

Working Paper Series
Congressional Budget Office
Washington, D.C.

Sources and Preparation of Data Used in HISIM2—CBO’s Health Insurance Simulation Model

Jessica Banthin
Congressional Budget Office
Jessica.Banthin@cbo.gov

Geena Kim
Congressional Budget Office
Geena.Kim@cbo.gov

Keren Hendel
Congressional Budget Office
Keren.Hendel@cbo.gov

Alexandra Minicozzi
Congressional Budget Office
Alexandra.Minicozzi@cbo.gov

Ben Hopkins
Congressional Budget Office
Ben.Hopkins@cbo.gov

Eamon Molloy
Congressional Budget Office
Eamon.Molloy@cbo.gov

Allison Percy
Congressional Budget Office
Allison.Percy@cbo.gov

Working Paper 2019-04
April 2019

To enhance the transparency of the work of the Congressional Budget Office and to encourage external review of that work, CBO’s working paper series includes papers that provide technical descriptions of official CBO analyses as well as papers that represent independent research by CBO analysts. Papers in this series are available at <http://go.usa.gov/xUzd7>.

The authors thank the following staff of the Congressional Office: Chad Chirico, Katherine Fritzsche, Zhuang Hao, Ed Harris, Jeffrey Kling, Sean Lyons, Sarah Masi, Kevin McNellis, Lyle Nelson, Romain Parsad, Kevin Perese, David Rafferty, Lisa Ramirez-Branum, Robert Stewart, David Weaver, Susan Yeh Beyer, and Chris Zogby for their helpful comments and suggestions; Ezra Porter for fact-checking; and Christine Bogusz and Christine Browne for editing. The authors also thank the staff of the Joint Committee on Taxation for their helpful comments and suggestions.

Linda Blumberg and Matthew Buettgens of The Urban Institute, Christine Eibner of RAND, and Jessica Vistnes of the Agency for Healthcare Research and Quality also provided helpful comments. (The assistance of external reviewers implies no responsibility for the final product, which rests solely with CBO.)

Abstract

The Congressional Budget Office’s new health insurance simulation model, HISIM2, is a microsimulation model, meaning that it uses individual-level input data to simulate population behavior. Beginning with the spring 2019 baseline budget projections, CBO completely revamped the way it models consumers’ and employers’ behavior; the agency also updated to new sources of individual-level input data. This paper describes the input data and the procedures for adjusting and adding to those data to form the sample used by HISIM2. The new model has been designed to use the Current Population Survey as its main source of base data. CBO builds the units of its analysis, such as those that make joint decisions about health insurance, based on family relationships. The agency also edits and imputes variables that are important for modeling health insurance decisions to create the data for the base year, 2015. These variables include demographic characteristics, employment characteristics, income, health coverage options, health spending, employer characteristics, and insurance premiums. Then, CBO projects those characteristics into the future—incorporating additional data from 2016 to 2018, expectations about the economy, and the effects of existing policy—to create input data for each year of the agency’s budget projection period, 2019 to 2029.

Keywords: health insurance simulation model, HISIM2, health insurance, health coverage, health insurance marketplaces, nongroup, employment-based insurance, premiums, Current Population Survey, Medicaid, Affordable Care Act (ACA)

JEL Classification: H51, I10, I13, I18

Notes

Unless otherwise specified, all values are for 2015.

This paper describes the methods that underlie the new version of the Congressional Budget Office's health insurance simulation model, HISIM2, which the agency used to create its spring 2019 baseline projections. CBO makes updates to its model at least annually to incorporate new data, changes in law, and improvements to methods.

Contents

Introduction.....	1
Modifications to CPS Data in the Base Year.....	3
Health Insurance Units.....	4
Marketplace Units.....	5
Tax-Filing Units.....	6
Demographics	7
Employment Characteristics	10
Income.....	14
Health Insurance Coverage Options	17
Health Insurance Coverage Targets	21
Health Spending.....	23
Synthetic Firms	27
Employment-Based Health Insurance Premiums	30
Nongroup Premiums	34
Projecting Data to the End of the Baseline Budget Projection Period.....	34
Income Growth	35
Health Spending and Cost Sharing Over Time.....	36
Adjustments to Person Weights	37
Additional Adjustments	37
Final Data Set.....	38
Appendix.....	39

Introduction

The Congressional Budget Office’s new health insurance simulation model, HISIM2, is designed to analyze the health insurance choices of individuals and families. The agency uses the model to produce a number of different estimates. Most importantly, CBO uses HISIM2 to estimate the major sources of health insurance coverage and associated premiums for U.S. residents under age 65. Those estimates are used, along with other models and tools, to determine the federal outlays, tax revenues, and net federal budgetary effects associated with that coverage to create baseline projections for the Congress for a specified 10-year period. In addition, CBO often uses its health insurance simulation model to estimate the effects of proposed policy changes on health insurance coverage, which then serve as the basis for estimating the effects of the proposed policy changes on the federal budget.

CBO updates its model at least once a year to incorporate information from the most recent administrative and survey data, CBO’s most recent macroeconomic forecast, and relevant judicial decisions, enacted legislation, and administrative actions. For 2019, in addition to making those regular updates, CBO completely changed the way it models consumers’ and employers’ behavior.¹ The individual-level input data have also been updated; CBO has designed HISIM2 to incorporate data from the Current Population Survey (CPS), a large, regularly updated source.

HISIM2 is a microsimulation model, meaning that it uses microdata (that is, individual-level data) to simulate behavior. CBO combines multiple data sets to form HISIM2’s input microdata, and it uses an extensive set of equations to simulate the behavior of individuals, families, and employers in choosing health insurance coverage. This working paper is focused on the model’s data—in particular, the sources of the microdata and the steps CBO has taken to prepare those microdata for use in the most recent version of HISIM2, which will underlie CBO’s spring 2019 baseline projections.² This paper is intended to enhance understanding of CBO’s coverage estimates and underlying methods used to produce those estimates. In addition to describing the preparation of the input data, this paper describes the additional data used to develop insurance coverage targets (or total counts of people with a given type of coverage), a key precursor to estimating parameters of the behavioral model. Other CBO publications describe the structure of

¹ See Congressional Budget Office, “HISIM2—The Health Insurance Simulation Model Used in Preparing CBO’s Spring 2019 Baseline” (April 2019), www.cbo.gov/publication/55097.

² Those projections will be published in May 2019. For the spring 2018 projections, see Congressional Budget Office, *Federal Subsidies for Health Insurance Coverage for People Under Age 65: 2018 to 2028* (May 2018), www.cbo.gov/publication/53826.

the model used to predict how individuals and families make decisions regarding health insurance coverage.³

The microdata come from the Current Population Survey’s Annual Social and Economic Supplement, which surveys about 95,000 occupied households annually on their social and economic characteristics. CPS data provide reliable, timely, and detailed information on many of the key variables needed to model health insurance coverage—including income, employment, and self-reported health status. The value of the CPS data is that they capture the relationships among the key variables listed above and provide a large, representative data set for the civilian, noninstitutionalized population living in the 50 states and the District of Columbia.⁴ For example, the CPS data show how eligibility for insurance coverage varies by employment and income. Compared to some other large household surveys, the CPS collects more detail on sources of income and family relationships, which is critical for determining eligibility for public insurance and premium subsidies. Modeling how individuals and families would respond to a change in policy in HISIM2 requires data on such associations.

To improve the accuracy of the CPS data and how well they match administrative data, such as the number of people enrolled in Medicaid, CBO modifies those data in several ways. Like the data from other household surveys, the CPS data reflect some reporting error. CBO adjusts variables in the CPS that are likely to be reported with some error so that the distributions of characteristics of people in the HISIM2 sample match those found in administrative data. CBO also supplements the CPS data with additional variables that are needed to model decisions about health insurance coverage. For example, CBO develops measures of health spending and premiums for each CPS respondent for each private insurance plan they could consider.

The first half of this paper describes in detail how CBO prepares CPS data to create the input data set for the initial year of the simulation, known as the base year, and it also lists the other sources of data used to develop health insurance enrollment targets. The second half of the paper describes how CBO projects the base year data into the future, creating an input data set for the full projection period. The data preparation described in this paper refers to CPS data collected in 2016, which provide information about individuals and households in calendar year 2015. That year serves as the base year for CBO’s baseline projections in the spring of 2019, for the years 2019 to 2029. The base year data are therefore projected forward from 2015 to 2029.

³ See Congressional Budget Office, “HISIM2—The Health Insurance Simulation Model Used in Preparing CBO’s Spring 2019 Baseline” (April 2019), www.cbo.gov/publication/55097.

⁴ The population represented by HISIM2 is the same as that covered by the CPS sample. It excludes members of the armed services, prisoners, and residents of long-term care hospitals and nursing homes.

CBO continually works to improve the accuracy of the model and to reflect real-world developments. HISIM2 is supplemented with other models, such as the Joint Committee on Taxation’s (JCT’s) tax models and CBO’s Medicaid and Medicare models. HISIM2 generates coverage estimates that are used by those models to estimate the costs associated with those coverage distributions. Those models are used together to develop CBO’s baseline and cost estimates.⁵

Modifications to CPS Data in the Base Year

The CPS provides the broad set of base data, but CBO relies on other administrative and survey data sets to improve the accuracy of specific measures or incorporate additional information. For example, the National Health Interview Survey (NHIS) and the Medical Expenditure Panel Survey—Household Component (MEPS-HC) both provide reliable estimates of the number of people who are uninsured, and the Medical Expenditure Panel Survey—Insurance Component (MEPS-IC) provides accurate estimates of employment-based coverage and health insurance offered to employees by firm size.⁶ CBO also uses data from the Centers for Medicare & Medicaid Services (CMS), including the National Health Expenditure Accounts (NHEA) and administrative Medicaid and Medicare data; the Kaiser Family Foundation and Health Research & Educational Trust Employer Health Benefits Survey (EHBS); the Robert Wood Johnson Foundation’s HIX Compare database; state-level marketplace data from New York and California; the National Association of Insurance Commissioners; and administrative tax data and data from information returns.

CBO modifies CPS data in two ways: editing variables to better match other survey and administrative data and adding variables critical for modeling health insurance decisions. For example, the agency edits age, income, and public insurance coverage in the base year. CBO adds variables such as decisionmaking units (referred to as health insurance units, or HIUs), which are based on family relationships and immigration status. See Table A-1 in the appendix for a complete overview of the edits and additions that CBO performs on the CPS data.

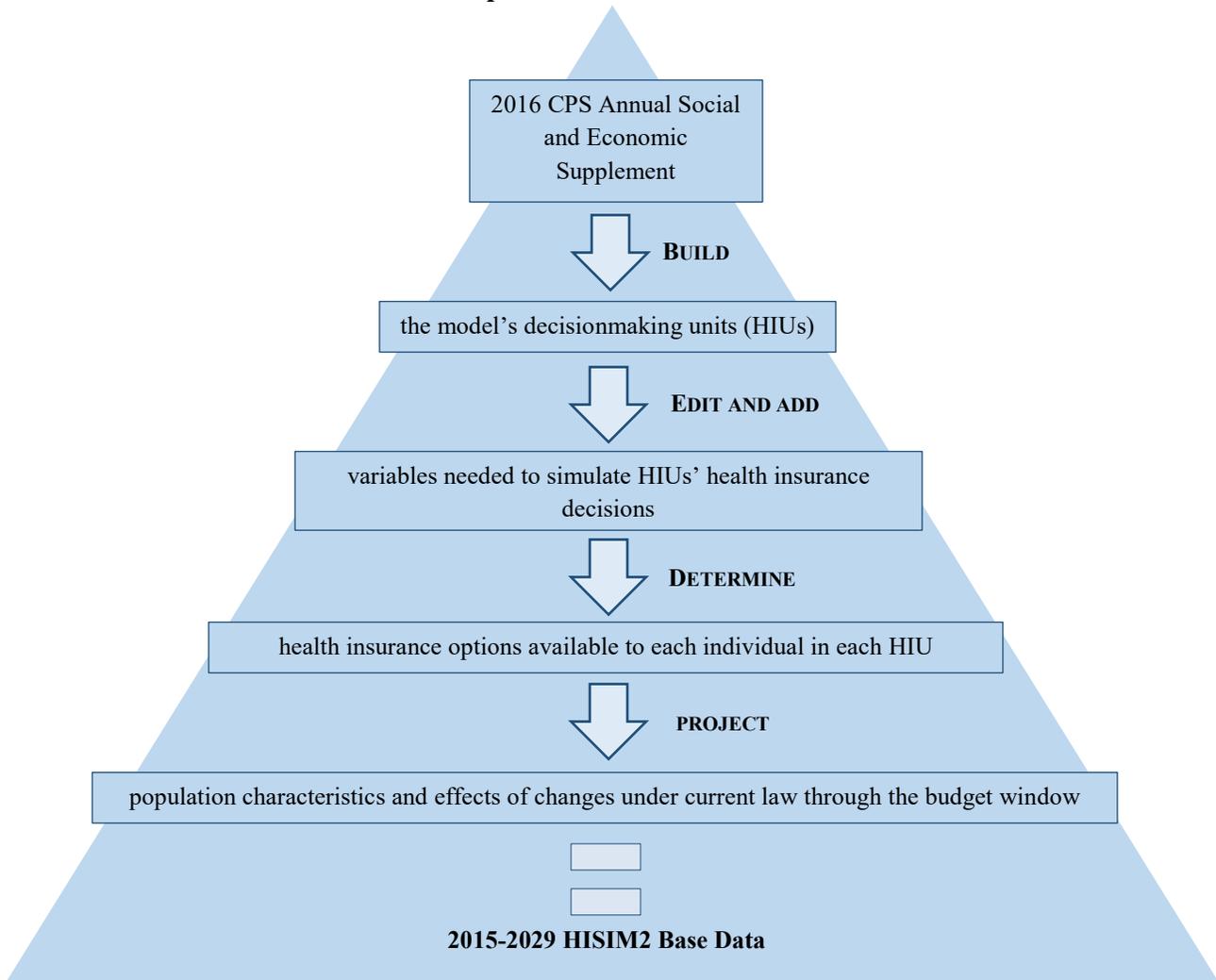
People’s decisions about health insurance also depend on the actual characteristics of their insurance options (for instance, plan premiums and deductibles). Therefore, CBO calculates expected health spending and estimates health insurance premiums for each person in the sample for each of their potential plan options. In addition, CBO creates synthetic firms to model firm-level characteristics for each worker, all of which are described in this section.

⁵ For more information, see Congressional Budget Office, *How CBO and JCT Analyze Major Proposals That Would Affect Health Insurance Coverage* (February 2018), www.cbo.gov/publication/53571.

⁶ The MEPS-IC does not include federal employees, so CBO supplements that information with data from the Office of Personnel Management. For more information, see the [Health Insurance Coverage Targets](#) section.

Figure 1 provides an overview of the major components of data preparation, including the major modifications to the CPS data.

Figure 1.
Overview of the HISIM2 Data Preparation Process



HISIM2 is CBO's health insurance simulation model.

CPS = Current Population Survey; HIU = health insurance unit.

Health Insurance Units

The fundamental objective of HISIM2 is to model the health insurance decisions of people under age 65 in the United States. Because many insurance products can cover multiple people in a

family, the insurance decisions of people in the same family are not independent.⁷ Therefore, CBO refers to household members who make health insurance decisions together as health insurance units. An HIU is the set of individuals who could generally be covered by a family plan if an employer offered that plan.⁸ People within the same HIU may not qualify for the same type of coverage and do not necessarily choose the same coverage option. Rather, the members of an HIU consider one another when making individual choices.

CBO uses CPS data on family relationships to build HIUs and uses those units to simulate decisions about coverage for each person in a given HIU. The CPS data are collected at the household level. Information on family relationships among the people who live in the household is then used by the Census Bureau to identify families. The CPS family unit includes everyone living in the household who is related by birth, marriage, or adoption. This family unit is too broad to be used as the decisionmaking unit in CBO's model. For example, a grandparent living with his or her adult child's family could be part of their CPS family but would generally not be eligible for coverage under the same family plan. An HIU may consist of an individual—for example, a single adult with no spouse or dependents who makes choices about his or her own coverage. Multiperson HIUs generally comprise a head of household; a spouse, if applicable; and any related children under the age of 26.⁹ For simplicity, children under age 26 who are married or have a child and are not claimed by their parents for tax purposes (based on CBO's imputed tax-filing status) are assigned to their own HIU. Similarly, individuals under age 26 who are not living in the same household as their parents are treated as their own household. The CPS does not provide enough information to connect those individuals to their parents' household.

Marketplace Units

HIUs are fundamental to modeling health insurance decisions in HISIM2, but CBO develops additional units of analysis known as marketplace units (MPUs) to help determine eligibility for subsidies in the health insurance marketplaces. Subsidies for insurance plans in the marketplaces depend, in part, on an MPU's modified adjusted gross income (MAGI) and the number of people in the MPU.¹⁰ An MPU comprises a head of household and any individuals for whom the head

⁷ For more information, see Irena Dushi and Marjorie Honig, "Price and Spouse's Coverage in Employee Demand for Health Insurance," *American Economic Review*, vol. 93, no. 2 (May 2003), pp. 252–256, <http://doi.org/10.1257/000282803321947146>.

⁸ Employment-based insurance coverage is usually offered for an employee only (single plan) or for the employee and any number of dependents (family plan). In this paper, a family plan is any plan that covers two or more people, including plans that are offered to an employee plus only one dependent.

⁹ Under the Affordable Care Act, dependents under the age of 26 may be covered by their parent's employment-based or nongroup insurance plan.

¹⁰ MAGI is adjusted gross income plus untaxed foreign income, nontaxable Social Security benefits, and tax-exempt interest, as applicable. For more information, see HealthCare.gov, "Modified Adjusted Gross Income (MAGI)" (December 10, 2018), www.healthcare.gov/glossary/modified-adjusted-gross-income-magi/.

of household claims a personal exemption (based on CBO's imputed tax-filing status), including a spouse, dependent children, and any other dependent relatives, in accordance with the Affordable Care Act (ACA).¹¹ In HISIM2, CBO uses CPS data on family relationships and reported income to determine who is the head of the household, who are tax dependents, and who belongs in an MPU. Tax dependents required to file taxes under CBO's imputed tax-filing status exist in their own tax-filing units (TFUs) but are included in the MPU with the head of household that claims them.

MPUs mainly differ from HIUs because they can include dependent relatives.¹² In addition, the two types of units have fundamentally different purposes in HISIM2. HIUs are the decisionmaking units in HISIM2, whereas MPUs are used to determine MAGI and the MPU's resulting eligibility for premium tax credits.

Tax-Filing Units

CBO develops TFUs for the purpose of calculating specific income measures. TFUs are mainly used to estimate marginal tax rates (MTRs) and adjusted gross income (AGI).¹³ MTRs and AGIs provide valuable information on disposable income used to pay for health insurance and help determine eligibility for public insurance. CBO constructs TFUs for all people in the sample, even those who do not file income tax returns because their income falls below the filing threshold.¹⁴ TFUs include people who would be eligible to file taxes together on a single tax return.

CBO constructs TFUs using self-reported information on household income, demographics, and family structure. CBO groups people into TFUs in accordance with tax law, adjusting for various family structures and ensuring that the population closely reflects the demographic characteristics in administrative tax data.¹⁵ Filers are TFUs who are imputed to file a federal tax return, including those with income above the filing threshold in 2015 and those with wages or another incentive to file taxes (for example, to claim the earned income tax credit). All other

¹¹ For more information on who is included in the calculation of subsidy eligibility under the ACA, which is the basis for determining who is included in MPUs in HISIM2, see HealthCare.gov, "Who to Include in Your Household" (December 12, 2018), www.healthcare.gov/income-and-household-information/household-size/.

¹² For more information on who qualifies as a dependent relative, see Internal Revenue Service, *Publication 501: Dependents, Standard Deduction, and Filing Information* (December 2018), pp. 17–22, www.irs.gov/pub/irs-pdf/p501.pdf.

¹³ AGI refers to total income for the tax year minus certain deductions, including contributions to individual retirement accounts, alimony paid, and student loan interest. For more information, see Internal Revenue Service, "Definition of Adjusted Gross Income" (November 21, 2018), www.irs.gov/e-file-providers/definition-of-adjusted-gross-income.

¹⁴ People who do not file income tax returns have MTRs from payroll taxes.

¹⁵ That administrative tax data includes information from the Statistics of Income and Internal Revenue Service Form W-2s.

TFUs are imputed to be nonfilers. Table 1 presents the number of TFUs that file taxes in HISIM2.

Table 1.
TFUs Imputed to Have Filed Taxes in 2015

Millions of TFUs

	Single Filers Without Dependents	Single Filer With Dependents and Heads of Household	Joint Filers Without Dependents	Joint Filers With Dependents
Age 65 or Older^a	10.9	0.7	11.8	1.0
Under Age 65	57.3	24.4	15.8	27.3

TFU = tax-filing unit.

^a This category comprises single filers and heads of household over the age of 65 as well as joint-filing units in which one or both members are over the age of 65.

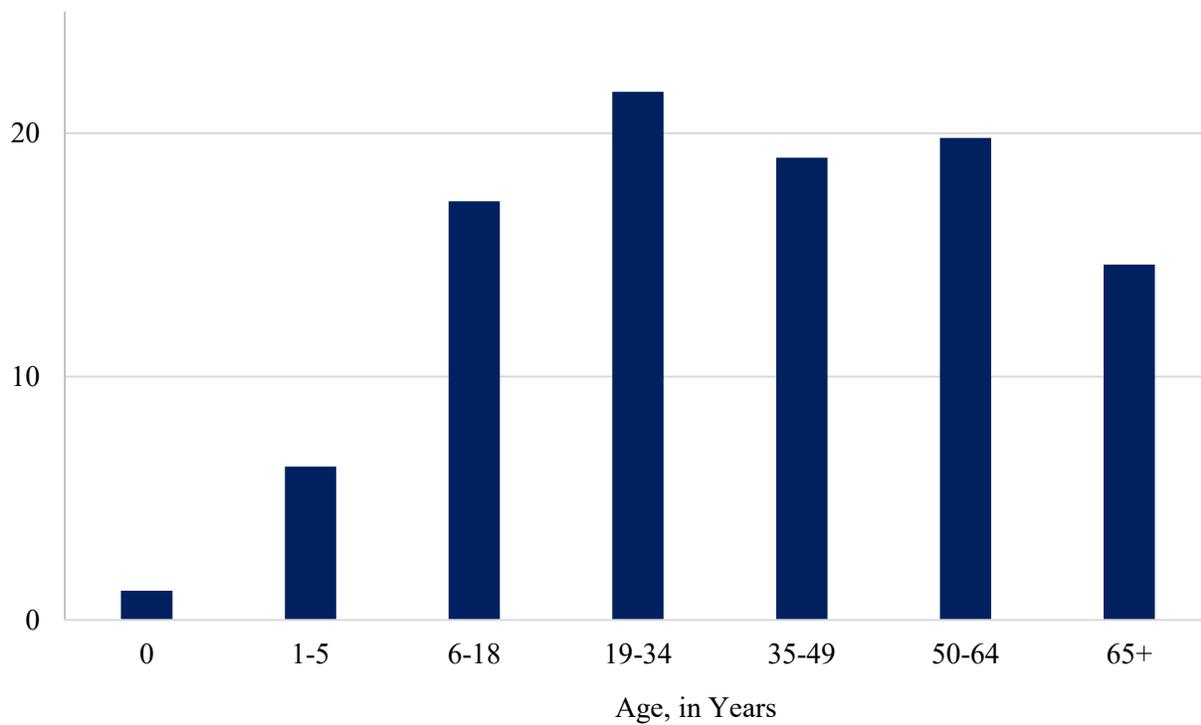
Demographics

Demographic variables affect eligibility for health insurance coverage options. For example, both age and immigration status affect eligibility for public and private insurance options. The CPS contains a broad range of questions about demographics and, in general, CBO uses the values provided in the survey. However, CBO modifies certain variables—age and immigration status, in particular—to more accurately reflect eligibility for health insurance coverage options.

Age. Although CPS respondents report their health insurance coverage in the prior year, they report their age at the time of the survey (which is conducted between February and April of each year). Because their age in the prior year is a critical factor in determining their eligibility for certain health insurance programs in that year and CPS does not report date of birth, CBO subtracts one year from the age reported in the CPS by some respondents to reflect their age on December 31 of the prior year. CBO makes that adjustment for a randomly chosen subset of respondents—about one-fourth of the population—because the problem exists only for people with birthdays between January 1 and the time of the survey. The adjustment is particularly relevant for infants and pregnant women, two groups of people who are eligible for Medicaid, although it is applied to people of all ages.¹⁶ After the age adjustment, some children have ages less than zero. CBO then removes those children from the sample and flags their mothers as being pregnant. Figure 2 presents the distribution of the HISIM2 sample by age and sex after all adjustments for age have been made.

¹⁶ Although eligibility requirements for Medicaid and the Children’s Health Insurance Program vary by state, most cover children age 18 or younger. In addition, young adults under age 26 usually qualify for coverage under their parent’s insurance.

Figure 2.
Ages of People in the HISIM2 Sample, 2015
Percent



HISIM2 is CBO’s health insurance simulation model.

Immigration Status. Immigration status affects eligibility for Medicaid, the Children’s Health Insurance Program (CHIP), Medicare, and marketplace subsidies and thus influences people’s decisions about health insurance. Most noncitizens who are not lawfully present are ineligible for most public insurance options and marketplace subsidies.¹⁷ CPS respondents are asked to report their citizenship status, but noncitizens are not asked details that would indicate whether their presence is lawful. Therefore, CBO imputes whether noncitizens are lawfully present or not. To do so, CBO defines a set of criteria indicating that a noncitizen is likely to be lawfully present. Some of those criteria are whether the person works for the government or in an occupation that requires licensing, whether the person is a veteran or currently in the armed forces, and whether

¹⁷ For example, under the regulations implementing the ACA, coverage and subsidies through the marketplaces are not available for most noncitizens who are not lawfully present. Noncitizens lawfully present owing to the Deferred Action for Childhood Arrivals policy are ineligible to purchase insurance through the marketplaces. For more information, see HealthCare.gov, “Immigration Status and the Marketplace” (accessed February 13, 2019), www.healthcare.gov/immigrants/immigration-status/.

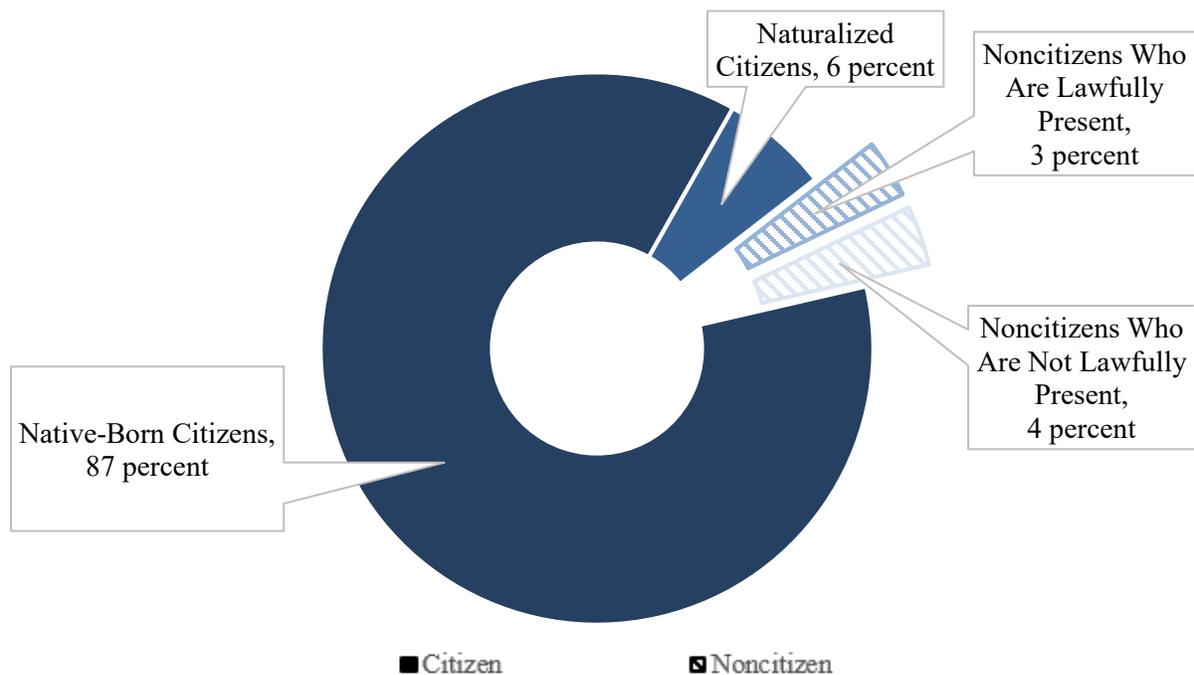
the person arrived in the United States before 1980.¹⁸ Noncitizens who meet one or more of those criteria are identified as lawfully present; the remaining noncitizens are deemed to be not lawfully present.

Following that imputation, CBO estimates that there were about 11 million noncitizens not lawfully present in the United States in 2015.¹⁹ In CBO's estimation, citizens (including naturalized citizens) constituted about 93 percent of the population at that time, whereas noncitizens constituted about 7 percent. Figure 3 presents the proportions of the HISIM2 sample by their immigration status.

¹⁸ That method is drawn from George J. Borjas, *The Labor Supply of Undocumented Immigrants*, *Labour Economics*, vol. 46 (June 2017) pp. 1–13, <https://doi.org/10.1016/j.labeco.2017.02.004>.

¹⁹ The Department of Homeland Security estimates that 12 million people were living in the United States as illegal aliens in 2015; see U.S. Department of Homeland Security, Office of Immigration Statistics, *Population Estimates, Illegal Alien Population Residing in the United States: January 2015* (December 2018), www.dhs.gov/immigration-statistics/population-estimates/authorized-resident. Pew Research Center estimates that 11 million unauthorized immigrants lived in the United States in 2015; see Jeffrey S. Passel and D'Vera Cohn, *U.S. Unauthorized Immigrant Total Dips to Lowest Level in a Decade* (Pew Research Center, 2018), www.pewhispanic.org/2018/11/27/u-s-unauthorized-immigrant-total-dips-to-lowest-level-in-a-decade/. CBO is investigating the extent to which the foreign-born population without legal status is underrepresented in the CPS historically, and the method described in this paper does not incorporate an adjustment for such undercounting.

Figure 3.
Immigration Status of the HISIM2 Sample, 2015



HISIM2 is CBO’s health insurance simulation model.

Employment Characteristics

Information about a person’s labor force participation and employment is important for modeling health insurance decisions because more than half of the U.S. population under age 65 has employment-based health insurance coverage. For example, certain employment characteristics help determine a person’s likelihood of being offered employment-based insurance and his or her eligibility for such coverage. In preparing the CPS data for use in HISIM2, CBO edits and imputes key employment characteristics for each person in the sample. For people who report employment in the survey, the agency uses the employment characteristics of their longest held job. CBO categorizes workers over age 15 as either full time or part time depending on whether they report working 30 or more hours a week (the same threshold used in the ACA to define full-time employment). Among people under age 65, less than 10 percent are categorized as part time, close to half are categorized as full time, and the remaining people are categorized as either unemployed or out of the labor force.

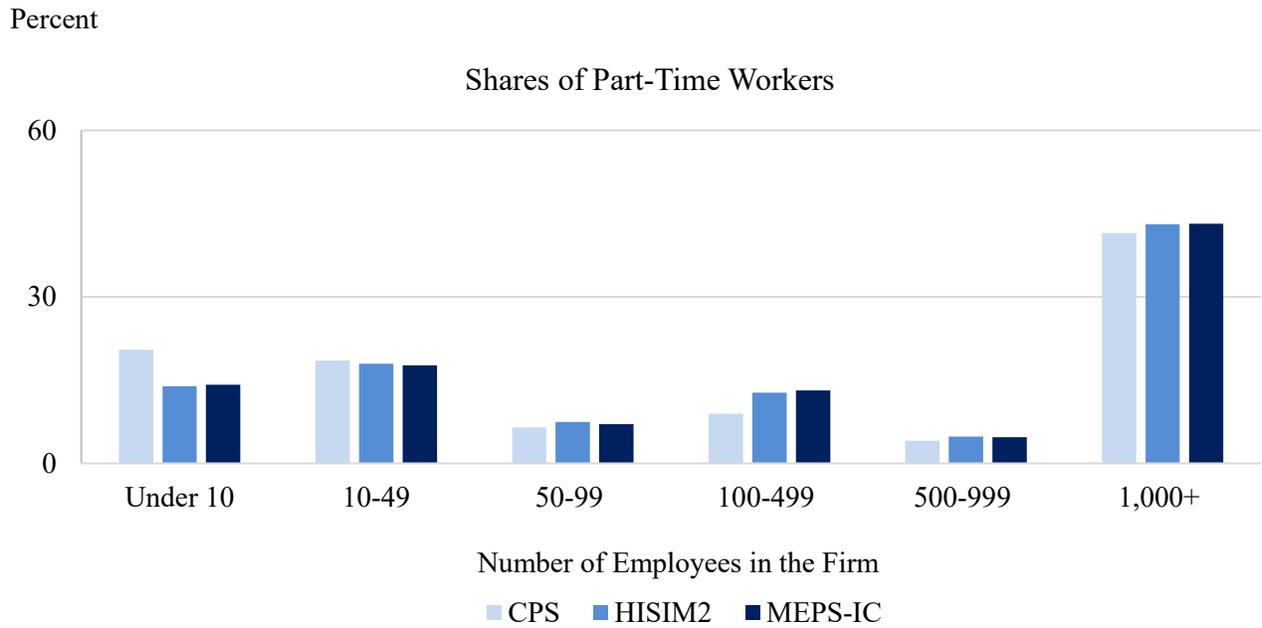
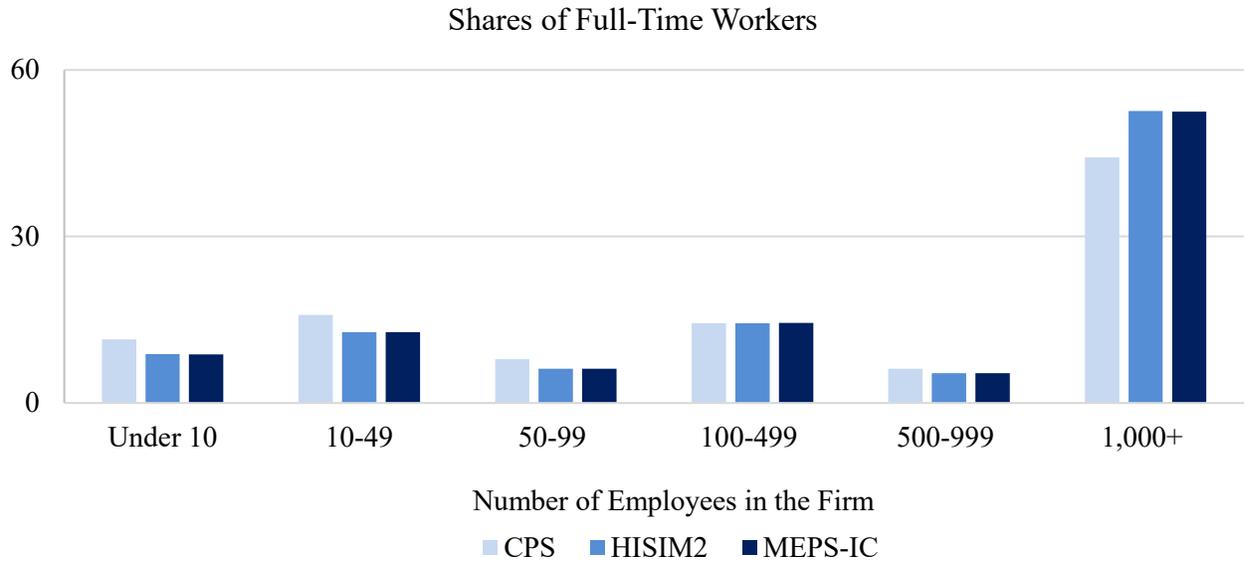
CBO also edits the CPS data on firm size and uses it to perform imputations. Comparisons with data from surveys of firms show that CPS respondents tend to underreport the number of

coworkers at their current firm.²⁰ Therefore, CBO edits and imputes the CPS firm size variable so that it matches the firm size distribution seen in the MEPS-IC data augmented with information on the federal government that is not included in the MEPS-IC.²¹ To do so, CBO first assigns all federal government employees to the largest firm size category, representing firms with 1,000 or more employees. Then, the agency defines targets for the share of workers within each firm size category by full-time status and sector (public or private). People are classified according to whether the Census Bureau imputed their firm size and then ranked by firm size. CBO then imputes a new firm size for some workers until the share of respondents matches the targets from the MEPS-IC for each of the four full-time status and sector categories (full-time public, part-time public, full-time private, and part-time private). Most workers that report a firm size in the CPS retain the value they reported, but workers with Census-imputed values and some workers with reported values are assigned new firm size values by CBO during the process. Figure 4 portrays the shares of full-time and part-time workers in various data sources by firm size. Even though the CPS has more workers in the smaller firm size categories and fewer workers in the larger firm size categories when compared with the MEPS-IC, the HISIM2 sample more closely matches the MEPS-IC distribution of firm sizes.

²⁰ For the purposes of this document, CBO defines a firm as a public or private organization under common management. Firm size is the number of employees working for an organization at all of its locations.

²¹ CBO uses MEPS-IC data to impute firm sizes because it is an employer-based survey, samples state and local governments, and includes measures of employment-based insurance offers. In CBO's view, employers are more likely than employees to accurately report firm sizes and offer status. Additionally, the MEPS-IC sample weights are adjusted to closely match firm size counts from the Census Bureau Business Registry; see D. Kashihara, *Methodology Report #28: Construction of Weights for the 2011 Medical Expenditure Panel Survey Insurance Component* (Agency for Healthcare Research and Quality, 2013), pp. 11–12, https://meps.ahrq.gov/data_files/publications/mr28/mr28.pdf.

Figure 4.
Distribution of Full-Time and Part-Time Workers by Firm Size, 2015
 Percent



HISIM2 is CBO’s health insurance simulation model.

CPS = Current Population Survey; MEPS-IC = Medical Expenditure Panel Survey—Insurance Component.

Employment-Based Health Insurance Offers and Eligibility. CBO imputes both whether a worker's firm offers employment-based health insurance to at least some of its employees and whether that worker is eligible for that insurance in the base year. Those indicators for the firm's offer status and the worker's eligibility are important for determining the insurance choices available to a worker.

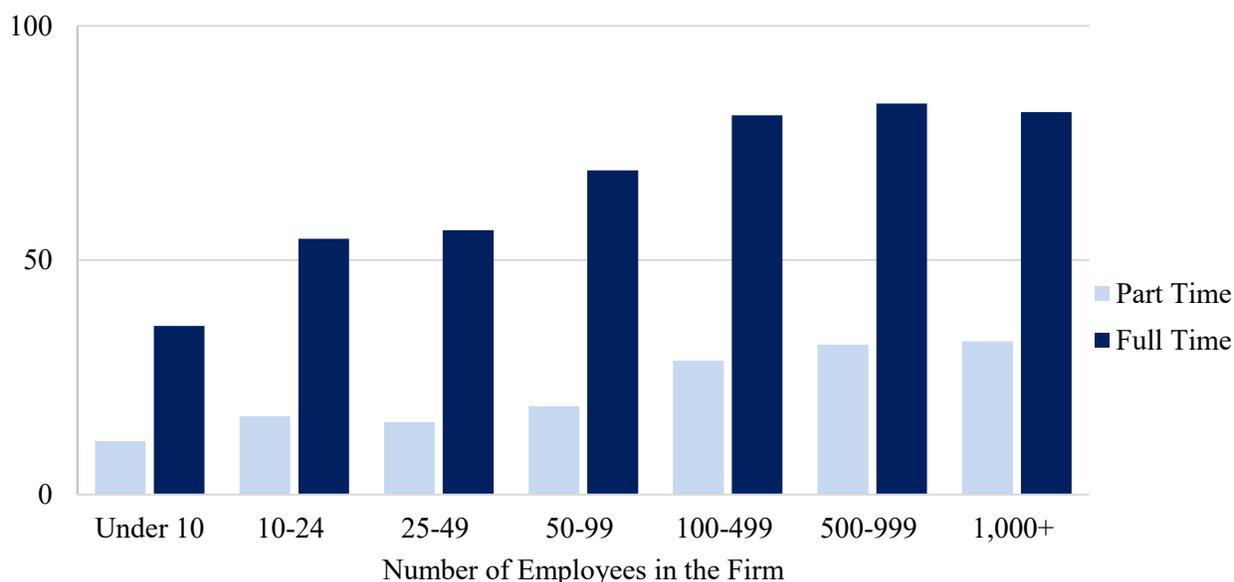
To impute offer status and eligibility, CBO uses MEPS-IC and MEPS-HC data in combination with CPS data. The CPS collects detailed information on a worker's employment characteristics, income, and health insurance coverage for the calendar year before the date of the survey. In contrast, the CPS question on whether a worker's firm offers employment-based coverage to at least some of its workers and whether the worker is eligible for that coverage is asked in reference to his or her status on the date of the survey. As a result, for workers who change jobs or change their employment status, the information on offer status and worker eligibility may not be accurate for the prior year. To address that potential inaccuracy, CBO used MEPS-HC data to estimate the likelihood that those measures would change over time. CBO first regresses a measure of whether a worker works for a firm that offers employment-based coverage to at least some of its employees in the prior year on a set of variables that are common between the CPS and the MEPS-HC (firm-level offer status in the current year, age, earnings, weekly hours worked, sex, marital status, industry, and firm size). The agency next regresses a second measure of whether a worker is eligible for employment-based coverage in the prior year on a similar set of covariates, which includes firm-level offer status in the current year. CBO then applies the regression output to predict a CPS respondent's firm-level offer and individual-level eligibility status in the prior year.

A second potential source of inaccuracy is that employees who are ineligible for employment-based health insurance (for example, part-time, seasonal, or temporary workers) are less likely than eligible employees to know whether their firm offers insurance to at least some of its workers. For that reason, the number of workers at firms that offer insurance may be underreported in the CPS data. To account for such underreporting, CBO uses MEPS-HC data to estimate the likelihood that a person who did not receive an offer of employment-based insurance works at a firm that offers insurance. To that end, CBO first regresses offer status on a set of variables that are common between the CPS and the MEPS-HC (age, earnings, weekly hours worked, sex, marital status, industry, and firm size). CBO then applies the regression output to predict whether a CPS respondent who did not receive an offer of employment-based health coverage worked at a firm that offered such coverage to other employees. The estimate is adjusted to ensure that the overall probability of working at a firm that offers insurance in the HISIM2 sample is consistent with MEPS-IC data. Finally, CBO uses a second imputation to designate which workers at firms that do not offer insurance would be either eligible or ineligible for an employment-based offer (if their firm decided to offer insurance) using the same MEPS-HC data and variables.

About 87 percent of people employed in the HISIM2 sample work at firms that offer employment-based insurance. About two-thirds of the working population work at a firm that offers insurance and qualify for that insurance offer. Figure 5 provides more detail on how access to employment-based coverage varies by firm size for full- and part-time workers in the HISIM2 sample.

Figure 5.
Percentage of Workers Who Work for Firms That Offer Employment-Based Insurance to Some of Their Workers and Who Are Eligible for That Offer, 2015

Percent



To construct premiums for all firms including those that do not offer insurance coverage in the base year, CBO predicts the probability that each worker would choose to enroll in an employment-based insurance plan, or “take up” that plan, if he or she received an offer. The predicted take-up probabilities are used as weights when calculating premiums for each firm.

Income

Income is another important variable for modeling health insurance decisions, because income determines eligibility for marketplace subsidies and various public insurance programs (such as Medicaid and CHIP). The CPS measures most major kinds of income, but CBO makes various adjustments to the CPS income variables in the base year so that the sample’s income measures better reflect the income distribution in the United States according to tax data.

CBO groups income into three categories, referred to for the remainder of this paper as wage, self-employment and property, and other income. Wage income comprises wage and salary income; self-employment and property income consists of self-employment income, capital

gains, interest, dividends, and rental income; and other income includes, for example, public assistance, pensions, and unemployment insurance.²² CBO first augments the measure of self-employment and property income by imputing values for self-employment income and capital gains.²³ The imputations help correct for the critical underreporting of those types of income in the CPS data.

Self-Employment Income. CBO, with help from JCT, uses administrative tax data to impute self-employment income because of the underreporting of self-employment income in the CPS data and the importance of such income to health insurance decisionmaking. In comparison with administrative tax data, CPS data reflect 8 million fewer tax filers under age 65 with self-employment income.

Undercounting the number of people with self-employment income would have particularly significant effects because of the self-employment health insurance deduction, which allows qualified filers to deduct their contribution to nongroup private health insurance premiums from their taxable self-employment income. To ensure that HISIM2 captures how that deduction affects decisions about health insurance, CBO imputes self-employment income to some people who report having zero self-employment income based on a method developed by researchers at the Internal Revenue Service (IRS).²⁴

CBO and JCT use tax data to estimate the relationships between a person’s characteristics and his or her likelihood of earning self-employment income—and, if applicable, how much. CBO uses those estimated relationships to assign to CPS respondents probable values of self-employment income. Then, CBO categorizes TFUs into cells based on wages, demographic characteristics, and imputed probable values of self-employment income. The final step is assigning TFUs, which are ranked within cells in order of their estimated likelihood of having self-employment income, varying amounts of such income. Imputations of self-employment income are constrained so that, for each wage-demographic cell, the final distribution of filers’

²² The “other income” category comprises the following variables, drawn directly from the CPS or derived from CPS variables: Social Security income, Supplemental Security Income, Railroad Retirement benefits, benefits from all other pensions, unemployment insurance, all veterans’ benefits, workers’ compensation, all disability support, child support, alimony, financial assistance, public assistance, and other miscellaneous sources of income.

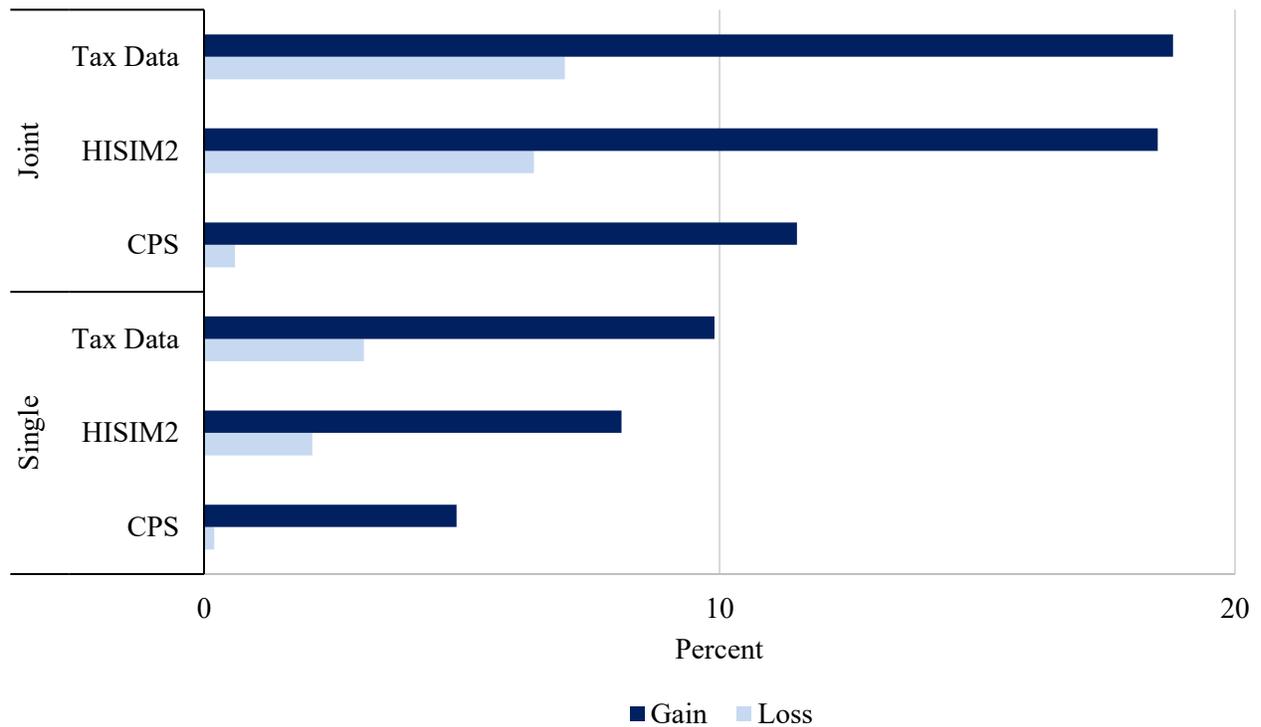
²³ In this paper, self-employment income generally refers to income subject to the Self-Employment Contributions Act tax. CBO also considers business income reported by CPS respondents that is not subject to the tax but might have been reported by a CPS respondent as self-employment income, especially for reporting self-employment losses.

²⁴ For more information, see Brian Erard and others, “Missing Returns vs. Missing Income: Estimating the Extent of Individual Income Tax Filing Noncompliance From IRS and Census Data” (paper presented at the National Tax Association’s 107th Annual Conference, 2014, Santa Fe, Nm., November 13–15, 2014), www.ntanet.org/conference/2014/11/107th-annual-conference-proceedings-2014/; and Internal Revenue Service, *The Individual Income Tax and Self-Employment Tax Nonfiling Gaps for Tax Years 2008–2010* (2016), www.irs.gov/pub/irs-soi/16resconpayne.pdf.

self-employment income matches data from the IRS’s Statistics of Income (SOI) and does not substantially alter the gross income distribution. CBO adds additional constraints so that TFUs in the CPS sample in which no members report employment of any kind are least likely to be imputed self-employment income. CBO expects that those people are less likely to have self-employment income.

Following the imputation, the distribution of self-employment income for 2015 more closely matches the distribution of self-employment income in the SOI data. (Because differences remain between those two distributions, CBO and JCT make additional adjustments in the final steps of adjusting income.) Figure 6 provides additional information about the share of filing TFUs under age 65 with self-employment gains and losses in the CPS, HISIM2 base data, and tax data.

Figure 6.
Share of TFUs Under Age 65 with Self-Employment Income, 2015



HISIM2 is CBO’s health insurance simulation model.

CPS = Current Population Survey; TFU = tax-filing unit.

Capital Gains. Using SOI data, CBO works with JCT to impute an indicator for the presence of capital gains or losses and a dollar value for those gains or losses for CPS respondents. Using JCT’s estimated parameters from a logit model using SOI data and six variables (total income quintile excluding capital gains and binary variables for positive earned income, positive

retirement income, income from interest or dividends exceeding \$1,000, positive rental income, and joint filing status), CBO predicts the likelihood of a capital gain or loss. Parameter estimates from that model are then applied to the CPS data to assign the presence of a capital gain or loss.

Each TFU with a capital loss is assigned the mean capital loss value (less than 6 percent of filers are assigned a capital loss), and TFUs with capital gains are assigned a capital gain value based on the distribution of capital gains for their age, income, and filing status group. The assigned values range from a loss of about \$2,800 to a gain of about \$49 million; the mean value is a capital gain of about \$38,000.

Final Adjustments to Total Income. Once those imputations are complete, CBO computes factors to adjust individual-level income such that income totals for cells based on age, filing status, and wages match totals for similar cells estimated using tax data. CBO classifies the sample into age-wage-filing status cells because there are differences between those groups in patterns of income growth.

Then, CBO uses the National Bureau of Economic Research’s TAXSIM model to estimate federal and state tax liabilities, federal and state MTRs, and AGI using the adjusted income values.²⁵ CBO uses TAXSIM’s AGI outputs to calculate MAGI by aggregating AGIs and taxable Social Security income for members of marketplace units.²⁶ MAGI is used to determine eligibility for subsidies in the marketplaces and various other tax credits. In addition, a TFU’s MTR is an important factor in determining that unit’s potential tax subsidy for purchasing employment-based insurance and demand for health insurance more broadly.

Health Insurance Coverage Options

In preparing the CPS data for use in HISIM2, CBO’s goal is to define the set of health insurance options available to each HIU rather than assign the unit’s coverage in a deterministic approach. Coverage probabilities are estimated at a later stage in the model.²⁷ Although the CPS data provide detailed information on enrollment in various types of coverage, CBO does not rely on that information exclusively for employment-based, Medicaid, CHIP, and nongroup coverage. The main reason CBO does not rely solely on CPS reports of insurance coverage is that CBO’s goal is for HISIM2 to produce estimates and projections of coverage that are consistent with the administrative enrollment and tax data used in federal budgets. In addition, the CPS measures

²⁵ TAXSIM is a collection of programs and data sets used to implement a microsimulation model for the U.S. federal and state income tax systems. See Daniel Feenberg and Elisabeth Coutts, “An Introduction to the TAXSIM Model,” *Journal of Policy Analysis and Management*, vol. 12, no. 1 (1993), <https://doi.org/10.2307/3325474>.

²⁶ See the [Marketplace Units](#) section for more information on MPUs.

²⁷ See Congressional Budget Office, “HISIM2—The Health Insurance Simulation Model Used in Preparing CBO’s Spring 2019 Baseline” (April 2019), www.cbo.gov/publication/55097.

insurance coverage throughout the entire previous year, whereas HISIM2 is designed to produce estimates of average annual enrollment, which are based on insurance coverage reported at a single point in time.²⁸

Therefore, instead of using CPS health coverage variables, CBO uses the information on family relationships, age, income, disability status, state of residence, and offers of employment-based coverage to define the coverage options that are potentially available to each individual within an HIU. CBO defines eligibility for four major categories of health insurance coverage, provided through the following sources: employment, Medicaid, CHIP, and nongroup markets. In addition, CBO allows every person in the sample to have the option of going without coverage.

Coverage through Medicare, military and veterans' health care programs, the Indian Health Service, and other smaller programs is treated differently. CBO does not model eligibility for those programs, as discussed at the end of this section.

Determining offers and eligibility for employment-based health insurance is discussed in the Employment-Based Health Insurance Offers and Eligibility section under [Employment Characteristics](#), and most people can purchase nongroup coverage.²⁹ Therefore, much of the remainder of this section will focus on determining eligibility for public insurance coverage.

Medicaid and CHIP Eligibility. CBO determines whether each person is eligible for full Medicaid benefits or CHIP benefits and the type of eligibility (by age, disability status, and whether the person was made eligible by the ACA).³⁰ This is a key step in the development of the input data, for several reasons. First, members of HIUs can choose to enroll in Medicaid only if they are eligible for it. Second, if the members of the HIU are eligible for Medicaid, they are not eligible for marketplace subsidies. Third, a person's enrollment in Medicaid and his or her type of eligibility affect the federal budget. That is because the federal government pays states for a specified percentage of Medicaid expenditures, called the federal medical assistance percentage (FMAP), and the FMAP for enrollees made eligible by the ACA is higher than the FMAP for Medicaid enrollees who qualify under other eligibility categories.

²⁸ For more information on how CBO determines health coverage, see Congressional Budget Office, *How CBO Defines and Estimates Health Insurance Coverage for People Under Age 65* (May 2018), www.cbo.gov/publication/53822.

²⁹ Noncitizens who are not lawfully present are unable to purchase insurance through the marketplaces. Those people can, however, purchase nongroup insurance outside of the federal and state marketplaces. See the [Immigration Status](#) section for more information.

³⁰ People who qualify for full benefits are eligible to receive the full Medicaid benefit package. This excludes those eligible for only partial Medicaid benefits, such as family planning services or assistance with Medicare cost sharing and premiums.

There are many issues to consider when determining Medicaid and CHIP eligibility. Eligibility is determined by state-specific requirements using monthly income. However, CPS respondents report annual income. Also, a person's Medicaid unit (that is, the members of a household considered for determining eligibility for Medicaid) may not be the same as his or her TFU, MPU, or HIU as the result of various Medicaid- and CHIP-specific exceptions.³¹

To assign eligibility, CBO first constructs a person's Medicaid unit and determines his or her Medicaid unit size, income, and income as a share of the federal poverty level. Medicaid unit income is defined using MAGI-based concepts and specific rules regarding countable income.³² A person is eligible for Medicaid if his or her Medicaid unit income as a share of the federal poverty level is lower than the eligibility income limits of the state in which that person resides.

However, some CPS respondents report being enrolled in Medicaid even though their reported Medicaid unit income exceeds the income thresholds for Medicaid eligibility. One possible reason for the discrepancy is that eligibility is determined by state-specific requirements using monthly income, but CPS respondents report annual income. Those people might have had monthly income low enough to qualify them for Medicaid at some point during the year, even though their annual income was above the threshold. Other reasons for the discrepancy are that those CPS respondents' medical spending might be high enough to allow them to be on Medicaid through their state's medically needy program, or they might have erroneously reported their Medicaid enrollment or income in the CPS.³³ CBO allows some of those seemingly ineligible people who reported being on Medicaid in the CPS to remain eligible in order to account for within-year variations in income.

For all remaining people in the population, CBO compares household income as a share of the federal poverty level with the eligibility thresholds of their state of residence to categorize eligibility into two main groups: people who were made eligible by the ACA and those otherwise eligible. To determine eligibility for CHIP, CBO compares household income as a share of the federal poverty level with the state-specific CHIP income thresholds for eligibility.

³¹ For such exceptions, see Tricia Brooks, *Getting MAGI Right: A Primer on Differences that Apply to Medicaid and CHIP* (Georgetown University Health Policy Institute, 2015), http://ccf.georgetown.edu/wp-content/uploads/2015/01/Getting-MAGI-Right_Jan-30-2015.pdf.

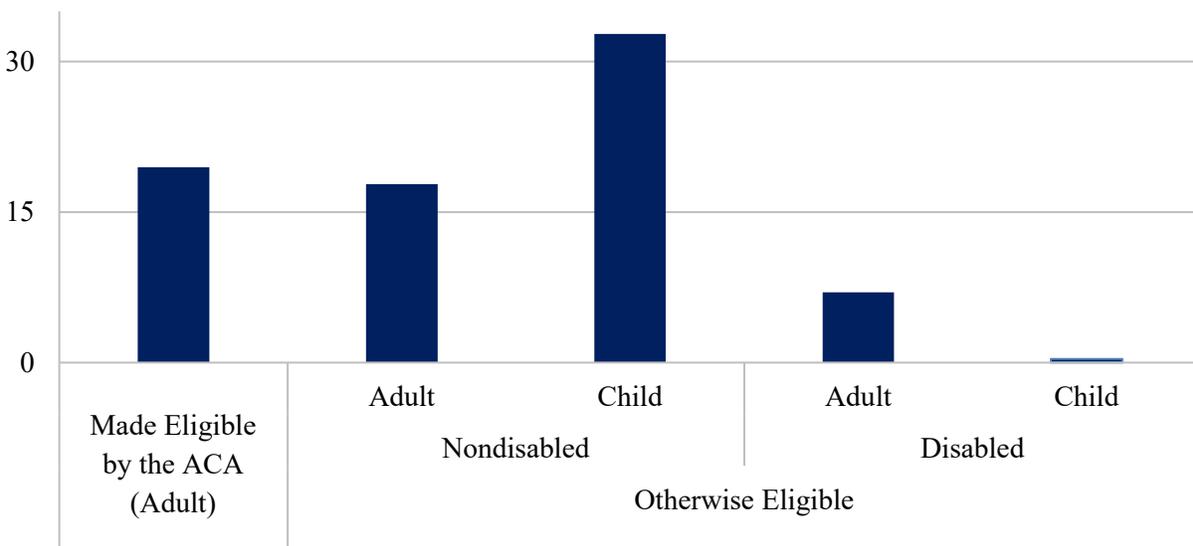
³² For more details, see the previous footnote.

³³ Medicaid enrollment is underreported in data from surveys such as the CPS in comparison with administrative data. For more information, see James M. Noon, Leticia E. Fernandez, and Sonya R. Porter, *Response Error and the Medicaid Undercount in the Current Population Survey*, Working Paper 2016-11 (Center for Administrative Records Research and Applications, U.S. Census Bureau, 2016), www.census.gov/content/dam/Census/library/working-papers/2016/adrm/carra-wp-2016-11.pdf; and Jacob A. Klerman and others, Understanding the Current Population Survey's Insurance Estimates and the Medicaid 'Undercount,' *Health Affairs*, vol. 26, no. 6 (2009), www.healthaffairs.org/doi/full/10.1377/hlthaff.28.6.w991.

Overall, about 30 percent of the entire sample of people under age 65 is flagged as eligible for full Medicaid benefits, and about 75 percent of that group are people who were not made eligible by the ACA. Furthermore, about 5 percent of HISIM2’s sample of people under age 65 is flagged as eligible for CHIP. CBO also categorizes people who are eligible for Medicaid by age and disability status to allow for simulations of policy changes. Figure 7 presents the number of people under age 65 in the HISIM2 sample who qualify for Medicaid, by their eligibility category.

Figure 7.
Number of People Under Age 65 Eligible for Medicaid, 2015

Millions



ACA = