Highlights of the individual findings for each program are as follows:

AFDC/TANF As shown in the top half of the AFDC/TANF section of Table 6, the new

---

<sup>9</sup> Note that the likelihood of nonresponse to the date questions is roughly equivalent across both programs and across both panels.

procedures resulted in the “repair” of the initially-reported start date in 5 cases, or about 5% of the total number of cases reporting a left-censored spell. As a result of these repairs, at the end of the process there were no 2004 cases in which the left-censored spell started when the respondent was suspiciously young (age 16 or under), and thus possibly reporting on an out-of-scope receipt episode.<sup>10</sup> This is perhaps an improvement over the 2001 results, but even without the additional questions and procedures the reporting of “suspiciously young” start dates in 2001 was quite rare (2 cases; 2.5% of all cases).

SSI The new procedures resulted in revised reports for 118 cases, or about 7% of the initially-reported start dates for left-censored SSI spells. Ultimately, about 3% of the final date reports were of the “suspiciously young” variety. Clearly, the new procedures resulted in improved 2004 spell start date data relative to what would have been the case without them. However, because of the 2001 instrument error which resulted in no questions being administered about left-censored SSI spells, there is no basis for assessing whether the new procedures resulted in any data quality improvement over the earlier results.

Food Stamps About 8% of the initially-reported start dates of left-censored Food Stamps start dates were revised as a result of the new procedures, leaving slightly less than 1% of the 2004 start dates indicating that the respondent/recipient’s age at the start of the spell was 16 or less. Significance testing finds no difference, however, between the 2004 estimate (0.9%) and the 2001 estimate (1.4%) obtained without the benefit of the new procedures.

Combined Summary The Table 6 results summarized above suggest that the changes to the 2004 RHTM’s left-censored spell start date questions may have had a modest positive impact on the quality of the resulting reports. Almost certainly the 2004 results were improved compared to what would have been the case had the originally-reported date been allowed to stand – see, especially, the results for SSI and Food Stamps, where several dozens of initially reported dates were revised following one or the other of the new probes.

But there is also some suggestion in Table 6 that the 2004 repairs fixed problems that were more extensive in 2004 than they had been in the earlier panel – note, for example, the Food Stamps results, where 90 (67 + 23) respondents (3.4%) initially reported a “suspiciously young” start date in 2004, compared to only 22 respondents (1.4%) in 2001. The 2001 approach, with its focus on “applying for” the benefits in question, appear to have elicited fewer “suspiciously young” reports to begin with than did 2004’s “start receiving” wording, in which case the 2004 repairs were to some extent merely making up for lost ground due to the negative effects of the revised question wording. This is perhaps going beyond the data, however, especially given the rough nature of the standard used here to assess ultimate quality. Additionally, although the 2001 question may have done a better job communicating the “as an adult” concept, it still was

---

<sup>10</sup> Here I use age 16 as an arbitrary cut-off between receipt that is likely to have been received as a child (and thus not intended to be in-scope for the RHTM), versus receipt that is likely to have been received “as an adult,” or “in your own name.”

unlikely to elicit accurate spell start dates because of the likely lag between “applying for” benefits and actually receiving them. And finally, the 2004 reports no doubt benefitted from the extra care to assure that the reported date was the start of the spell in question, and not some earlier spell – such errors went unexamined and undetected in 2001.

#### 4.3.2. The impact of changes to the RHTM’s first-ever receipt date questions

The RHTM questions about respondents’ “first ever” receipt also suffered from the same lack of clarity as the left-censored spell questions with regard to their “on your own/as an adult” focus, and as a result the same new procedures as those described above were also applied in this section of the module. One impact of the new procedures was the occasional identification of “false positives” – people who, had they responded similarly in the 2001 RHTM, would have been recorded as having had a history of receipt, but who, through the new probes, were identified in 2004 as child recipients only, without any relevant (i.e., as an adult) receipt history. As noted in Table 1, the new probes resulted in the exclusion of 17 pre-wave 1 AFDC/TANF false positive cases, 26 SSI cases, and 70 Food Stamps cases.

The primary purpose of the new logic and new probes, however, was to improve the quality of respondents’ first-ever receipt date reports by helping ensure that the receipt episodes reported about represented “adult/on your own” receipt. These results are also summarized in Table 6 – see the lower part of each separate section of the table for each program type. Again, the new procedures seem to have produced modest gains in quality, resulting in repairs to reported dates in 2.5% of the AFDC/TANF cases, 4.1% of the SSI cases, and 1.7% of the Food Stamps cases. In each case, the 2004 final estimate of the rate of first-ever receipt at an age of 16 or under is less than its 2001 counterpart, suggesting that the new procedures were more successful at preventing receipt as a child from slipping through. The differences are quite small, however, and only achieve statistical significance in the case of SSI.

## 5. SUMMARY AND CONCLUSIONS

SIPP’s 2004 RHTM questionnaire included many changes to improve the reporting of the module’s key data items – the presence of receipt of AFDC/TANF, SSI, and Food Stamps before the wave 1 reference period, the total number of receipt spells, and the earliest start date for receipt of these programs. The design of the research to evaluate the questionnaire changes does not support definitive conclusions, but the weight of the evidence suggests that on all three fronts they had positive effects, albeit to varying degrees.

### 5.1. Pre-Wave-1 Receipt Reports

Perhaps the clearest evidence of improved reporting in 2004 is with regard to respondents’ reports of any pre-wave-1 receipt of any of the three target programs. For all three programs, the rate of reported pre-wave-1 receipt was substantially and significantly higher in 2004 than it was

in 2001. Of course, whether the “more reporting” in 2004 is the right amount of reporting, or still too little, or too much, remains unknown. However, the basic assumption that more reporting signals better reporting is rather easily justified by the fact that many of the changes implemented in the new panel were explicitly designed to remove restrictions which had in earlier panels prevented some respondents from reporting their past receipt.

For example, reporting of pre-wave-1 AFDC/TANF receipt – which increased from 2.3% of all respondents in 2001 to 4.0% in 2004 – benefitted somewhat from the removal of unnecessary screening criteria which had limited access to the entire module. By far the more important change, however, to which almost the entire 2001-2004 difference is attributable, was the opening of the AFDC/TANF question series to anyone who had ever been a parent or guardian (as opposed to just current guardians). This change clearly permitted many respondents to report past receipt episodes who would not even have been asked about past receipt in 2001.

The removal of the unnecessary RHTM screening criteria also allowed more reports of pre-wave-1 SSI receipt, which almost tripled from 2001 (1.1%) to 2004 (3.1%). In fact, because of the specific nature of the old screening criteria, their removal in 2004 had a larger impact on SSI than on either of the other programs. But an equally important (if not more important) cause of the SSI increase, was the repair to the questionnaire logic which had prevented the core-embedded questions about left-censored spells from even being asked. Not surprisingly, when the appropriate change was made in the 2004 questionnaire, and the left-censored questions were asked, they elicited quite a number of pre-wave-1 receipt reports.

Pre-wave-1 Food Stamps receipt reports show a smaller gain from 2001 (9.1%) to 2004 (10.0%) than the other programs, in both absolute and proportional terms; in addition, unlike the other programs, here there is no clear “smoking gun” explanation for the cause of the increase. As with AFDC/TANF, there appears to have been a modest boost to Food Stamps reporting due to the removal of the RHTM screening criteria, and a similar boost due to the re-ordering of the RHTM sub-sections, which allowed the addition of several new probes exploiting the links between eligibility for the other programs and eligibility for Food Stamps. However, even the combination of these two changes only explains about half of the observed effect. One possible additional factor which might have played a role in the Food Stamps increase is the rewording of the key RHTM question asking about pre-wave-1 receipt, presumably clarifying the intent of the question. Perhaps, in the absence of any other changes with major implications for Food Stamps reporting, this more subtle change had a greater opportunity to affect response than was the case for AFDC/TANF or SSI. This is, of course, pure conjecture.

## 5.2. Number of Spells Reports

Historically, suspect data quality in respondents’ reports of the total number of receipt spells they have experienced has been evidenced by implausibly high reports, indicating (supposedly) multiple dozens of receipt episodes. HHES subject-matter specialists have expressed the view that even 10 lifetime spells would be a great many, and have speculated that the cause of the

tendency to overreport so enormously is respondents' misunderstanding of the response task – they thought they were being asked to report the number of payments received, not the number of continuous *spells* of payments.

To counter this problem, generic changes were implemented with regard to both AFDC/TANF and Food Stamps spells. The primary change was to try to avoid asking the “number-of-spells” question directly, asking instead simpler questions from which total spells could be inferred; and second, if a direct question could not be avoided, to re-word the question for increased clarity. These procedures appear to have had the desired effect – extreme reports in 2004 were greatly reduced from the levels observed in 2001, as was nonresponse, as were spell reports that gave evidence of other types of problems. As a result, about 95% of those who reported any spells in 2004 reported 3 or fewer (vs. about 80% in 2001), and the average number of spells reported was cut in half – from 2.9 in 2001 to 1.3 in 2004 for AFDC/TANF; and from 2.6 to 1.3 for Food Stamps. Of course, absent any validating information, the only conclusion that can be offered with some confidence is that the results moved in the right direction – whether they moved just the right amount, or too far, or too little, is impossible to assess with the current data.

### 5.3. Reporting of Left-Censored Spell and First-Ever Receipt Start Dates

Because the evaluation criteria are not strong, it is somewhat more difficult to draw conclusions concerning the impacts of questionnaire changes aimed at improving the reporting the start dates for various events – specifically, the start dates for spells which continued into the wave 1 reference period, and for respondents' first-ever receipt experiences. In general, though, it appears that the changes produced small but positive effects. In part, the modest level of success is a function of the fact that the problems to be solved turned out to be rather modest themselves – not very many respondents appear to have needed help in reporting an appropriate start date relative to their age, and the new probe to verify that the supposed spell start date was indeed the start of a continuous period of receipt was verified 95% of the time, or more. On the other hand, across all three programs, both of the new probes generated at least some repairs to both types of spell start dates (left-censored and first-ever). Thus, the very least that can be said is that in those individual cases there is some certainty that the data were improved.

### 5.4. Future Research

The absence of a true experimental design places important limits on the conclusions to be drawn from this research. Perhaps its most glaring shortcoming, however, is the absence of objective criteria which would permit real understanding of the data quality impacts of the 2004 questionnaire changes. We can make educated guesses about the meaning of observed differences between the 2001 and 2004 panel data, but without those criteria it is very difficult to rise above the educated guess level. Adding an administrative record component to this research – even if only for one of the three programs – would add tremendously to the information value of the results.

## References

- Boies, J. (2003), "SIPP 2004+: Cross-Sectional Weighting Specifications for Wave 1," unpublished U.S. Census Bureau memorandum from A. Tupek to C. Bowie, June 3, 2003.
- Finke, S., Downs, R., and Forsythe, S. (1999), "TransCASES – The Definitive CASES to SAS Processing System User's Guide," version 3.02, Washington, DC: Demographic Surveys Division, U.S. Census Bureau, May 26, 1999.
- Hughes, K., and Hunter, J. (2003), "Results and Recommendations from the Cognitive Pretesting of the 2003 [sic] SIPP Reciprocity History Module," Washington, DC: U.S. Census Bureau, Study Series (Survey Methodology #2003-15), issued July 26, 2003.
- Miller, E., Bogen, K., and Moore, J. (1996), "Documentation of the SIPP Redesign Pretest and Content Test Wave 1 Behavior Coding Project," unpublished U.S. Census Bureau report, August 8, 1996.
- Moore, J. (2002), "Recommended Improvements to the Reciprocity History Topical Module," unpublished U.S. Census Bureau memorandum to the SIPP Executive Committee (with attachment), January 31, 2002
- Moore, J., Pascale, J., Doyle, P., Chan, A., and Griffiths, J. (2004), "Using Field Experiments to Improve Instrument Design: The SIPP Methods Panel Project," Chapter 10 in Presser, S., Rothgeb, J., Couper, M., Lessler, J., Martin, E., Martin, J., and Singer, E. (eds.), Methods for Testing and Evaluating Survey Questionnaires, New York: John Wiley & Sons, 189-207.
- Moore, J., Stinson, L., and Welniak, E. (2000), "Income Measurement Error in Surveys: A Review," Journal of Official Statistics, 16: 331-361.
- SIPP Continuous Instrument Improvement Group (1998), "Summary of Recommendations for Revisions to the 'General Income' Receipt (Part 1) and Amounts (Part 2) Sections of the Wave 1 SIPP Instrument," unpublished U.S. Census Bureau memorandum to the SIPP Executive Committee (with attachment), September 21, 1998.



**TABLE 1: Estimates from the Reciprocity History Topical Module (RHTM) of Pre-Wave-1 Receipt of AFDC/TANF, SSI, and Food Stamps, Unweighted and Weighted, in the 2001 and 2004 SIPP Panels**

	2001			2004		
	n	% (unweighted)	WTD %	n	% (unweighted)	WTD %
TOTAL N – interviewed adults 18+	65630			81015		
<b>AFDC/TANF –</b>						
<b>Any receipt pre w1m1?</b>						
Yes – first reported in core	80	0.1	0.1	101	0.1	0.1
<u>Yes – first reported in RHTM</u>	<u>1,629</u>	<u>2.5</u>	<u>2.5</u>	<u>3,503</u>	<u>4.3</u>	<u>4.4</u>
<b>Yes – TOTAL</b>	1,709	2.6	2.7	3,604 <sup>2/</sup>	4.4	4.5
No <sup>1/</sup>	12,240	18.7	19.1	75,804	93.6	95.5
skipped AFDC/TANF Qs <sup>2/</sup>	50,274	76.6	78.3	–	–	–
screened out of RHTM <sup>3/</sup>	1,135	1.7		–	–	
(missing) <sup>4/</sup>	(272)	(0.4)		(1,607)	(2.0)	
<b>T-TEST RESULTS, 2001 vs. 2004</b>	% Yes (weighted):		2.3% vs. 4.0%	t = 17.93	p<.0001	
<b>SSI (Fed or state) –</b>						
<b>Any receipt pre w1m1?</b>						
Yes – first reported in core	0 <sup>6/</sup>	0	0	1,626	2.0	2.0
<u>Yes – first reported in RHTM</u>	<u>732</u>	<u>1.1</u>	<u>1.2</u>	<u>1,163</u>	<u>1.4</u>	<u>1.5</u>
<b>Yes – TOTAL</b>	732	1.1	1.2	2,789 <sup>2/</sup>	3.4	3.5
No <sup>1/</sup>	62,048	94.5	98.8	76,603	94.6	96.5
screened out of RHTM <sup>3/</sup>	1,142	1.7		–	–	
(missing) <sup>4/</sup>	(1,708)	(2.6)		(1,623)	(2.0)	
<b>T-TEST RESULTS, 2001 vs. 2004</b>	% Yes (weighted):		1.1% vs. 3.1%	t = 28.21	p<.0001	
<b>Food Stamps –</b>						
<b>Any receipt pre w1m1?</b>						
Yes – first reported in core	1,583	2.4	2.5	2,693	3.3	3.4
<u>Yes – first reported in RHTM</u>	<u>4,698</u>	<u>7.2</u>	<u>7.5</u>	<u>6,312</u>	<u>7.8</u>	<u>7.9</u>
<b>Yes – TOTAL</b>	6,281	9.6	10.0	9,005 <sup>3/</sup>	11.1	11.3
No <sup>1/</sup>	56,502	86.1	90.0	70,397	86.9	88.7
screened out of RHTM <sup>3/</sup>	1,142	1.7		–	–	
(missing) <sup>4/</sup>	(1,705)	(2.6)		(1,613)	(2.0)	
<b>T-TEST RESULTS, 2001 vs. 2004</b>	% Yes (weighted):		9.1% vs. 10.0%	t = 5.81	p<.0001	

[Table 1 continued on next page.....]

TABLE 1 NOTES:

<sup>1/</sup> For the sake of analytical simplicity, the “no” category includes both the explicit “no’s” and those few cases in which D or R nonresponse to one or more questions obscured a person’s pre-wave 1 receipt status. D/R nonresponse for the items used to establish receipt status was rare in both the 2001 and 2004 panels, and is unlikely to have any discernable impact on the results and conclusions.

<sup>2/</sup> In 2001, as noted in the text, respondents who were not a current “guardian” of a child in the household were skipped out of the RHTM’s AFDC/TANF question series. (The current analysis treats these cases as the equivalent of a “no” response.) This flaw was corrected in 2004, and thus does not affect the 2004 results.

<sup>3/</sup> As noted in the text, the 2001 RHTM screened some people age 18+ out of the module entirely – namely, those who (a) did not report Food Stamps receipt in the core interview, and (b) did report receipt of SSI (either federal or state), and either (c) reported AFDC/TANF receipt or (d) were not a guardian. (The current analysis treats these cases as the equivalent of missing data.) This flaw was corrected in 2004, and thus does not affect the 2004 results.

<sup>4/</sup> “Missing” cases are probably due to partial interviews – respondents who quit the interview before this point in the questionnaire. The rate of “missing-ness” in the AFDC/TANF results appears lower in 2001 only because the vast majority of cases were non-guardians, and thus were skipped out of this question series entirely. Missing data rates are comparable for the SSI and Food Stamps results.

<sup>5/</sup> The “yes” category excludes 17 respondents who originally reported pre-wave 1 AFDC/TANF receipt, but only as a child beneficiary (not “on his/her own,” or “in his/her own name”); these cases are included with the “no’s.” As noted in the text, this extra step to ensure correct understanding of the intent of the questions – receipt in one’s own name – was not used in 2001.

<sup>6/</sup> As noted in the text, due to an error in the 2001 questionnaire, no core data were collected on the start date of SSI spells which were in progress in month 1 of wave 1.

<sup>7/</sup> The “yes” category excludes 26 respondents who originally reported pre-wave 1 SSI receipt, but only as a child beneficiary (not “in his/her own name,” or as a parent of a child beneficiary); these cases are included with the “no’s.” As noted in the text, this extra step to ensure correct understanding of the intent of the questions – receipt in one’s own name – was not used in 2001.

<sup>8/</sup> The “yes” category excludes 70 respondents who originally reported pre-wave 1 Food Stamps receipt, but only as a child beneficiary (not “on his/her own,” or “in his/her own name”); these cases are included with the “no’s.” As noted in the text, this extra step to ensure correct understanding of the intent of the questions – receipt in one’s own name – was not used in 2001.

**TABLE 2: The Effects of the 2001 RHTM Screening Procedures on Differences in Reported Pre-Wave-1 Receipt of AFDC/TANF, SSI, and Food Stamps, Unweighted and Weighted, in the 2001 and 2004 SIPP Panels**

	2001			2004				
	n	% (unweighted)		WTD %	n	% (unweighted)		WTD %
TOTAL N – interviewed adults 18+	65,630				81,015			
<b>AFDC/TANF –</b>								
<b>Any receipt pre w1m1? [Table 1]</b>								
Yes	1,709	2.6	2.7	<b>2.3</b>	3,604	4.4	4.5	<b>4.0</b>
No	12,240	18.7	19.1	18.1	75,804	93.6	95.5	96.0
skipped AFDC/TANF Qs	50,274	76.6	78.3	79.6	–	–		
screened out of RHTM	1,135	1.7			–	–		
(missing)	(272)	(0.4)			(1,607)	(2.0)		
ADJUSTED DATA FOR 2004:								
[identifies cases which would have been screened out, had the 2001 procedures been retained]				Yes	3,495	4.3	4.5	<b>3.9</b>
				No	74,781	92.3	95.5	96.1
				“screened out”	1,145	1.4		
				(missing)	(1,594)	(2.0)		
<u>T-TEST RESULTS</u>								
2004 (“screened out”*) vs. 2004 (adjusted)		% Yes (weighted):	8.9% vs. 3.9%	t = 5.72				p<.0001
2001 (observed) vs. 2004 (adjusted)		% Yes (weighted):	2.3% vs. 3.9%	t = 17.21				p<.0001
2004 (observed) vs. 2004 (adjusted)		% Yes (weighted):	4.0% vs. 3.9%	t = 0.69				n.s.
* - data not shown								
<b>SSI (Fed or state) –</b>								
<b>Any receipt pre w1m1? [Table 1]</b>								
Yes	732	1.1	1.2	<b>1.1</b>	2,789	3.4	3.5	<b>3.1</b>
No	62,048	94.5	98.8	98.9	76,603	94.6	96.5	96.9
skipped AFDC/TANF Qs	1,142	1.7			–	–		
screened out of RHTM	1,142	1.7			–	–		
(missing)	(1,708)	(2.6)			(1,623)	(2.0)		
ADJUSTED DATA FOR 2004:								
[identifies cases which would have been screened out, had the 2001 procedures been retained]				Yes	1,943	2.4	2.5	<b>2.2</b>
				No	76,308	94.2	97.5	97.8
				“screened out”	1,145	1.4		
				(missing)	(1,619)	(2.0)		
<u>T-TEST RESULTS</u>								
2004 (“screened out”*) vs. 2004 (adjusted)		% Yes (weighted):	73.8% vs. 2.2%	t = 53.73				p<.0001
2001 (observed) vs. 2004 (adjusted)		% Yes (weighted):	1.1% vs. 2.2%	t = 16.81				p<.0001
2004 (observed) vs. 2004 (adjusted)		% Yes (weighted):	3.1% vs. 2.2%	t = 12.14				p<.0001
* - data not shown								
[Table 2 continued on next page.....]								

	2001				2004			
	n	% (unweighted)		WTD %	n	% (unweighted)		WTD %
TOTAL N – interviewed adults 18+	65,630				81,015			
<b>Food Stamps –</b>								
<b>Any receipt pre w1m1? [Table 1]</b>								
Yes	6,281	9.6	10.0	<b>9.1</b>	9,005	11.1	11.3	<b>10.0</b>
No	56,502	86.1	90.0	90.9	70,397	86.9	88.7	90.0
screened out of RHTM	1,142	1.7			–	–		
(missing)	(1,705)	(2.6)			(1,613)	(2.0)		
ADJUSTED DATA FOR 2004:				Yes	8,638	10.7	11.0	<b>9.7</b>
[identifies cases which would have been				No	69,632	85.9	89.0	90.3
screened out, had the 2001 procedures				“screened out”	1,145	1.4		
been retained]				(missing)	(1,600)	(2.0)		
<b>T-TEST RESULTS</b>								
2004 (“screened out”*) vs. 2004 (adjusted)	% Yes (weighted):	29.9% vs. 9.7%	t = 14.43	p < .0001				
2001 (observed) vs. 2004 (adjusted)	% Yes (weighted):	9.1% vs. 9.7%	t = 4.06	p < .01				
2004 (observed) vs. 2004 (adjusted)	% Yes (weighted):	10.0% vs. 9.7%	t = 1.83	n.s.				
* - data not shown								

**TABLE 3: The Effects of the 2001 RHTM’s Restricted (“Current Guardians Only”) Eligibility for the AFDC/TANF Question Series on Differences in Reported Pre-Wave-1 Receipt of AFDC/TANF, Unweighted and Weighted, in the 2001 and 2004 SIPP Panels**

	2001			2004				
	n	% (unweighted)		WTD %	n	% (unweighted)		WTD %
TOTAL N – interviewed adults 18+	65,630				81,015			
<b>AFDC/TANF –</b>								
<b>Any receipt pre w1m1? [Table 1]</b>								
Yes	1,709	2.6	2.7	<b>2.3</b>	3,604	4.4	4.5	<b>4.0</b>
No	12,240	18.7	19.1	18.1	75,804	93.6	95.5	96.0
skipped AFDC/TANF Qs	50,274	76.6	78.3	79.6	–	–	–	–
screened out of RHTM	1,135	1.7			–	–		
(missing)	(272)	(0.4)			(1,607)	(2.0)		
PROJECTED DATA FOR 2004:			Yes		1,969	2.4	2.4	<b>2.1</b>
[identifies “non-guardian” cases which would have			No		14,462	17.9	17.9	17.2
been skipped out of the AFDC/TANF series, had the			“skipped”		64,376	79.5	79.7	80.7
2001 procedures been retained]			(missing)		(208)	(0.3)		

**T-TEST RESULTS**

2004 (“skipped”*) vs. 2004 (adjusted)	% Yes (weighted):	2.3% vs 2.1%	t = 3.17	p<.05
2001 (observed) vs. 2004 (adjusted)	% Yes (weighted):	2.3% vs 2.1%	t = 3.36	p<.05
2004 (observed) vs. 2004 (adjusted)	% Yes (weighted):	4.0% vs 2.1%	t = 22.21	p<.0001

\* - data not shown

**TABLE 4: The Effects of the Re-Ordering of the 2004 RHTM's Questions (Food Stamps Last) on Differences in Reported Pre-Wave-1 Receipt of Food Stamps, Unweighted and Weighted, in the 2001 and 2004 SIPP Panels**

	2001			2004				
	n	% (unweighted)		WTD %	n	% (unweighted)		WTD %
TOTAL N – interviewed adults 18+	65,630				81,015			
<b>Food Stamps –</b>								
<b>Any receipt pre w1m1? [Table 1]</b>								
Yes	6,281	9.6	10.0	<b>9.1</b>	9,005	11.1	11.3	<b>10.0</b>
No	56,502	86.1	90.0	90.9	70,397	86.9	88.7	90.0
screened out of RHTM	1,142	1.7			–	–		
(missing)	(1,705)	(2.6)			(1,613)	(2.0)		
ADJUSTED DATA FOR 2004:								
[recodes 140 “yes” cases – those captured			Yes		8,865	10.9	11.2	<b>9.8</b>
by the new RHTM probes – to “no”]			No		70,537	87.1	88.8	90.2
			(missing)		(1,613)	2.0		
<b>T-TEST RESULTS</b>								
2001 (observed) vs. 2004 (adjusted)		% Yes (weighted):	9.1% vs 9.8%	t = 4.85	p < .001			
2004 (observed) vs. 2004 (adjusted)		% Yes (weighted):	10.0% vs 9.8%	t = 1.01	n.s.			

**TABLE 5: Summary of Number of Spells Reported (Including Spells Within the Wave 1 Reference Period), Among Respondents with Reported Receipt in Either the Core or the RHTM, by Program Type, in the 2001 and 2004 SIPP Panels (unweighted)<sup>1/</sup>**

	AFDC/TANF		Food Stamps	
	2001	2004	2001	2004
Total (non-blank) #-of-Spells Reports <sup>2/</sup>	1,700 <sup>3/</sup>	3,624	6,632 <sup>4/</sup>	8,547
number of #-of-spells reports filled automatically, based on responses to other items (%)	59 (3.5%)	2,768 (76.4%)	1,273 (19.2%)	6,649 (77.8%)
	t=87.46, p<.0001		t=88.75, p<.0001	
maximum number of spells reported	99	20	99	30
number reporting 1 spell <sup>4/</sup> (%)	1,008 (59.3%)	3,100 (85.5%)	4,499 (67.8%)	6,982 (81.7%)
	t=19.78, p<.0001		t=19.51, p<.0001	
number reporting 1-3 spells (%)	1,353 (79.6%)	3,444 (95.0%)	5,541 (83.5%)	8,118 (95.0%)
	t=14.82, p<.0001		t=22.29, p<.0001	
number reporting 10+ spells (%)	79 (4.6%)	17 (0.5%)	276 (4.2%)	26 (0.3%)
	t=7.99, p<.0001		t=15.29, p<.0001	
number of D/R nonresponses (%)	155 (9.1%)	82 (2.3%)	450 (6.8%)	159 (1.9%)
	t=9.26 p<.0001		t=14.41, p<.0001	
number of “suspicious” <sup>5/</sup> #-of-spells reports (%)	54 (3.2%)	12 (0.3%)	194 (2.9%)	14 (0.2%)
	t=6.53, p<.0001		t=13.06, p<.0001	
mean # of spells reported (excluding reports of 99 spells (2001 only))	2.9	1.3	2.6	1.3
	t=8.36, p<.0001		t=14.10, p<.0001	
cases with “first rec’d” or “last rec’d” date nonresponse (in 2004 these cases skipped the # of spells question)	196	326	781	841
ADJUSTED* 2001 DATA mean # of spells reported (excluding reports of 99 spells (2001 only))	2.9	1.3	2.6	1.3
* - sets 2001 # of spells to missing for missing date cases	t=8.13, p<.0001		t=13.59, p<.0001	

[Table 5 continued on next page.....]

TABLE 5 NOTES:

<sup>1/</sup> Weighted estimates are very similar to those reported here, and the results of analyses of weighted data are virtually identical [data not shown].

<sup>2/</sup> For both 2001 and 2004, respondents who reported receipt only within the wave 1 reference period (no pre-wave-1 receipt) are considered to have experienced 1 spell, even though they skipped the number-of-spells question in the RHTM. This is one source of the discrepancy between the number cases with a non-missing #-of-spells report and the frequency of pre-wave-1 spell reports in Table 1. Other sources of discrepancy include the presence of extra screening criteria for the module as a whole, or the AFDC/TANF question series (2001 only), and nonresponse to key RHTM questions, resulting in an exit from the RHTM questions and thus a blank #-of-spells variable.

<sup>3/</sup> Excludes 2 reports of zero spells.

<sup>4/</sup> Excludes 18 reports of zero spells.

<sup>5/</sup> “Suspicious” reports are defined as those divisible by 10 (i.e., rounded) or 12, which are likely to indicate reporting of monthly payments for some number of years.



**TABLE 6: Start Date Repairs in the 2004 SIPP Panel, by Type of Repair, and the Extent of “Suspiciously Young” Start Date Reporting for Left-Censored and First-Ever Spells of Receipt of AFDC/TANF, SSI, and Food Stamps in the 2001 and 2004 SIPP Panels (unweighted)<sup>1/</sup>**

	2001		2004	
	n	%	n	%
<b>AFDC/TANF –</b>				
pre-w-1 spells first reported in core – Total [Table 1]	80		101	
<u>Total with non-missing spell start year</u>	<u>79</u>	<u>100%</u>	<u>101</u>	<u>100%</u>
start year repaired following “suspiciously young” probe	x		1	1.0
start year repaired following “non-continuous” probe	x		4	4.0
<b>“suspiciously young” final spell start year (age ≤ 16)</b>	<b>2</b>	<b>2.5</b>	<b>0</b>	<b>0.0</b>
T-TEST RESULTS: n/a				
pre-w-1 spells first reported in RHTM – Total [Table 1]	1,629		3,503	
<u>Total with non-missing spell start year</u>	<u>1,522</u>	<u>100%</u>	<u>3,303</u>	<u>100%</u>
start year repaired following “suspiciously young” probe	x		84	2.5
<b>“suspiciously young” final spell start year (age ≤ 16)</b>	<b>46</b>	<b>3.0</b>	<b>80</b>	<b>2.4</b>
T-TEST RESULTS: 3.0% vs. 2.4% t = 1.17 n.s.				
<b>SSI (Fed or state) –</b>				
pre-w-1 spells first reported in core – Total [Table 1]	0		1,626	
<u>Total with non-missing spell start year</u>			<u>1,614</u>	<u>100%</u>
start year repaired following “suspiciously young” probe			94	5.8
start year repaired following “non-continuous” probe			24	1.5
<b>“suspiciously young” final spell start year (age ≤ 16)</b>			<b>49</b>	<b>3.0</b>
T-TEST RESULTS: n/a				
pre-w-1 spells first reported in RHTM – Total [Table 1]	732		1,163	
<u>Total with non-missing spell start year</u>	<u>661</u>	<u>100%</u>	<u>1,094</u>	<u>100%</u>
start year repaired following “suspiciously young” probe	x		45	4.1
<b>“suspiciously young” final spell start year (age ≤ 16)</b>	<b>46</b>	<b>7.0</b>	<b>25</b>	<b>2.3</b>
T-TEST RESULTS: 7.0% vs. 2.3% t = 4.30 p<.01				
<b>Food Stamps</b>				
pre-w-1 spells first reported in core – Total [Table 1]	1,583		2,693	
<u>Total with non-missing spell start year</u>	<u>1,578</u>	<u>100%</u>	<u>2,676</u>	<u>100%</u>
start year repaired following “suspiciously young” probe	x		67	2.5
start year repaired following “non-continuous” probe	x		149	5.5
<b>“suspiciously young” final spell start year (age ≤ 16)</b>	<b>22</b>	<b>1.4</b>	<b>23</b>	<b>0.9</b>
T-TEST RESULTS: 1.4% vs. 0.9% t = 1.55 n.s.				
pre-w-1 spells first reported in RHTM – Total [Table 1]	4,698		6,312	
<u>Total with non-missing spell start year</u>	<u>4,236</u>	<u>100%</u>	<u>5,830</u>	<u>100%</u>
start year repaired following “suspiciously young” probe	x		101	1.7
<b>“suspiciously young” final spell start year (age ≤ 16)</b>	<b>70</b>	<b>1.7</b>	<b>83</b>	<b>1.4</b>
T-TEST RESULTS: 1.7% vs. 1.4% t = 0.92 n.s.				