Modeling a Bridge When Survey Questions Change: Evidence from the Current Population Survey Health Insurance Redesign

Brett O'Hara[1], Carla Medalia[1], and Jerry J. Maples[1]

Most research on health insurance in the United States uses the Current Population Survey Annual Social and Economic Supplement. However, a recent redesign of the health insurance questions disrupted the historical time trend in 2013. Using data from the American Community Survey, which has a parallel trend in the uninsured rate, we model a bridge estimate of the uninsured rate using the traditional questions. Also, we estimate the effect of changing the questionnaire. We show that the impact of redesigning the survey varies substantially by subgroup. This approach can be used to produce bridge estimates when other questionnaires are redesigned.

Key words: Health insurance; redesigned survey; aggregate model; unit model.

1. Introduction

Health insurance is the primary avenue to receive health care in the United States. People without health coverage are less likely to go to the doctor or hospital than people who have health coverage (O'Hara and Caswell 2013). As such, people with health insurance have higher economic well-being than people who lack health insurance because they have greater access to health care services (Kaestner and Lubotsky 2016). Health insurance coverage also acts as a buffer against the adverse effect of health shocks (McGeary 2009; Bradley et al. 2012), but means that some workers have less job-mobility because of a dependence on their employment-sponsored insurance (Bailey and Chorniy 2016). Furthermore, health insurance is linked to public policy practices and changes in the United States. A key example is the 2010 Patient Protection and Affordable Care Act, the health care law that shaped availability and access to health insurance for millions of Americans.

The Current Population Survey Annual Social and Economic Supplement (CPS) generates widely used estimates on health insurance coverage in the United States, is used to calculate official poverty estimates, and serves as the basis for many policy-related decisions (Blewett and Davern 2006). However, estimates of the uninsured population from the CPS have been historically higher than estimates from other federal surveys (Smith and Medalia 2015). This runs counter to expectation, since the CPS measures health insurance

coverage in the previous calendar year, while the majority of other surveys measure coverage at the time of interview. When someone is more likely to be uninsured on a particular day than on all days in the year, the calendar year estimate of uninsured persons in the CPS should be *lower* than the uninsured estimates from other surveys, not higher. Research indicated several reasons why this was the case. For example, estimates from the CPS may have actually reflected a mixture of current and past year coverage (Kenney and Lynch 2010), or respondents may have had difficulty with the long recall period (Pascale et al. 2009). Respondents may also be confused about the type of coverage they have. For example, Medicaid coverage has been shown to be misreported as another type of health insurance coverage, but also misreported as being uninsured, which could contribute to an overestimate in the uninsured rate (Call et al. 2008). Another possible explanation focuses on suboptimal imputations of missing data (Davern et al. 2007).

To address these issues, the U.S. Census Bureau implemented a redesign of the health insurance questions in the CPS in 2014, which measured coverage during the 2013 calendar year. One of the major changes implemented in the redesigned survey was improvements to the way data were collected about coverage during the previous calendar year, thereby reducing potential recall bias and clarifying the reference period (Pascale et al. 2016). The redesigned survey asks about health insurance coverage on the day of the interview, and then asks follow-up questions to determine monthly coverage from January 1 of the previous calendar year through the interview date. Additional changes to the questionnaire are explained in Table 1.

The redesigned questions in the CPS lowered estimates of the uninsured rate and brought health insurance estimates more in line with other federal surveys (Smith and Medalia 2015). The timing of this questionnaire change, which completely replaced the traditional questions, was particularly important because it established a strong baseline for measuring health insurance coverage in calendar year 2013, before the implementation of many provisions of the Affordable Care Act. However, at the same time, CPS estimates from prior to 2013 are not comparable to estimates for the period 2013 and beyond, and there is no direct survey-based estimate of the effect of the questionnaire change.

While there are other surveys that measure health insurance coverage that did not undergo questionnaire redesigns during this period, the CPS is unique because it produces the official poverty estimate for the United States; due to the strong association between health insurance and income, it is important to continue the time trend in health insurance coverage in the CPS. Predicting the uninsured rate in 2013 and beyond using the traditional questions makes it possible to continue the historic time series forward, which is necessary to provide an estimate of the effect of changing the questionnaire, and is a central issue when questionnaires are redesigned. This is important because it will allow researchers and policymakers to take a broader view of trends in health insurance coverage from before the ACA's implementation in 2010. In addition, a better understanding of the effect of the CPS questionnaire change provides a measure of the percentage-point difference between the traditional and redesigned estimates for micro-simulation models. Research is needed to derive a model-based bridge estimate between the redesigned and traditional health insurance questionnaires. Our goal is to produce a reliable counterfactual: if the U.S. Census Bureau had kept the traditional health insurance questions in the CPS, what would the uninsured rate have been?

Table 1.   Comparison of the traditional and redesigned health insurance questions in the CPS.

	Traditional	Redesigned
Reference period of estimates	Previous calendar year	Previous calendar year
Reference period of questions	Only asks about previous calendar year	Starts with current and then goes back to previous calendar year
Types of coverage	Laundry-list style questions	Starts with general question and then gets more specific
How questions are asked	Collected at the household level	Collected by person; also asks if others in household were covered by fsplan
New content	n/a	Participation in the health insurance marketplace
		Employer-sponsored insurance offers and take-up
		Revised medical out-of-pocket expenses

Notes: Health insurance coverage for both is captured at the time of the survey but estimates reflect coverage during the previous calendar year.

For the details on the question wording changes from the traditional to redesigned CPS, see (Pascale 2016).

2.   Data

The data for this article come from two sources, the CPS and the American Community Survey (ACS). The CPS is an annual survey of about 98,000 addresses and includes detailed questions regarding health insurance coverage, income received and place of residence. Interviews are conducted from February through April each year, either in person or by phone. We use data from the 2010 to 2013 CPS files, which collected health insurance using the traditional questions about coverage during the previous calendar year. We also use the 2014 to 2015 CPS files, which used the redesigned questions about the months of coverage.

The ACS is a survey of about 3.5 million addresses annually, which collects social, demographic, and housing information. We use the restricted access data that are available through the U.S. Census Bureau's Research Data Centers. The ACS are collected continuously from January to December each year, and interviews are either self-administered (conducted by paper or on the internet), or interviewer-administered (in person and by phone). Note that because there were no changes to interview mode over time, differences in mode between the CPS and the ACS do not affect our results. The health insurance questions in the ACS ask about coverage on the day of the interview.

Both surveys have post-stratified weights and the standard errors are computed using successive difference replication (Fay and Train 1995).

3.   Methods

We use two methods to predict what the CPS estimate would have been if the health insurance questions had not been changed: the first is based on yearly aggregates and the second is based on a difference-in-difference model on person-level data. Both methods

rely on the assumption that the time trend in the CPS is parallel to the trend in the ACS over the period 2009 to 2012 (pre-redesign) and 2013 to 2014 (post-redesign), an assumption that we test and validate (Figure 1).

The first approach uses aggregate-level data to predict what the CPS uninsured rate may have been in 2013, had the questionnaire remained the same. To do this, we use the uninsured rate in 2013 from our auxiliary data source, the ACS, plus the difference between the uninsured rates in the CPS and ACS in 2012 (see Equation 1). The difference between the estimates in the CPS and ACS is stable over time; we average the differences between the point estimates of the uninsured rate and the variances between the two surveys over 2009 to 2012 to improve the stability of the estimates. Subsection 6.2., Appendix 2 details the method for calculating the standard errors (Equation A1).

$$Pred\big(Rate_{CPS_{2013}}\big) = Rate_{ACS_{2013}} + \frac{1}{4} \sum_{k=2009}^{2012} (Rate_{CPS} - Rate_{ACS})_k \tag{1}$$

This approach has been used in other estimates of health insurance coverage, which used the CPS rate in one year together with the growth rate in other surveys and administrative records (Centers for Medicare and Medicaid Services 2014). As shown in Equation 1, the aggregate approach estimates the uninsured rate primarily, but can also be used to estimate the effect of the questionnaire change secondarily by subtracting the estimated uninsured rate of the traditional CPS in 2013 from the observed redesigned uninsured rate in the same year.

