

# An Initial Evaluation of the 2010 Field Test of Re-Engineered SIPP

SIPP-EHC Data Evaluation Workgroup\*

U.S. Census Bureau

SEHSD Working Paper Number 2011-9

March 2011

## Abstract

We analyze data from the 2010 field test of a re-engineered instrument for the Survey of Income and Program Participation (SIPP). This instrument, called SIPP-EHC, uses an event history calendar and extends the retrospective interview reference period from four months in SIPP to one year in SIPP-EHC. We evaluate this new interviewing method by comparing rates of reported employment and program participation in data from the 2010 field test with those from 2008 SIPP data for the same period. Most of these rates are not statistically different across the two surveys, with differences of only a few percentage points where differences exist. We also test for time trends in the differences between monthly reported participation rates across the surveys

---

\*The workgroup members are Rebecca Chenevert, Ashley Edwards, Renee Ellis, Jason Fields, Graton Gathright, David Hedengren, Carolyn Hronis, Jeongsoo Kim, Lindsay Monte, Daniel Perez-Lopez, Michelle Sandhoff, Marina Vornovytsky, and Rachael Walsh. This paper is released to inform interested parties of ongoing research and to encourage discussion of work in progress. All views expressed in this paper are those of the authors and not necessarily those of the U.S. Census Bureau. Any errors or omissions are the sole responsibility of the authors. E-mail: [graton.m.gathright@census.gov](mailto:graton.m.gathright@census.gov) or [marina.s.vornovytsky@census.gov](mailto:marina.s.vornovytsky@census.gov)

and find no evidence of memory decay in the SIPP-EHC in reported employment or reported participation in WIC, Food Stamps, or Medicare.

## 1 Introduction

The Survey of Income and Program Participation (SIPP) is a major longitudinal household survey conducted by the U.S. Census Bureau. SIPP data is currently used by researchers and government agencies to study annual and sub-annual income and program participation dynamics in the United States. In an effort to reduce costs and improve data quality, the Census Bureau has undertaken a re-engineering of SIPP. A re-engineered SIPP instrument (SIPP-EHC) is scheduled to replace the current SIPP instrument in 2014. A key component of SIPP-EHC is a longer retrospective reference period: one year instead of four months. To help elicit accurate response about the entire reference period, the new instrument incorporates an event history calendar (EHC) as part of the interviewing strategy.

This paper presents an analysis of data from the 2010 field test of the SIPP-EHC instrument. We compare rates of reported employment and program participation between the 2010 SIPP-EHC data and data from 2008 Panel of SIPP (2008 SIPP). Most of the rates across the two surveys are not statistically different, with difference of only a few percentage points where differences exist.<sup>1</sup>

We also address the concern of some stake-holders (National Research Council, 2009) that SIPP-EHC would be less accurate about events earlier in the one year retrospective reference period than about events in a four months reference period. We test for time trends in the differences between monthly reported employment and program participation rates across the 2010 SIPP-EHC and 2008 SIPP and find no significant evidence of memory decay in 2010 SIPP-EHC for reported employment

---

<sup>1</sup>The proportions and other statistics in the text, figures, and tables of this report describe the samples considered here for the purpose of evaluating differences between the two surveys. The weighting and design effects necessary to interpret the reported statistics as estimates of underlying population parameters have not been incorporated into these results. Apparent differences may not be statistically significant, but all comparative statements in this report have undergone statistical testing and are significant at the 95% confidence level.

or participation in WIC, Food Stamps, or Medicare.

Administrative records provide a complementary approach to the evaluation of 2010 SIPP-EHC data that we present in this report. Administrative records for CY2009 Medicare enrollment and IRS-Form-W2-reported earnings recently been received by the Census Bureau, and evaluation of 2010 SIPP-EHC using these records is in process. It is anticipated that CY2009 administrative records for Medicaid, Food Stamps (selected states) and TANF (selected states) will eventually become available as well. We discuss these projects in more detail in the Conclusion (Section 6).

This paper is structured as follows: Section 2 provides background on SIPP and SIPP-EHC. Section 3 describes our approach to comparing data from 2010 SIPP-EHC and 2008 SIPP. Section 4 describes our analytical framework and empirical approach. Section 5 discusses the results, and Section 6 offers concluding remarks.

## 2 Background

### 2.1 The Survey of Income and Program Participation (SIPP)

SIPP is a nationally-representative, interviewer-administered, longitudinal survey conducted by the U.S. Census Bureau. In its current design, each SIPP panel consists of multiple waves of interviewing, with waves administered three times a year at four month intervals. The SIPP core instrument, which contains the survey content that is repeated in each wave, is detailed, long, and complex. It collects information about household structure, labor force participation, income sources and amounts, educational attainment, and health insurance over the four month retrospective reference period. Various “topical modules,” appended to the core with each wave, collect detailed data on taxes, assets, liabilities, receipt of program benefits as well as marital, fertility, and employment histories.<sup>2</sup>

---

<sup>2</sup>Additional information on the SIPP can be found at the following websites: [www.census.gov/sipp/](http://www.census.gov/sipp/) (main SIPP website), [www.census.gov/sipp/workpapr/wp230.pdf](http://www.census.gov/sipp/workpapr/wp230.pdf), (SIPP Quality Profile) and [www.census.gov/sipp/usrguide.html](http://www.census.gov/sipp/usrguide.html) (SIPP User’s Guide).

## 2.2 The Re-engineering of SIPP

In early 2006, the Census Bureau was faced with a projected shortfall of approximately 40 million dollars in its fiscal year 2007 budget. The agency proposed to offset the shortfall by eliminating the SIPP program. Policy-makers, data users, and other stake-holders protested strongly to Congress, emphasizing the unique value of the survey. Ultimately, Congress opted to continue the survey, providing funding for both its continuation and a broad array of improvements that became the current re-engineering effort.

A key component of the re-engineered SIPP (or SIPP-EHC) is a proposed shift away from a three-times-per-year data collection schedule to a single annual interview with a one year retrospective reference period. Such a change should yield several benefits including reduced field costs and respondent burden. To maintain data quality while accomplishing this change, event history calendar (EHC) methods have been incorporated into the SIPP-EHC instrument.

The development and evaluation of the SIPP-EHC has included a paper-and-pencil reinterview feasibility study conducted in 2008 (Moore et al, 2008) and a field test of an initial computer-assisted interview (CAPI) instrument in 2010. This report presents an analysis of the data from the 2010 field test. A 2011 field test is currently in the field. A field test in 2012 will test the wave 2 version of the instrument. The SIPP-EHC instrument is scheduled to become the production instrument beginning in 2014.

## 3 Data

The primary goal of this paper is to assess the accuracy of reports collected using the new 2010 SIPP-EHC instrument. We compare rates of reported employment and program participation across 2010 SIPP-EHC and 2008 SIPP. We have endeavored to make this comparison in a manner that isolates the differences in the rates that are due to the data collection method used and not due to differences in sample design or variable construction between the two surveys. To this end, we

have constructed a dataset that we call, “Mapped SIPP” (MSIPP)<sup>3</sup>, in which we have mapped both surveys to a common standard and included only those units from 2008 SIPP that form a suitable comparison group for evaluating 2010 SIPP-EHC.

All of the results presented in this paper are from the 2010 MSIPP file. All comparisons between SIPP and SIPP-EHC presented in this paper should be interpreted as a comparison between 2010 SIPP-EHC and our comparison 2008 SIPP sub-sample.

This section is organized as follows: Subsection 3.1 and Subsection 3.2 describe how sample design and sample composition differences between 2010 SIPP-EHC and 2008 SIPP have been addressed in constructing MSIPP. Subsection 3.3 describes the steps that were taken to ensure that variables used to produce estimates for the two surveys were comparable. In other words, Subsection 3.1 and Subsection 3.2 address comparability of respondents included in the study, while Subsection 3.3 addresses comparability of responses provided by these individuals.

### 3.1 MSIPP Sample Design

The 2010 MSIPP dataset consists of all units assigned to 2010 SIPP-EHC and all units from 2008 SIPP that matched the entire list of 2010 SIPP-EHC sample restrictions (listed in Figure 1). Figure 1 also documents a subtle difference in sample design between 2010 SIPP-EHC and 2008 SIPP that was easily addressed with an appropriate re-weighting scheme.<sup>4</sup>

To understand this sample design difference, it is helpful to understand certain aspects of SIPP/SIPP-EHC sample design. Following the 2000 Census, the Census Bureau created twelve SIPP sample designations. A *sample designation* is a list of addresses (or *designated units*) that constitute a suitable sample for a SIPP panel. The manner in which these sample designation are created ensures

---

<sup>3</sup>Release of a public-use version of 2010 MSIPP is intended. The internal version of 2010 MSIPP will be made available to external researchers through the Census Bureau’s network of secure Research Data Centers. The MSIPP framework will also be used to compare estimates produced by future SIPP-EHC field tests to SIPP and to administrative records.

<sup>4</sup>This re-weighting should not be confused with the official SIPP sampling weights.

that units are assigned to the sample designations at random and that the sample designations are created without regard to the panel for which they may ultimately be employed.

For the 2008 SIPP Panel, sample designation S05 was included in its entirety. Since there was a desire to have a sample that was state-representative for the 20 largest states, state-specific numbers of reduction groups from S11 were also included.<sup>5</sup> For 2010 SIPP-EHC, two entire sample designations (S16 and S17) were employed.<sup>6</sup>

Since the proportion of S11 that was included in the 2008 SIPP sample varied by state, 2008 SIPP effectively under-samples certain states relative to 2010 SIPP-EHC. To account for this state under-sampling, we employ a state-specific adjustment factor so that those interviewed in each state represent the whole *sample* from that state. The adjustment factor associated with 2010 SIPP-EHC is equal to 1.00, since there was no state-specific under-sampling. The adjustment factor associated with 2008 SIPP is also equal to 1.00 for some states (such as Maryland and Wisconsin) that were not under-sampled in 2008 SIPP relative to 2010 SIPP-EHC. Figure 1 lists adjustment factors associated with each state.

### 3.2 MSIPP Sample Composition

We include in 2010 MSIPP all sample households and persons associated with the included sample units. We view this as a study design decision since there are other defensible approaches for comparing the two surveys. The need to take a decision here arises because the initial interview periods for the two surveys were over a year apart. Sample units for 2010 SIPP-EHC were rostered in early 2010 whereas units for 2008 SIPP were rostered in late 2008. This temporal offset creates differences in the sample composition between 2008 SIPP and 2010 SIPP-EHC.

As an illustration of the type of differences that arise from this temporal offset in rostering periods,

---

<sup>5</sup>Each SIPP sample designation is partitioned into 101 reduction groups designed to facilitate sample expansion or reduction without compromising the required properties of the sample.

<sup>6</sup>Of course, only sample units satisfying the entire list of criteria shown in Figure 1 were included in the 2010 SIPP-EHC sample.

consider a unit that was assigned to 2008 SIPP, but which satisfied all of the 2010 SIPP-EHC sample restrictions. The persons rostered at this unit in late 2008 could relocate at some point in CY2009 to a residence outside of the 2010 SIPP-EHC geographic coverage area. At this location they would not have been sampled had they been assigned to 2010 SIPP-EHC and rostered in early 2010.

A radically different alternative to the approach that we have taken would have been to attempt to construct from 2008 SIPP a comparison subsample of those households (not units) whose addresses in early 2010 satisfied the 2010 SIPP-EHC sample restrictions. For both practical and sample design reasons this cannot be done perfectly, and doing so would not exploit as cleanly the random assignment of units to the two surveys.

### **3.3 MSIPP Topics and Variable Construction**

The discussion so far has focused on which sample units, households, and persons were included in 2010 MSIPP. In this subsection, we describe decisions taken in mapping comparable concepts from both surveys to common MSIPP variables.

MSIPP was created using “unedited” SIPP and SIPP-EHC data. At the time of this study, only unedited data was available for 2008 SIPP wave 5 and the 2010 SIPP-EHC data of interest. Using unedited data also has the advantage of allowing us to eliminate differences in surveys arising from post-collection data processing.

The variables included in 2010 MSIPP dataset include administrative variables (identifiers for sample unit, household, person and respondent; codes for wave, rotation, reference month, and interview month; our state-specific adjustment factor variable and the sampling weights), demographic variables (age, sex, race, and ethnicity), and monthly participation indicators for several topics (WIC, Food Stamps, TANF, Medicaid, Medicare, and employment).

The monthly indicator for employment signals participation in any job or business activity for any number of days in the given month.

For Food Stamps, TANF, WIC, and Medicaid, there can be multiple reports for a given sample persons since a sample person’s participation can be attributed to other members of the household to which they belong.<sup>7</sup> We aggregate these reports by treating a person as having participated in the program if anyone reports participation for him or her. We treat “refused” or “don’t know” reports as reports of non-participation.

For all topics we focused on groups that were asked for a report in both surveys. For employment we restrict attention to sample persons aged 16 or older. We measure age at time of SIPP-EHC or SIPP wave 5 interview. For Medicare, we focus on sample persons aged 65 or older. For WIC, we focus on children age 5 and under and females aged 16 to 45.

In mapping the data, we discovered two important differences between the surveys in the attribution of reported spells of program participation to others in the household. In SIPP-EHC, a sample person’s report about his or her own participation is necessarily coupled with any report about participation by others in the household. In SIPP, by contrast, a person can report participation for another but not for herself. In robustness checks discussed below, we make use of an alternate participation indicator which we forced to indicate participation for anyone who reports participation for someone else.

A second difference is that in SIPP, only a parent or guardian of a child can report WIC participation for that child. In SIPP-EHC, any adult sample person in the household can report participation for any child in the household. The impact of this difference is something that future work could explore.

## 4 Methods

The process by which units (addresses) were sampled for SIPP-EHC and SIPP produces a convenient experimental framework for comparing the two surveys. Effectively, a single sample of addresses was created, and the sampled addresses were randomly assigned to one of two treatments, SIPP-EHC or

---

<sup>7</sup>For Medicaid multiple reports can arise only in SIPP.



SIPP. The assigned survey was administered to the household living at each sampled address, and the reports were observed.

Let  $SIPPEHC_i$  equal 1 when unit  $i$  is assigned to SIPP-EHC and  $SIPPEHC_i = 0$  otherwise. Let  $R_i^e$  be the report that household at unit  $i$  would give if assigned to SIPP-EHC and let  $R_i^s$  be the report if assigned to SIPP. We will refer to these as *potential reports*.

We would like to know the value of  $E[R_i^e - R_i^s]$ , but we observe only the report for the method that was administered. If assignment to survey is independent of the potential reports, then we can compare the methods by comparing the expected values of reports across the two methods. That is,

$$E[R_i^e | SIPPEHC_i = 1] - E[R_i^s | SIPPEHC_i = 0] = E[R_i^e - R_i^s] \quad (1)$$

The same argument can be applied to any conditional means statistic used to describe the response distributions for each method. In this paper, we investigate patterns of monthly reported status using linear regression.

The random assignment of sample addresses to treatment ensures no correlation between  $SIPPEHC_i$  and unobserved potential reports for sample persons. The difference in conditional expectations of observed reports for the two surveys will reflect the net effect of all of the design differences between the two surveys.

Our basic empirical specifications is to estimate the following model using person-month data:

$$R_{ij} = \alpha + \beta SIPPEHC_i + \varepsilon_{ij} \quad (2)$$

where  $R_{ij}$  is the report on employment or program participation (topic) obtained from respondent  $i$

concerning month  $j$ .  $SIPPEHC_i$  takes the value of 1 if the report from respondent  $i$  was obtained using 2010 SIPP-EHC and the value of 0 if the SIPP method was used. The model for each topic has been estimated using weighted least squares (weighting by the state-specific adjustment factor) with standard errors clustered at the person level.

To investigate the hypothesis of memory decay in SIPP-EHC, we also estimate the following specification for each topic:

$$R_{ij} = \alpha + \beta_1 SIPPEHC_i + \beta_2 MONTHS_{ij} + \beta_3 SIPPEHC_i * MONTHS_{ij} + \varepsilon_{ij} \quad (3)$$

where  $MONTHS_{ij}$  is the number of months since December 2008. As above, the model for each topic has been estimated using weighted least squares (weighting by the state-specific adjustment factor) with standard errors clustered at the person level.

## 5 Results

This section discusses the major findings of the paper. In particular, Subsection 5.1 presents a comparison of wave 1 unit response rates for 2008 SIPP (hereafter, SIPP) and 2010 SIPP-EHC (hereafter, SIPP-EHC), as well as a comparison of the characteristics of non-interviewed households from the two surveys. In Subsection 5.2, we compare the characteristics of respondents. In Subsection 5.3, we compare rates of reported employment and program participation in the two surveys. In Subsection 5.4, we disaggregate these rates by several categories. Finally, Subsection 5.5 presents trends in reported participation over the course of the reference period to address the hypothesis of memory decay in SIPP-EHC.

These results from 2010 MSIPP have been constructed for the sole purpose of analyzing how the statistics associated with two different data collection methods differ after accounting for sample

design and other design differences. The statistics reported in this section characterize the *MSIPP sample*, and are *not* designed to describe either the national population of households or any particular geography.

## 5.1 Response Rates and Characteristics of Non-Interview Households

In each survey, there were sample units for which no interview was obtained. Figure 2 presents wave 1 unit response rates and characteristics of non-interviewed households for 2010 MSIPP by survey.<sup>8</sup>

The first column shows the response rates for both surveys. The rate is higher for SIPP-EHC. One possible explanation for the higher response rate in 2010 SIPP-EHC relative to a comparison sub-sample of 2008 SIPP is the fact that 2010 SIPP-EHC was administered right before the 2010 Decennial Census when the Census Bureau had launched an intensive advertising campaign.

The next four columns of Figure 2 compare the proportion of non-interview households across surveys by gender of the reference person, tenure, and household size. Overall, the percent of non-responding households with a male reference person is greater in SIPP, although this result seems to be driven by the Dallas region alone. Additionally, there is no significant difference between the two surveys with respect to the tenure status of non-respondents. Average household size reported in SIPP-EHC is greater than SIPP, although the result appears to be driven by the Boston and Dallas regions.

We have repeated our analysis of response rates and characteristics of non-interview households (for all regions) both with sampling weights and without any weights (see Figure 3). This analysis showed similar patterns.

---

<sup>8</sup>Characteristics of non-interview households are based on the best guess of the field representative attempting to make the interview.

## 5.2 Characteristics of Sample Persons

Figure 4 presents a comparison of the characteristics of respondents included in 2010 MSIPP by survey. There is no significant difference between SIPP and SIPP-EHC in the proportion of household members who are adults. The percentage reporting race as White in SIPP is greater than in SIPP-EHC overall and for New York, Philadelphia and Dallas regions. For all regions, there is no significant difference between the percentage reporting race as Black in SIPP and SIPP-EHC. The Chicago and Dallas regions report a greater percentage reporting race as Black in SIPP while the Philadelphia region reports a greater percentage reporting race as Blacks in SIPP-EHC. Except for Philadelphia and the Dallas regions, there is no significant difference in the percentage reporting Hispanic origin in SIPP and SIPP-EHC.

There also appear to be more Type-Z<sup>9</sup> individuals in SIPP-EHC than in 2008 SIPP sub-sample, and this is true for every region. There are two factors that we conjecture could have contributed to this difference. The first is that interviews using this prototype SIPP-EHC instrument proved long.<sup>10</sup> The second is that field representatives in 2010 SIPP-EHC may not have tried as hard to obtain an interview, given that they may have understood that the survey was a field test instead of a production survey.<sup>11</sup>

There are more proxy interviews for SIPP than for SIPP-EHC which also seems to be true across all regions. This could be due to differences in the way that the respondents or field representatives reacted when an individuals in the household was not available for an interview. It seems plausible that many interviews that would have been proxy interviews in a SIPP interview became Type-Z interviews in SIPP-EHC.

We repeated our analysis of the characteristics of sample persons (for all regions) both with sampling weights and without any weights (see Figure 5). This analysis showed similar patterns with the

---

<sup>9</sup>A rostered person in an interviewed household for whom no (self or proxy) interview is provided is classified as “Type-Z”

<sup>10</sup>This is something that was specifically addressed in refinements incorporated into the 2011 SIPP-EHC instrument.

<sup>11</sup>This is something that was addressed in the development of training materials for the 2011 SIPP-EHC survey.

exception of the proportions of sample persons reporting race as Black and reporting Hispanic origin. In the unweighted analysis, a higher proportion of sample persons reported race as Black in SIPP than in SIPP-EHC. Also, a lower proportion of sample persons reported Hispanic origin in SIPP than in SIPP-EHC. When analyzed using the sampling weights, both of these relationships were reversed.

### **5.3 Probabilities of Reported Employment and Program Participation**

Figure 6 reports probabilities of reported employment and program participation for SIPP and SIPP-EHC. In all regions together, 2010 SIPP-EHC has higher levels of participation in each of WIC, Food Stamps, and Medicaid. However, the reported participation levels in SIPP are higher for TANF and Medicare. There is no statistically significant difference in reported participation for employment between the two surveys.

We repeated our analysis of the the probability of reported employment and program participation (for all regions) both with sampling weights and without any weights (see Figure 7). This analysis showed similar patterns with the exception of Medicaid and TANF. In the unweighted analysis, the rates of reported Medicaid participation were not statistically different across the two surveys. When using the sampling weights, it is TANF that no longer shows a statistically significant difference in rates of reported participation.

We also repeated this analysis using the alternative participation indicator discussed in Section 3 (see fourth row of Figure 7). This reversed the relationship between the rates of reported participation across the surveys for Food Stamps and WIC. The differences in this alternative specification were 1, 2, and 5 percentage points for TANF, Food Stamps and WIC, respectively.

## 5.4 Monthly Employment and Program Participation Probabilities by Categories

In Figure 8, we present a comparison of rates of reported employment and program participation by several categories. For women and children, the probability of participating in WIC is higher in SIPP than SIPP-EHC. Food stamps participation probabilities from SIPP and SIPP-EHC are not statistically different for Type Z persons, for interviews by proxy, or for women, children, White, and Black. Probability of reported Medicare participation was smaller in SIPP-EHC for not Type Z, self, proxy, male, adult, and Hispanic. Probability of reported Medicaid participation was smaller in SIPP-EHC for children, Whites and Non-Hispanics. Differences in probability of reported employment were not statistically significant for any sub-group.

## 5.5 Trends in Month-by-Month Rates of Reported Employment and Program Participation

In this subsection, we investigate whether memory decay is a problem in SIPP-EHC by looking at month-by-month probabilities of reported employment and program participation between the two surveys. A particular concern in moving from a 4-month recall period in SIPP to a longer 12-month recall period in SIPP-EHC is that respondents may have difficulty remembering events that happened earlier in the one year reference period, as opposed to within four months of the interview which would be similar to the recall burden required in SIPP. The term used for this possibility is “memory decay”. If memory decay is an issue for 2010 SIPP-EHC, then we would expect to observe differences in rates of reported employment between SIPP and SIPP-EHC early in the CY2009 with these differences declining over the course of reference period.

Figure 9 plots month-specific rates of reported employment and participation for all regions and by regional office. If memory-decay has an important impact on the reporting of employment and program participation in SIPP-EHC, we would expect to see differences between monthly rates across the two surveys early in the reference period that decrease over the course of the reference period.

This data visualization suggests that memory decay is not a problem in 2010 SIPP-EHC for most of the topics studied here since there either appears to be no difference across surveys (reported employment) or there appears to be differences that increase over the course of the reference period (Food Stamps, Medicare, WIC). Decreasing differences appear to be a possibility for Medicaid and TANF.

Table 1 presents evidence which corroborates the conclusions of this data visualization. We present estimates of linear trends in the differences between rates across the months of the reference period. In this specification,  $SIPPEHC$  is an indicator for assignment to the SIPP-EHC survey, and  $t$  is the number of months since December 2008. We would take it as evidence of memory-decay if the coefficients on  $sippehc$  and on the interaction term  $sippehcXt$  indicated that the difference between the monthly rates were declining over the course of the reference period. For WIC, we find no statistically significant change over time in the difference between the rates for the two surveys. For Food Stamps, the difference grows slightly over the course of the reference period which is also not consistent with the hypothesis of memory-decay. Similarly, we find no evidence of memory decay for reported employment and reported Medicare participation. For TANF, the rate of reported participation is lower in SIPP-EHC at the beginning of the reference period, and this difference decreases over time. This pattern is consistent with the hypothesis of memory-decay. The difference between the surveys in reported Medicaid participation also decreases over time in this specification.

Since attrition in 2008 SIPP over the reference period could produce time trends in observed differences that are unrelated to memory decay in SIPP-EHC, we repeat our tests for trends with a sample limited to those respondents who were in the surveys for the entire calendar year (see Table 2). This analysis showed similar patterns except for Medicaid which showed no evidence of memory decay in this alternative specification.

## 6 Conclusion

We have made an initial evaluation of data from the 2010 field test of 2010 SIPP-EHC by comparing rates of reported employment and program participation 2010 SIPP-EHC and 2008 SIPP. The surveys appear to generate similar patterns in responses for comparable groups of respondents. Most of the rates across the two surveys are not statistically different, with differences of only a few percentage points where differences exist.

We have given particular attention to the possibility that SIPP-EHC respondents would have difficulty recalling events early in the one year reference period. We find no evidence of memory decay in reported employment or reported participation in Medicare, WIC, or Food Stamps. We find evidence of memory decay only for reported participation in TANF and (in one of two specifications) Medicaid.

At the time of this writing, a 2011 SIPP-EHC field test is underway. It will be informative to compare the data from 2011 SIPP-EHC to 2008 SIPP as has been done here for 2010 SIPP-EHC. It will also be helpful to see how the comparison differs across the 2010 and 2011 field tests. Further analysis of 2010 SIPP-EHC data as well as in 2011 SIPP-EHC data could address the impact of reporting for other household members, responses of “don’t know” and “refused”, and could consider alternative compositions for comparison samples.

As mentioned in the introduction, the results of the 2010 SIPP-EHC field test are also being evaluated using administrative records. Using IRS Form W2 data for persons in the sample, we are assessing the accuracy of reported job counts and reported earnings. These same administrative records also facilitate a link to the Census Bureau’s Business Register allowing a validation that we are conducting of the employer characteristics reported in the surveys. Monthly data on Medicare enrollment are being used to assess the accuracy of reported monthly Medicare participation, including the accuracy of reported transitions into Medicare participation. Other topics for which administrative records analysis is scheduled for this year are SNAP, SSI, and Medicaid.



## References

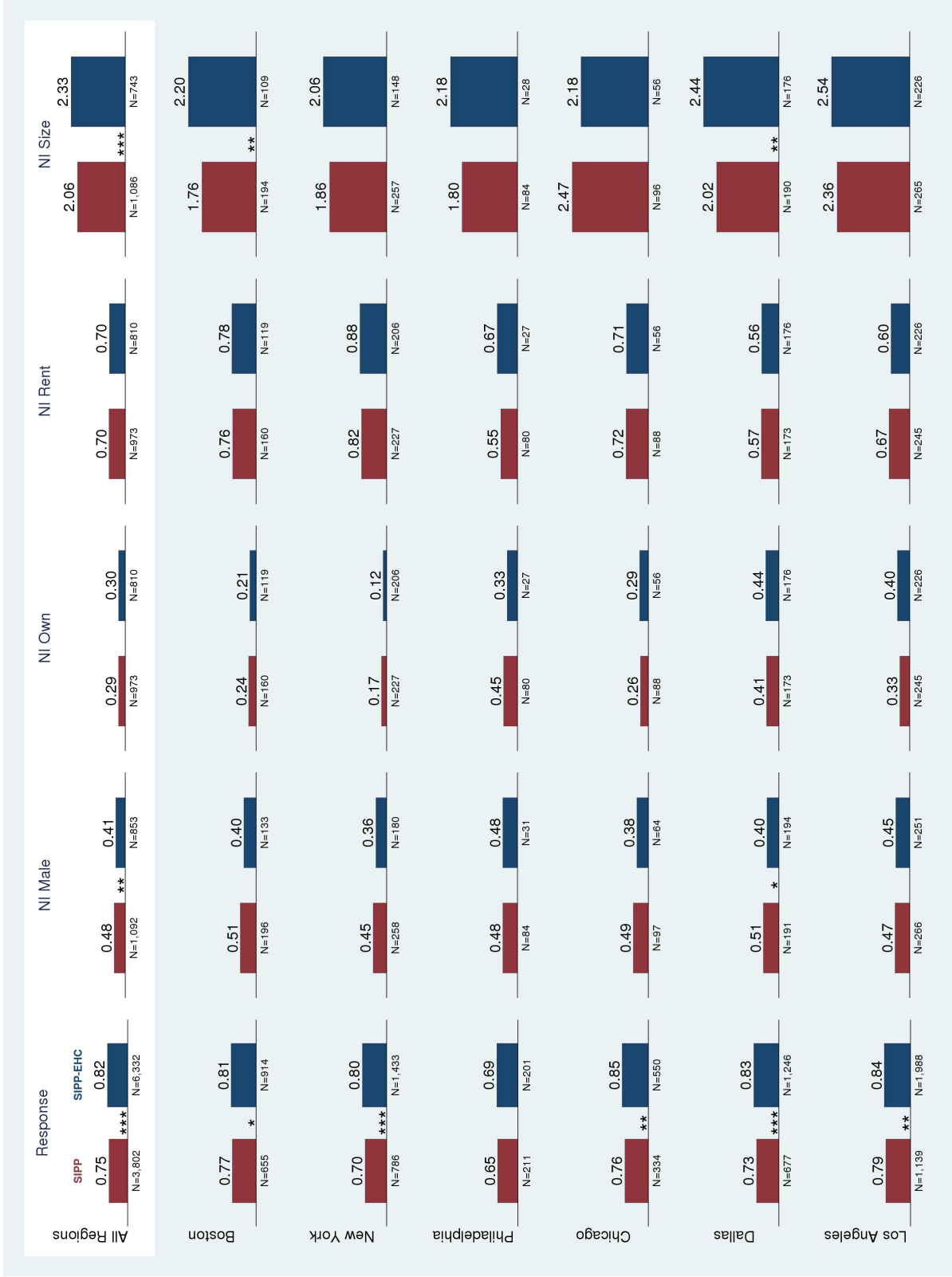
- [1] National Research Council. *Reengineering the Survey of Income and Program Participation*. Panel on the Census Bureaus Reengineered Survey of Income and Program Participation, Constance F. Citro and John Karl Scholz, Editors. Committee on National Statistics, Division of Behavioral and Social Sciences and Education, Washington, DC, 2009.
  
- [2] Jeffrey Moore, Jason Fields, Gary Benedetto, Martha Stinson, Anna Chan, and Jerry Maples. The 2008 survey of income and program participation event history calendar field test: Study design and initial results. Technical report, U.S. Census Bureau, 2009.

Sample Designation	Designated Units		Sample Restrictions	Wave 1 Household Interviews Attempted	Wave 1 Household Interviews	MSIPP Households	MSIPP Restrictions	Wave 1 MSIPP Households	MSIPP State-Specific Adjustment Factor*
	Assigned To	2008 SIPP							
S05	53,600	2008 SIPP	None	53,600	32,160	2,254	Same as Sample Restrictions on S16	2,254	CA, NY, CT, IL, LA, RI, TX MA MD, WI 1.8532 1.0466 1.0000
S11	45,000	2008 SIPP	State-specific number of deleted reduction groups:	11,800	7,080	2,487	Same as Sample Restrictions on S16	2,487	
			CA, NY, CT, IL, LA, RI, TX	93					CA, NY, CT, IL, LA, RI, TX
			MA	9					MA
			MD, WI	0					MD, WI
			Other states	various					
S16	45,000	2010 SIPP-EHC	Unit frame	3,991	3,219	3,219	None	3,219	All states
			Self-representing PSU						
			High poverty stratum						
			Located in CA, CT, IL, LA, MA, MD, NY, RI, TX, or WI						
			Served by Boston, Chicago, Dallas, Los Angeles, Philadelphia, or New York Regional Office						
S17	45,000	2010 SIPP-EHC	Same as S16 above	3,991	3,186	3,186	None	3,186	All states

\* MSIPP state-specific adjustment factor = 202/(202-# of deleted reduction groups).

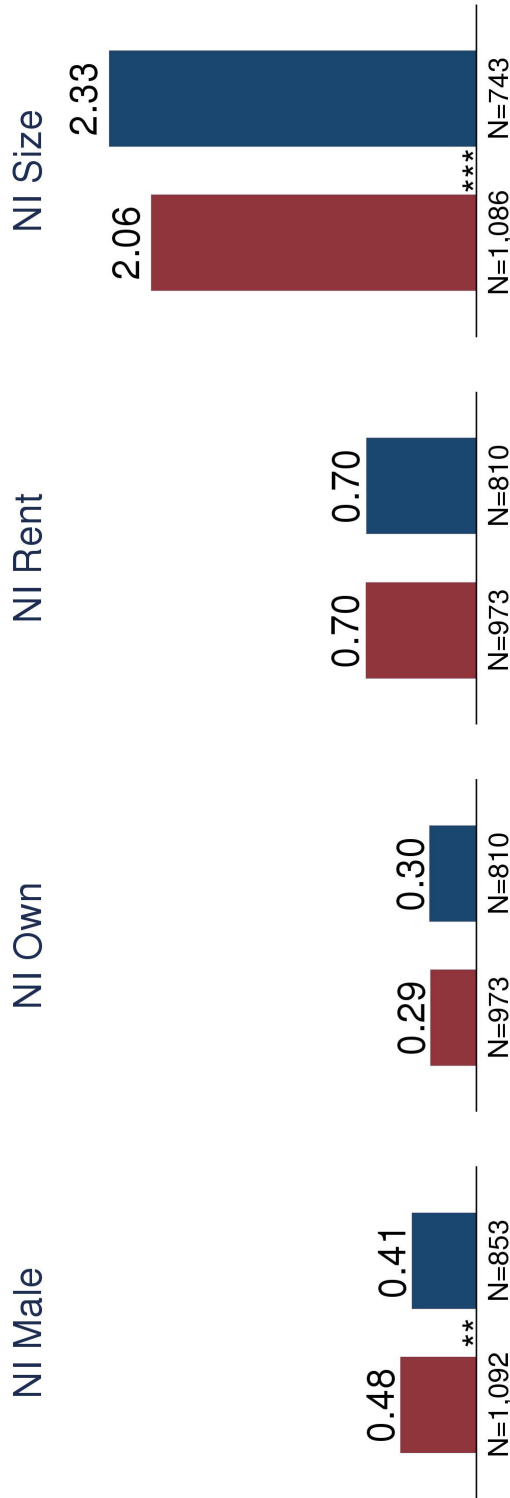
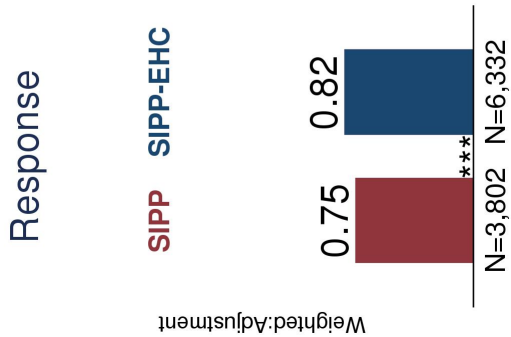
Figure 1: 2010 MSIPP Sample Design.

This table describes aspects of the sample design of 2008 SIPP and 2010 SIPP-EHC that are relevant to the construction of 2010 MSIPP. In particular, the design-phase sample restrictions of the surveys are listed. In addition, restrictions on which 2008 SIPP sample units are included in 2010 MSIPP are listed in the “MSIPP Restrictions” column. The table also describes an adjustment factor for state-under-sampling in 2008 SIPP relative to 2010 SIPP-EHC in the “MSIPP State-Specific Adjustment Factor”.



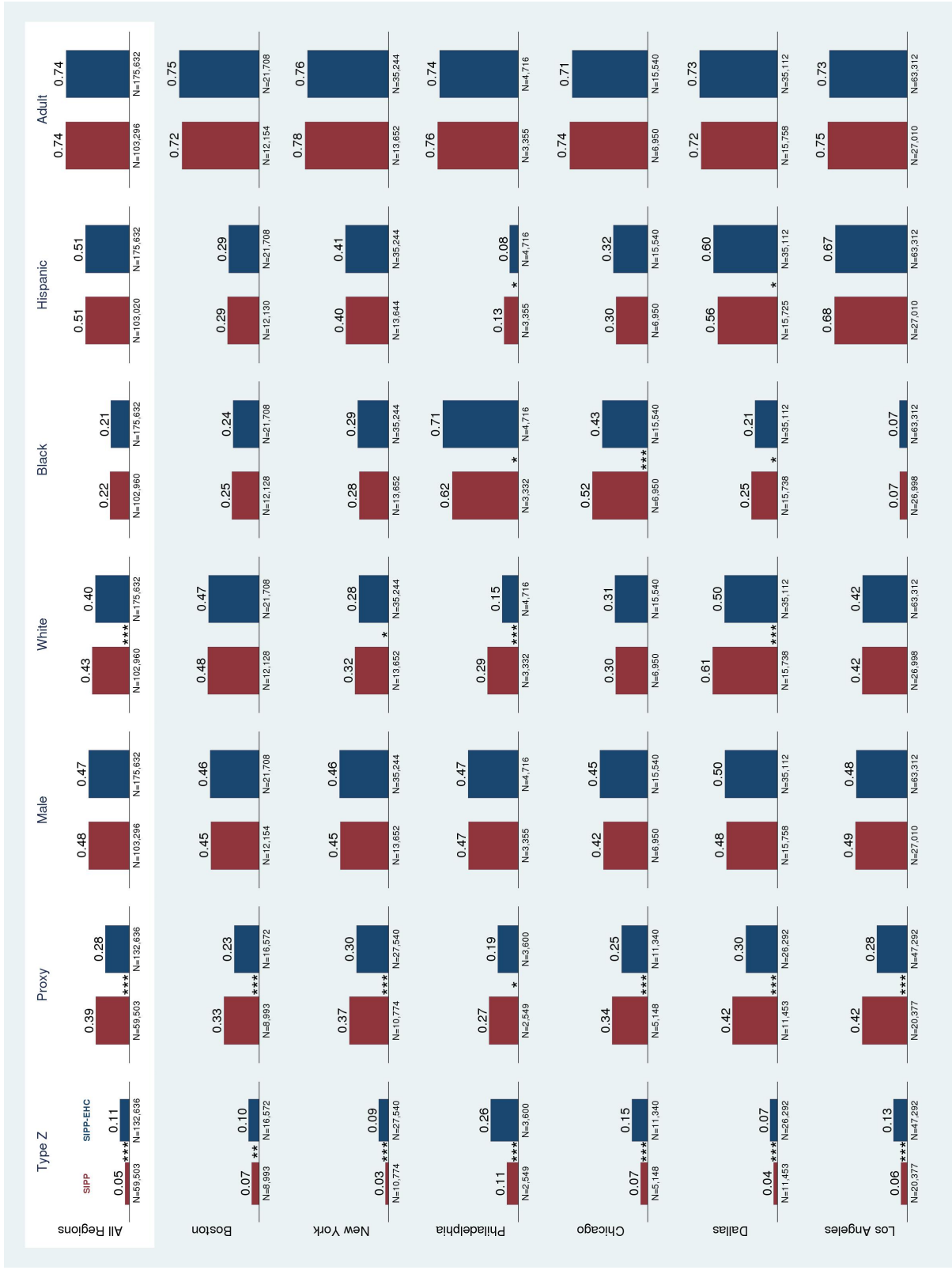
**Figure 2: Wave 1 Unit Response Rates for SIPP and SIPP-EHC and Characteristics of Non-Interviewed Households**

This figure reports comparisons of wave 1 unit response rates in 2010 MSIPP by survey type: 2008 SIPP (in red) or 2010 SIPP-EHC (in blue) and characteristics of non-responding households. Response rates are reported in the first column for all regions and by regional office. Response rates were calculated as the proportion of sampling units eligible for an interview from whom an interview was obtained. Non-responding sampling units are considered to have been eligible for interview if no one was at home, occupants were temporarily absent, refused to give any information, or could not be located. The number of eligible sampling units in MSIPP from each survey is indicated beneath the corresponding bar. Columns 2 to 5 report characteristics of non-responding households as reported by the field representative who attempted the interview. NI Male is the proportion of non-interviewed sampling units where the reference person was male. NI Own and NI Rent are proportion of non-responding households who owned and rented their residence, respectively. NI Size is the number of persons who belong to the household. Statistical significance of the difference between reported sample proportions is indicated with one, two, or three asterisks between the bars, representing statistical significance at the 5%, 1%, and 0.1% levels, respectively. Proportions for each survey were calculated using linear regression of an indicator for the given characteristic on a constant and an indicator for the survey to which the sample unit was assigned. The sample units were weighted in the regression by the MSIPP state adjustment factor. These statistics characterize the MSIPP sample and are not designed to describe the national population of households.



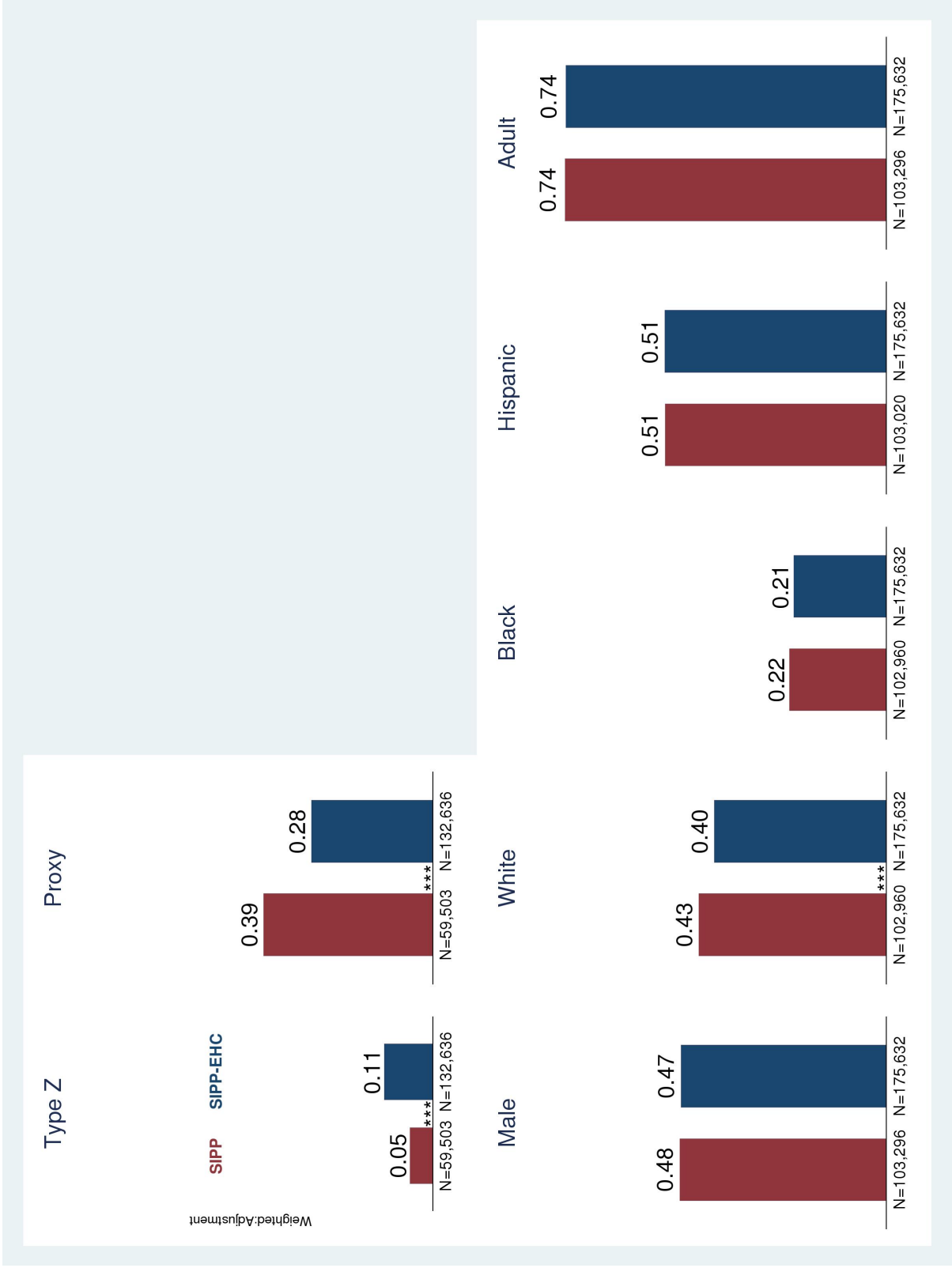
**Figure 3: Wave 1 Unit Response Rates for SIPP and SIPP-EHC and Characteristics of Non-Interviewed Households: Alternative Specifications**

This figure reports comparisons of wave 1 unit response rates in 2010 MSIPP by survey type: 2008 SIPP (in red) or 2010 SIPP-EHC (in blue) and characteristics of non-responding households. Response rates are reported in the first column. Response rates were calculated as the proportion of sampling units eligible for an interview from whom an interview was obtained. Non-responding sampling units are considered to have been eligible for interview if no one was at home, occupants were temporarily absent, refused to give any information, or could not be located. The number of eligible sampling units in MSIPP from each survey is indicated beneath the corresponding bar. Columns 2 to 5 report characteristics of non-responding households as reported by the field representative who attempted the interview. NI Male is the proportion of non-interviewed sampling units where the reference person was male. NI Own and NI Rent are proportion of non-responding households who owned and rented their residence, respectively. NI Size is the number of persons who belong to the household. Each row presents results using the indicated weighting. Statistical significance of the difference between reported sample proportions is indicated with one, two, or three asterisks between the bars, representing statistical significance at the 5%, 1%, and 0.1% levels, respectively. Proportions for each survey were calculated using linear regression on a constant and an indicator for the survey to which the sample unit was assigned. These statistics characterize the MSIPP sample and are not designed to describe the national population of households.



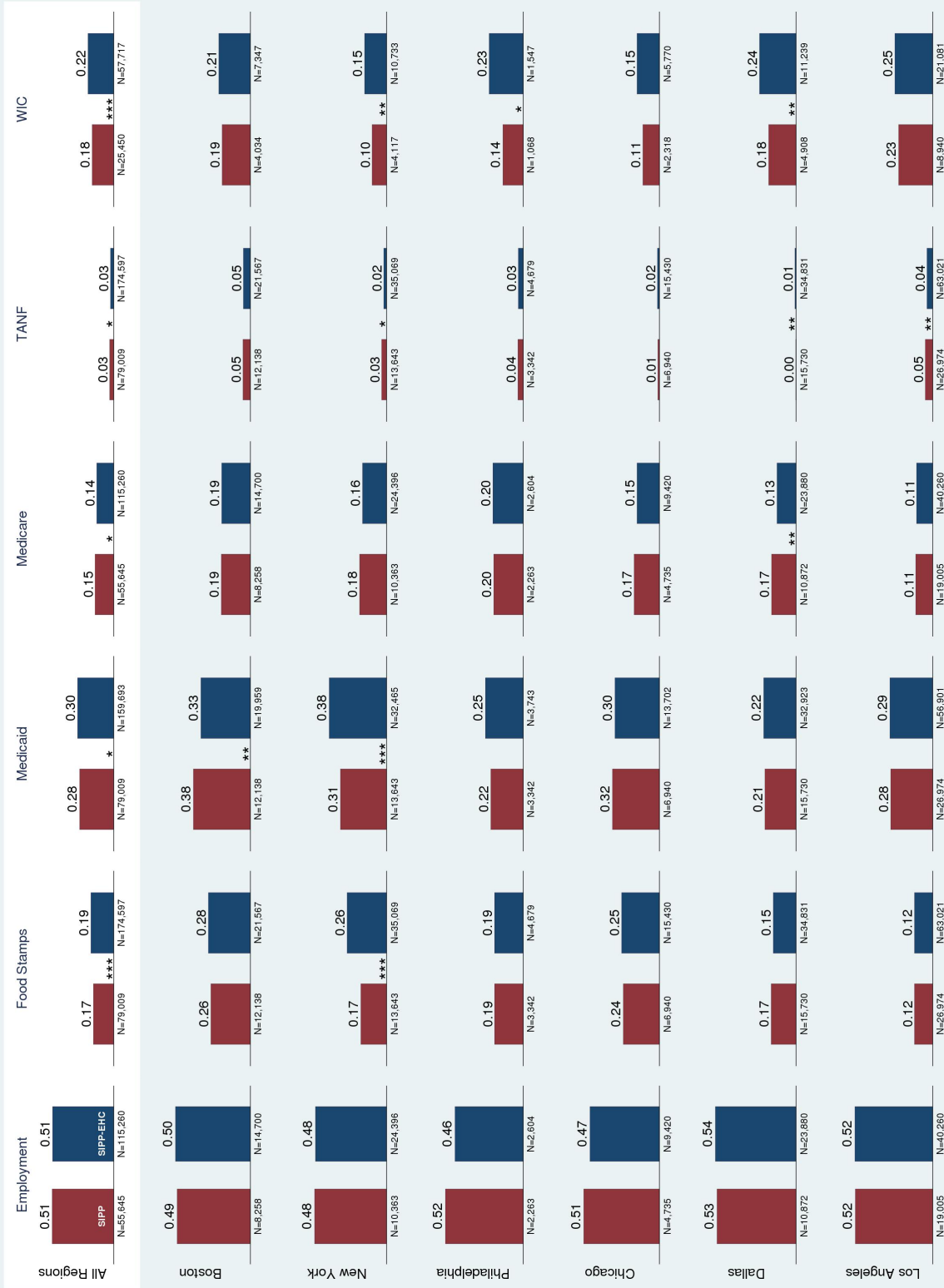
**Figure 4: Characteristics of Sample Persons.**

This figure reports characteristics of respondents in the sample for whom an interview or a proxy interview was obtained. These characteristics are reported across the two surveys and across regions. An adjustment has been made to account for state under-sampling in 2008 SIPP relative to 2010 SIPP-EHC. All statistics reported here characterize the sample and are not nationally representative of households. The asterisks represent statistically significant differences at the 5%, 1%, and 0.1% levels for one, two, and three stars, respectively. “Type Z” stands for person-level non-response within a responding household. “Proxy” stands for whether a proxy interview was granted.



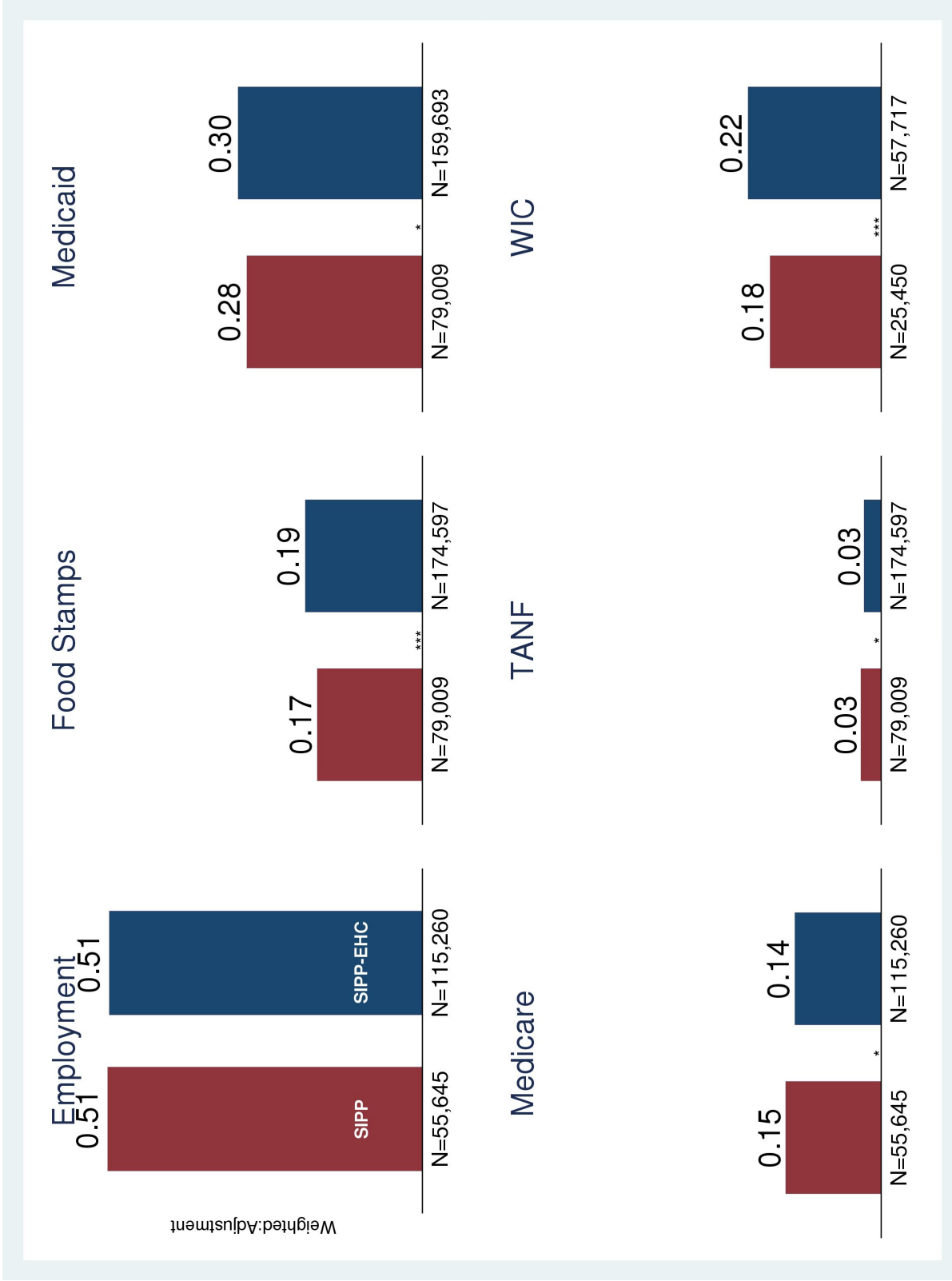
**Figure 5: Characteristics of Sample Persons: Alternative Specifications.**

This figure reports characteristics of respondents in the sample for whom an interview or a proxy interview was obtained. All statistics reported here characterize the sample and are not nationally representative of households. Each row presents results using the indicated weighting. The asterisks represent statistically significant differences at the 5%, 1%, and 0.1% levels for one, two, and three stars, respectively. “Type Z” stands for person-level non-response within a responding household. “Proxy” stands for whether a proxy interview was granted.



**Figure 6: Probability of Reported Employment and Program Participation in any Given Month.**

This figure reports probability of reported employment or program participation in any given month. The sample size is the number of person-month observations where respondents were in-universe to provide the corresponding report. The sample sizes vary across topics because of varying eligibility restrictions. For example, we restrict attention to sample persons above a cutoff age for possible employment participation. An adjustment has been made to account for state under-sampling in 2008 SIPP relative to 2010 SIPP-EHC. All statistics reported here characterize the sample and are not nationally representative of households. The asterisks represent statistically significant differences at the 5%, 1%, and 0.1% levels for one, two, and three stars, respectively. In this context, "Employed" means that an individual worked for pay in a job or business at some point during the month.



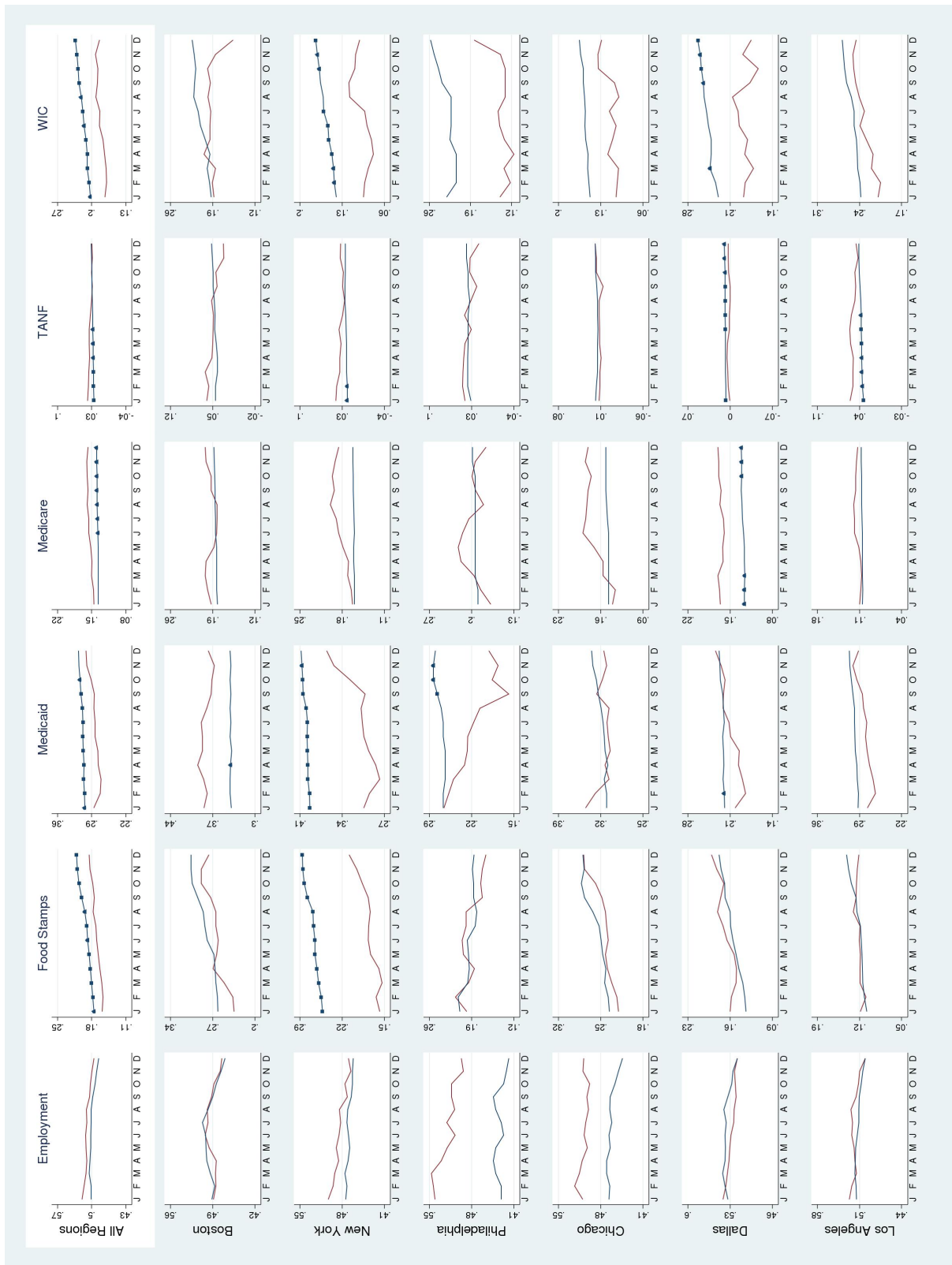
**Figure 7: Probability of Reported Employment and Program Participation in any Given Month: Alternative Specifications.**

This figure reports probability of reported employment or program participation in any given month. The sample size is the number of person-month observations where respondents were in-universe to provide the corresponding report. The sample sizes vary across topics because of varying eligibility restrictions. For example, we restrict attention to sample persons above a cutoff age for possible employment participation. The first three rows present results using the indicated weighting. The fourth row presents results using an alternative participation indicator discussed in the text. All statistics reported here characterize the sample and are not nationally representative of households. The asterisks represent statistically significant differences at the 5%, 1%, and 0.1% levels for one, two, and three stars, respectively. In this context, "Employed" means that an individual worked for pay in a job or business at some point during the month.

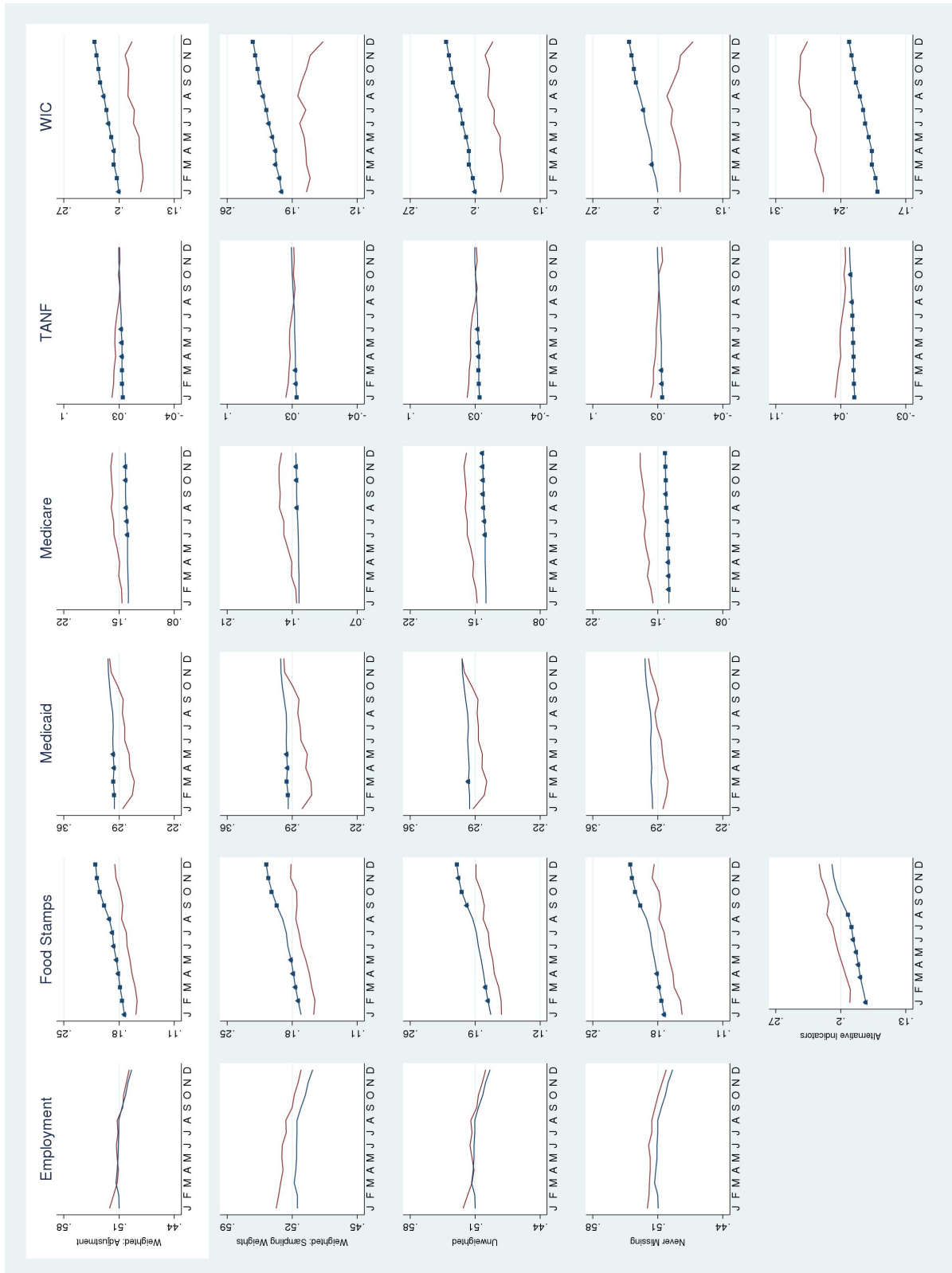




**Figure 8: Probability of Reported Employment and Program Participation by Several Categories.** This figure reports probability of reported employment and program participation in any given month, by different sample person characteristics. The sample size is the number of person-month observations where respondents were "in universe" and the characteristic being considered (i.e., gender) was known for them. The sample sizes vary across topics because of different restrictions imposed on who could be asked a particular question. For example, those under 15 were not asked about employment, and person-months associated with these individuals would not be considered "in universe" for calculating employment rates. An adjustment has been made to account for state under-sampling in 2008 SIPP relative to 2010 SIPP-EHC. All statistics reported here characterize the sample and are not nationally representative of households. The asterisks represent statistically significant differences at the 5%, 1%, and 0.1% levels for one, two, and three stars, respectively. "Type Z" stands for person-level non-response within a responding household. "Proxy" stands for whether a proxy interview was granted. "Employed" means that an individual worked for pay in a job or business at some point during the month. The fraction of those participating was calculated as the number of person-months reporting participation over the number of in-universe person-months.



**Figure 9: Month-by-month Rates of Reported Employment and Program Participation.** Each graph presents the time paths of rates of employment or a program participation for SIPP and SIPP-EHC in red and blue, respectively. In all of the graphs presented in this figure the scale is constant, although the vertical distance relative to the origin varies (that is, the portion of the graph displayed is shifted up or down based on participation levels). An adjustment has been made to account for state under-sampling in 2008 SIPP relative to 2010 SIPP-EHC. All statistics reported here characterize the sample and are not nationally representative of households. Throughout, the squares, triangles, the squares, triangles and circles denote statistically significant differences between the two surveys at the 0.1%, 1% and 5% levels, respectively. "Employed" means that an individual worked for pay in a job or business at some point during the month.



**Figure 10: Month-by-month Rates of Reported Employment and Program Participation: Alternative Specifications.**

Each graph presents the time paths of rates of employment or a program participation for SIPP and SIPP-EHC in red and blue, respectively. In all of the graphs presented in this figure the scale is constant, although the vertical distance relative to the origin varies (that is, the portion of the graph displayed is shifted up or down based on participation levels). The first three rows present results using the indicated weighting. The fourth row presents results for only sample persons for whom a report is available for all months of 2009. The fifth row presents results using an alternative participation indicator discussed in the text. All statistics reported here characterize the sample and are not nationally representative of households. Throughout, the squares, triangles and circles denote statistically significant differences between the two surveys at the 0.1%, 1% and 5% levels, respectively. "Employed" means that an individual worked for pay in a job or business at some point during the month.

Table 1: **Testing for time-trends in month-specific differences**

VARIABLES	(1) WIC	(2) Food Stamps	(3) TANF	(4) Medicaid	(5) Employment	(6) Medicare	(7) Employment
sippehc	0.0290*** (0.0106)	0.0144** (0.00564)	-0.0145*** (0.00278)	0.0239*** (0.00695)	-0.00499 (0.00921)	-0.00846 (0.00637)	-0.0117 (0.00973)
sippehcXt	0.000914 (0.000893)	0.000807* (0.000455)	0.00139*** (0.000223)	-0.00145** (0.000570)	0.000397 (0.000685)	-0.000931** (0.000409)	0.00327 (0.00208)
sippehcXt2							-0.000220 (0.000152)
t	0.00196** (0.000837)	0.00264*** (0.000417)	-0.000910*** (0.000211)	0.00220*** (0.000551)	-0.00176*** (0.000615)	0.00133*** (0.000403)	-0.000696 (0.00189)
t2							-8.25e-05 (0.000139)
Constant	0.168*** (0.00870)	0.153*** (0.00460)	0.0390*** (0.00244)	0.270*** (0.00566)	0.521*** (0.00756)	0.146*** (0.00529)	0.519*** (0.00806)
Observations	83,167	253,606	253,606	238,702	170,905	170,905	170,905
R-squared	0.002	0.001	0.000	0.000	0.000	0.000	0.000

This table presents the results for all regions together of ordinary least squares regressions of indicators for program participation and employment on an indicator for SIPP-EHC instrument (sippehc, vs. SIPP instrument) and its interaction with a linear (or quadratic) time trend, and a linear (or quadratic) time trend. The regressors  $t$  is the number of months since December 2008;  $t^2$  is its square. We take it as evidence of memory-decay if the coefficients on sippehc and on the interaction term sippehcXt indicate that the difference between the monthly rates is declining over the course of the reference period.

Table 2: **Testing for time-trends in month-specific differences: Alternative Indicators**

VARIABLES	(1) WIC	(2) TANF	(3) Food Stamps
sippehc	-0.0601*** (0.0117)	-0.0213*** (0.00295)	-0.0179*** (0.00594)
sippehcXt	0.000444 (0.000965)	0.00149*** (0.000237)	0.000218 (0.000470)
t	0.00243*** (0.000914)	-0.00101*** (0.000225)	0.00323*** (0.000433)
Constant	0.257*** (0.0101)	0.0458*** (0.00263)	0.186*** (0.00497)
Observations	83,167	253,606	253,606
R-squared	0.005	0.001	0.001

This table presents the results for all regions together of ordinary least squares regressions of alternative indicators (see Section 3) for program participation and employment on an indicator for SIPP-EHC instrument (sippehc, vs. SIPP instrument) and its interaction with a linear (or quadratic) time trend, and a linear (or quadratic) time trend. The regressors  $t$  is the number of months since December 2008;  $t^2$  is its square. We take it as evidence of memory-decay if the coefficients on sippehc and on the interaction term sippehcXt indicate that the difference between the monthly rates is declining over the course of the reference period.