We validate this approach by predicting the uninsured rate in years that we also have observed data, such as 2012, and find that the predicted and observed lines fall on top of each other (not shown). While the aggregate method can provide a good benchmark, it is
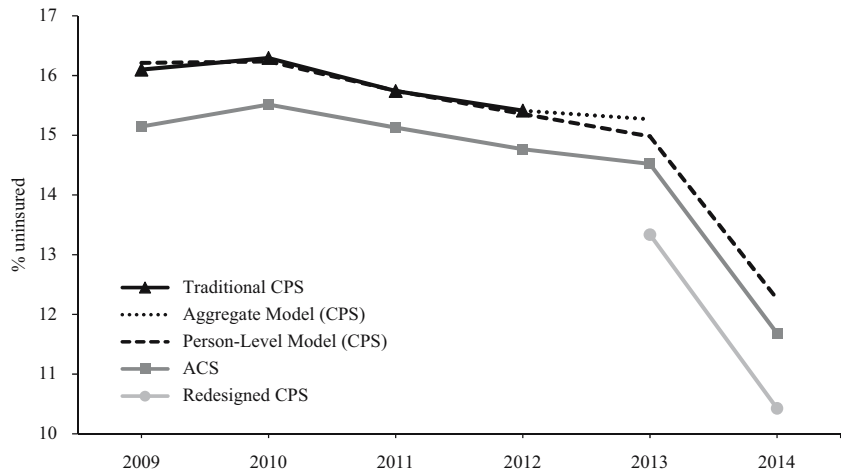
Fig. 1. Time series of the uninsured rate from 2010 to 2014, by data source. Source: 2009–2014 one-year American Community Surveys (restricted data), 2010–2015 Current Population Survey Annual Social and Economic Supplements. Note: Traditional CPS refers to the official uninsured estimate from the CPS for calendar years from 2012 and before. Redesigned CPS refers to the official uninsured estimate from the CPS for calendar years 2013 and beyond.

not practical for the analysis of subgroups due to high variance in the survey. Another limitation to the aggregate model is specific to the case explored here (e.g., health insurance coverage). As many provisions of the Affordable Care Act went into effect in 2014, the relationship between demographic characteristics and the uninsured rate changed between 2013 and 2014, so the aggregate model cannot be extended past 2013.

The second approach uses a linear probability difference-in-difference regression, hereafter referred to as the person-level model, to control for other factors that might have affected the uninsured rate in addition to the questionnaire change (Equation 2). The person-level model also enables us to examine differences in the predicted uninsured rate between subgroups, something that we could not do using the aggregate-level approach due to sample size. This person-level model assumes that certain effects are constant over time, but we also test this assumption using interactions between key demographic characteristics and a time component. In Subsection 6.2., Appendix 2, the variance formulas and the weighting procedure are discussed for the person-model.

$$P(UNINS_i) = \alpha + \gamma_t + \beta_1 CPS + \beta_2 \text{CPS} * \text{QCHANGE}_{CPS} + \beta_3 X + \varepsilon \qquad (2)$$

Where UNINS is a dummy for being uninsured (i.e., the probability of being uninsured is the average between the ACS and the CPS), γ_t is calendar year (controlling for the effect of the year), CPS is an indicator for the data used (CPS=1 if CPS data are used and CPS = 0 if ACS data are used), QCHANGE_{CPS} is a dummy representing the redesigned survey questions (for CPS in 2013 and beyond), X is a vector of covariates (listed below) that are controlled for in the model, and i denotes the individual. In this application, the primary parameter of interest is β_2; it represents the effect of the questionnaire change in the CPS. The regression is estimated separately for the full sample and by subgroup. At the individual level, we use the estimated coefficients to predict the probability of being uninsured for each individual in the data set. Now, we have a predicted individual value for the traditional questions by subtracting the parameter that is due to the questionnaire change (β_2) (see Equation 3). By subtracting the effect of the questionnaire effect from the redesigned estimate, this model also produces a predicted estimate of the traditional uninsured rate of the CPS in 2009 through 2014. All ACS data are removed from the analysis at this point.

$$if \ year \geq 2013 \ then \ Pred(Traditional \ Question_{i,}) = Pred(UNINS_i) - \beta_2$$

$$else \ Pred(Traditional \ Question_i) = Pred(UNINS_i) \qquad (3)$$

After the individual prediction of the traditional questions is done, we take the weighted mean for our final national estimate, by year. In the regression, we control for age, race, sex, Hispanic origin, disability status, citizenship, receipt of Supplemental Nutrition Assistance Program (SNAP) benefits, income-to-poverty ratios, and living in a metropolitan statistical area. In addition to the above characteristics, we control for several interactions between terms. Details are listed in Subsection 6.1., Appendix 1.

We denote this collection of control variables as X in Equation 2. Without these covariates in the model, the results of the person-level model would be equivalent to those of the aggregate-level approach (Equation 1) if the person-level model had only used

2009–2013 data (the model uses 2009 to 2014 data). We use the aggregate-level model as a robustness check for the person-level model, since the simplest form of the person-level model is very close to the aggregate-level model. If the person-model was based on just 2009–2013 data and $B_3 = 0$ (i.e., no covariates except for time, survey, and the years that the CPS redesigned survey questions were in effect), then the person-level estimate of the uninsured in 2013 should be very close to the aggregate model. This robustness check was done and the intuition is borne out in the data (not shown).

4.   Results

The first part of the analysis uses the aggregate method to estimate what the uninsured rate would have been, using the traditional questions (i.e., had the questions not changed). The time trend in the estimated uninsured rate for each survey is shown in Figure 1. Recall that one of the primary reasons the CPS was redesigned was to address the overestimate of the uninsured rate compared to other federal surveys. The uninsured rate in the traditional CPS was higher than in the ACS, despite the fact that the uninsured in the CPS rate reflects the entire calendar year, while the uninsured rate in the ACS reflects a point in time. The uninsured rate in the redesigned CPS, on the other hand, is lower than the uninsured rate in the ACS, which is consistent with the expectations that someone is less likely to be uninsured for an entire calendar year than on any given day in a year.

  The aggregate model predicted the uninsured rate in 2013 to be 15.3% (Table 2). The predicted year-to-year change in the uninsured rate from 2012 to 2013 is, by design, equivalent to the change in the uninsured rate in the ACS over that period: 0.1 percentage point (not statistically significant). As the 2013 CPS uninsured rate using the redesigned questions was 13.3%, the estimated questionnaire effect was 1.9 percentage points.

  The next part of the analysis uses person-level models to predict the questionnaire effect and the traditional uninsured rate in 2013, while controlling for possible confounding covariates. Because the fully specified model has 90 variables and there are 21 subgroups, we do not show all of the regression parameters (available upon request). The person-level model shows that the estimate of the questionnaire effect was 1.7 percentage points (15.0% for the predicted traditional CPS less 13.3% for the redesigned CPS), not statistically different from the questionnaire effect derived from the aggregate-level model (see Table 2). The predicted uninsured rate for 2013 using the person-level model is 15.0%, slightly lower than the person-level prediction in 2012. In addition, when comparing the predicted estimates for 2013, we find that the person-level model produces a slightly lower estimate than the aggregate-level model (see Subsection 6.2., Appendix 2 for a discussion of how the standard errors were calculated). Overall, the predictive power of the person-level model was 21%. The degree of the predictive power varied by subgroup. For example, the population aged 65 and over had an R-squared of 2%, which was in line with the expectation that this group would be relatively unaffected by the questionnaire change.

  The person-level model allows us to examine variation in both the questionnaire effect (in 2013) and change in the uninsured rate (2012 to 2013) by subgroup, including race and Hispanic origin, age, low-income status, and labor force status (for adults aged 19 to 64). The data show variation in the effect of the questionnaire change by subpopulation group,

Table 2. Observed and predicted uninsured rates, by year and survey.

| | Observed | | | | | | Predicted traditional CPS | | | |
| | Traditional CPS | | Redesigned CPS | | ACS | | Aggregate model | | Person-level model | |
Year	Percent	SE	Percent	SE	Percent	SE	Percent	SE	Percent	SE
2009	16.1	(0.14)			15.1	(0.05)			16.2	(0.08)
2010	16.3	(0.15)			15.5	(0.05)			16.2	(0.08)
2011	15.7	(0.12)			15.1	(0.04)			15.7	(0.08)
2012	15.4	(0.13)			14.8	(0.04)			15.4	(0.08)
2013			13.3	(0.12)	14.5	(0.04)	15.3	(0.08)	15.0	(0.08)
2014			10.4	(0.13)	11.7	(0.04)			12.3	(0.08)

Source: 2009–2014 one-year American Community Surveys (restricted data), 2010–2015 Current Population Survey Annual Social and Economic Supplements.

Note: Traditional CPS refers to the official uninsured estimate from the CPS for calendar years from 2012 and before. Redesigned CPS refers to the official uninsured estimate from the CPS for calendar years 2013 and beyond.

Note: The effect of questionnaire change in 2013 is calculated as follows. For the aggregate model, the questionnaire effect is 1.9 percentage points and is calculated by subtracting the redesigned CPS estimate (13.3) from the predicted traditional CPS estimate (15.3%). For the person-level model, the questionnaire effect is 1.7 percentage points, which is β_2 in Equation. 2.

Table 3. Predicted questionnaire effect and year-to-year change for subgroups, person-level model, 2012 to 2013.

	Questionnaire effect			Predicted uninsured rate						
	2013			2012		2013		Diff. 2012 to 2013		
	β_2	SE		%	SE	%	SE	% pts	SE	
Total Population	−1.7	−0.04	***	15.4	−0.08	15.0	−0.08	0.4	−0.12	**
Race and Hispanic origin										
Non-Hispanic White	−1.0	−0.05	***	10.9	−0.10	10.6	−0.10	0.3	−0.15	+
Non-Hispanic Black	−3.0	−0.21	***	18.7	−0.29	18.5	−0.32	0.2	−0.43	
Hispanic	−3.6	−0.16	***	29.4	−0.22	28.1	−0.22	1.2	−0.31	***
Age										
Under 65	−2.0	−0.05	***	17.6	−0.09	17.2	−0.09	0.4	−0.13	**
Under 19	−1.5	−0.09	***	9.2	−0.16	8.9	−0.15	0.3	−0.22	
Aged 19 to 64	−2.2	−0.06	***	21.0	−0.12	20.6	−0.13	0.4	−0.18	*
Aged 65 and over	−0.2	−0.04	***	1.5	−0.19	1.6	−0.17	−0.1	−0.25	
Low income by age										
All ages	−3.8	−0.12	***	26.5	−0.18	25.9	−0.18	0.7	−0.25	**
Under 65	−4.3	−0.15	***	30.3	−0.20	29.5	−0.20	0.8	−0.28	**
Under 19	−2.7	−0.17	***	13.6	−0.26	13.1	−0.26	0.5	−0.37	
Aged 19 to 64	−5.2	−0.20	***	40.2	−0.30	38.9	−0.31	1.3	−0.43	**
Aged 65 and over	−0.8	−0.10	***	2.7	−0.37	2.9	−0.35	−0.1	−0.51	
Labor force status (19–64 year olds)										
Full time	−1.8	−0.07	***	16.9	−0.17	16.7	−0.19	0.2	−0.25	
Part time	−2.4	−0.23	***	24.2	−0.44	23.8	−0.41	0.4	−0.60	
Unemployed	−3.2	−0.48	***	40.5	−0.88	38.1	−0.94	2.4	−1.29	+
Not in labor force	−3.3	−0.18	***	23.3	−0.34	23.2	−0.3	0.1	−0.48	

Source: Authors' calculations from 2013–2014 Current Population Survey Annual Social and Economic Supplements.

Note: ***p < .001, **p < .01, *p < .05, + p < .1 for t-test.

but little variation in the year-to-year change in the uninsured rate (Table 3). Among the race and Hispanic origin subgroups, the effect of the questionnaire change on the uninsured rate was the lowest for non-Hispanic Whites, who experienced a change of almost half of the effect on the total population. Non-Hispanic Blacks and Hispanics, on the other hand, had decreases in their uninsured rates due to the questionnaire change that were greater than the average for the population. Non-Hispanic Whites and Hispanics experienced increases in their uninsured rates from 2012 to 2013, though the change was greater for Hispanics.

The age group with the largest questionnaire effect on the uninsured rate was adults aged 19 to 64. Children under age 19 had a difference that was about half of the change for the working-age adults, and adults aged 65 and over showed the smallest difference in the uninsured rate between the traditional and redesigned questions. When examining the low-income population (family income is less than or equal to 200% of the Income-to-Poverty-Ratio), the effect of the questionnaire change is about twice as large as the effect for the total population. The age pattern of the questionnaire effect on the uninsured rate for the low-income population is consistent with the age pattern for the total population: the smallest effects are associated with the elderly and the largest effects are associated with working-age adults. In these low-income groups, only the low-income population overall, working-age adults, and people under 65 experienced a change in the uninsured rate from 2012 to 2013.

We also examined labor force status for working-age adults and found that the uninsured rate differed less for workers than nonworkers when the questionnaire changed. The questionnaire change had the largest effect for the unemployed and adults not in the labor force, a smaller effect for part-time workers, and the smallest effect on the uninsured rate for full-time workers. Between 2012 and 2013, unemployed adults were the only labor force category to experience a change in the uninsured rate.

5. Conclusion

In 2014, new health insurance questions replaced the traditional questions in the CPS. This change established a disruption in the time trend for the CPS, where estimates from 2012 and earlier cannot simply be compared to the estimates from 2013 and beyond without disaggregating the effect of the questionnaire change from the time trend. While the redesigned questions improved health insurance estimates by making them more in line with other federal surveys, it is also important to maintain the historical time trend, so that researchers and policymakers can take a broader view of trends in health insurance coverage from before the ACA's implementation in 2010. This article fills that gap, by predicting what the uninsured rate would have been if there was not a change in the health insurance questions.

Using the year-to-year change in the uninsured rate as measured by the ACS, together with the uninsured rate from the traditional questions in the CPS in 2009 through 2012, we predicted that the uninsured rate would have been 15.3% in 2013 using aggregate-level data, unchanged from the level in 2012. Using person-level data, we employed a difference-in-difference model to control for demographic and socioeconomic changes in the population, and predicted that the uninsured rate would have been 15.0% in 2013, not

different from the aggregate-level prediction but slightly lower than the person-level prediction in 2012. Both the aggregate- and person- level models could be used to evaluate other survey redesigns. For the former application, additional research would be needed to determine just how far the model could be extended without over-specifying the model. For the latter, one would need to update the data every couple of years and evaluate the goodness of fit to assure that the bridge still fit.

In addition to continuing the historic time trend, we also estimated what effect changing the CPS questionnaire had on the uninsured rate in 2013, the first year of data that collected health insurance using the redesigned questions. We found that the redesigned questions reduced the uninsured rate by about 1.7 percentage points using the person-level model to control for confounding covariates (not statistically different from the aggregate-level estimate). The questionnaire effect varied by subpopulation, and was greater, in general, for the groups that had higher uninsured rates in 2012 as measured by the traditional questions. For instance, the questionnaire change had the smallest effect on the uninsured rate for children, seniors, and adults working full time. These findings were consistent with the expectation that populations with higher rates of coverage would be less affected by the questionnaire change. Another finding was that the effect of the questionnaire change for non-Hispanic Whites was lower than average and lower than for the other race and Hispanic origin groups. This finding may be explained by a disproportionate number of non-Hispanic Whites over age 65, which is consistent with previous research (Day 2013). Finally, we found that the questionnaire change had the largest impact for low-income working-age adults. This means that low-income adults are reporting insurance more often using the redesigned questions as compared with the traditional questions.

Due to changes in the demographic and socioeconomic composition of the population between 2012 and 2013, the aggregate-level model produced a different predicted uninsured rate in 2013 than did the person-level model. These changes in the population also explain why the aggregate-level model did not produce a change in the uninsured rate between 2012 and 2013, while the person-level model showed a slight decrease. Therefore, when possible, it is preferable to use a difference-in-difference regression on person-level data to control for overall population changes, but when it is not possible, aggregate-level data produce similar results.

It is important to note that the person-level model is a linear probability model instead of a logit model. We chose to use a linear probability model because it is the standard model in the difference-in-difference context. However, we could have used a logit model.

In this article, we take advantage of the stable relationship between estimates from multiple surveys over time in order to fill in the gap during a disruption in the time series.

Surveys need to change questions for many reasons: to reduce respondent burden, improve validity, and to harmonize questions with other surveys. When survey questions change, there is always a balance between maintaining the time trend and improving the questions. However, as long as there are other sources of data that track in parallel over time with the survey, there does not have to be a tradeoff. This approach can be used when other surveys change questions.

6.   Appendix

6.1.   Appendix 1: Covariates Included in Person-Level Model

Table A1.   Covariates included in person-level model.

Demographic covariates	Health covariates	Interaction terms
Age	Disability status	IPR by survey[1]
0–5	Has a disability	Age 19–25 by year[2]
6–18	Does not have a disability	Receives SNAP
		by 0–138% IPR[3]
19–25		IPR by year[4]
26–34	**Socioeconomic covariates**	SNAP by year[4]
35–44	Supplemental Nutrition	SNAP by 0–138%
	Assistance Program (SNAP)	IPR by year[4]
45–54	Receives SNAP benefits	
55–64	Does not receive SNAP benefits	
65–74	Income-to-poverty ratio (IPR)	
75 and over	0–138% IPR	
Sex	139–199% IPR	
Male	200–299% IPR	
Female	300–399% IPR	
Race and Hispanic Origin	400–499% IPR	
Non-Hispanic	500% and over IPR	
White alone		
Non-Hispanic		
Black alone		
Non-Hispanic other	**Geographic covariates**	
Hispanic	Metropolitan Statistical	
	Area (MSA)	
Citizenship	Lives in MSA	
Citizen	Does not live in MSA	
Not a citizen	Medicaid Expansion	
	State (as of 1/1/14)	
	Lives in expansion state	
	Does not live in expansion state	

Notes:

[1]We include interactions between each IPR level and the survey because income is measured differently between the two surveys (the CPS collects over 50 types of income while the ACS collects only eight).

[2]Interactions between the age group 19 to 25 and each calendar year account for changes in the relationship between health insurance coverage and age over the period due to the 2010 implementation of the dependent coverage provision of the ACA.

[3]An interaction between receipt of SNAP benefits and IPR less than or equal to 138% of the poverty threshold is included because only low-income families are eligible for SNAP benefits.

[4]We control for time-effects (in addition to the direct effect of γ_t), by including interactions for year by all levels of the IPR, year by receipt of SNAP benefits, and a three-way interaction between year, receipt of SNAP benefits, and low-income status (IPR is between 0 and 138% of the poverty threshold).

6.2.   Appendix 2: Variance Estimation

All of the estimates presented in this article, including both the mean and the variance, are weighted estimates. In the article, we focus on the means of the uninsured. Appendix 2 focuses on the modeled variance. We will not focus on direct estimates of variance. Direct estimates of the means and variances use the person data and the post-stratified weights and the standard errors are computed using successive difference replication (Fay and Train 1995).

Aggregate Model, Tabular

For the aggregate model, the Rate and Var(Rate) is calculated directly from the survey. The second step is estimating the variance of the standard error of the prediction.

$$SE_{Pred(CPS2013)} =$$

$$\text{SQRT}\left(Var(Rate_{ACS2013}) + \frac{1}{16}\sum_{k=2009}^{2012} (Var(Rate_{ACS}) + Var(Rate_{CPS}))_k \right) \tag{A1}$$

Person-Model, Regression

The sample size for the ACS is roughly 35 times the sample size of CPS. If we ignore the sample size difference, the contribution of the ACS to the final prediction in the CPS estimate will be overstated. Therefore, we must adjust the final estimates of the mean and the variance. We do this in steps: first, we calculate a person-weight adjustment, and second, we use the adjusted person weight to calculate the variance of the prediction.

Step 1: Person-Weight Adjustment

To account for the complex survey designs in both the ACS and CPS in the regression-based person-model, each year of data for each survey is weight-adjusted to match their effective sample size (Kish 1965) for the uninsured rate. The Rate and Var(Rate) is calculated directly from the survey. The ESS is calculated (Equation A2a) on the uninsured rate.

$$ESS_j = (Rate_j)(1 - Rate_j)/Var(Rate_j) \tag{A2a}$$

where j = calendar year, subgroup, and survey.
    An adjustment factor is applied (Equation A2b).

$$Adjust_j = ESS_j/Sum\ of\ person\ weights_j \tag{A2b}$$

An adjusted/final person-weight (FPW) is used for the regression (Equation A2c).

$$Final\ Person\ Weight_{i,j} = Adjust_j \times Person\ Weight_i \tag{A2c}$$

Where i = person in group j.

Step 2: Final Estimate of Variance

Using the FPW as the weight for the regression model, we run the model. For each observation that comes from the CPS, the relevant output for calculating the standard error of the traditional questions ($SE(\hat{Y}_j)$) are: FPW, prediction error under the model ($SE(\hat{Y}_{i,j})$), calendar year, and subgroup. All of the ACS observations are dropped. At the person-level, calculate an adjustment to the person-level variance of the prediction of being uninsured.

$$P1\_VAR_{i,j} = FPW_{i,j}^2 \times SE(\hat{Y}_{i,j})2 \tag{A3}$$

Equation A4 is the standard error of the weighted predicted mean for our final national estimate, by year and subgroup.

$$SE(\hat{Y}_j) = \text{SQRT}\left(\sum P1\_VAR_{i,j}\right) / \sum FPW_{i,j}) \tag{A4}$$

7. References

Bailey, J. and A. Chorniy. 2016. "Employer-Provided Health Insurance and Job Mobility: Did the Affordable Care Act Reduce Job Lock?" *Contemporary Economic Policy* 34(1): 173–183. Doi: http://dx.doi.org/10.1111/coep.12119.

Blewett, L.A. and M.E. Davern. 2006. "Meeting the Need for State-Level Estimates of Health Insurance Coverage: Use of State and Federal Survey Data." *Health Services Research* 41(3): 946–975. Doi: http://dx.doi.org/10.1111/j.1475-6773.2006.00543.x.

Bradley, C.J., D. Neumark, and M. Motika. 2012. "The effects of health shocks on employment and health insurance: the role of employer-provided health insurance." *International Journal of Health Care Finance and Economics* 253–267. Doi: http://dx.doi.org/10.1007/s10754-012-9113-2.

Call, K.T., G. Davidson, M. Davern, E.R. Brown, J. Kincheloe, and J.G. Nelson. 2008. "Accuracy in self-reported health insurance coverage among Medicaid enrollees." *INQUIRY: The Journal of Health Care Organization, Provision, and Financing* 45(4): 438–456. Doi: http://dx.doi.org/10.5034/inquiryjrnl_45.04.438.

Centers for Medicare and Medicaid Services. 2014. *National Health Expenditure Accounts: Methodology Paper, 2013*. Centers for Medicare and Medicaid Services. Available at: https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/downloads/dsm-13.pdf (accessed March 2018).

Davern, M., H. Rodin, L.A. Blewett, and K.T. Call. 2007. "Are the Current Population Survey Uninsurance Estimates Too High? An Examination of the Imputation Process." *Health Services Research* 42(5): 2038–2055. Doi: http://dx.doi.org/10.1111/j.1475-6773.2007.00703.x.

Day, J. 2013. "Medicare and Medicaid, Age and Income." *Random Samplings*. U.S. Census Bureau. Available at: https://www.census.gov/newsroom/blogs/random-samplings/2013/09/medicare-and-medicaid-age-and-income-2.html (accessed March 2018).

Fay, R.E. and G.F. Train. 1995. "Aspects of Survey and Model-Based Postcensal Estimation of Income and Poverty Characteristics for States and Counties."

Proceedings of the American Statistical Association Conference. Orlando, August 13–17, 1995. Available at: https://www.census.gov/content/ dam/Census/library/working-papers/1995/demo/faytrain95.pdf (Accessed March 2018).

Kaestner, R. and D. Lubotsky. 2016. "Health Insurance and Income Inequality." *The Journal of Economic Perspectives* 30(2): 53–77. Doi: http://dx.doi.org/10.1257/jep.30.2.53.

Kenney, G. and V. Lynch. 2010. "Monitoring Children's Health Insurance Coverage Under CHIPRA Using Federal Surveys." Chapter. 8, In *Databases for Estimating Health Insurance Coverage for Children: A Workshop Summary*, edited by Thomas Plewes, 65–82. Washington, DC: National Academies Press. Doi: http://dx.doi.org/10.17226/13024.

Kish, L. 1965. *Survey Sampling*. New York: John Wiey & Sons, Inc.

McGeary, K.A. 2009. "How Do Health Shocks Influence Retirement Decisions?" *Review of Economics of the Household* 7(3): 307–321. Doi: http://dx.doi.org/10.1007/s11150-009-9053-x.

O'Hara, B. and K. Caswell. 2013. *Health Status, Health Insurance, and Medical Services Utilization: 2010*. Current Population Reports, P70-133RV. Washington: U.S. Census Bureau.

Pascale, J. 2016. "Modernizing a Major Federal Government Survey: A Review of the Redesign of the Current Population Survey Health Insurance Questions." *Journal of Official Statistics* 32(2): 461–486. Doi: http://dx.doi.org/10.1515/jos-2016-0024.

Pascale, J., M.I. Roemer, and D. Resnick. 2009. "Medicaid Underreporting in the CPS: Results from a Record Check Study." *Public Opinion Quarterly* 73(3): 497–520. Doi: http://dx.doi.org/10.1093/poq/nfp028.

Pascale, J., M. Boudreaux, and R. King. 2016. "Understanding the New Current Population Survey Health Insurance Questions." *Health Services Research* 51(1): 240–261. Doi: http://dx.doi.org/10.1111/1475-6773.12312.

Smith, J.C. and C. Medalia. 2014. *Health Insurance Coverage in the United States: 2013*. Current Population Reports, P60-250, U.S. Census Bureau. Washington, DC: U.S. Government Printing Office.

Smith, J.C. and C. Medalia. 2015. *Health Insurance Coverage in the United States: 2014*, Current Population Reports, P60-253, U.S. Census Bureau, Washington, DC: U.S. Government Printing Office.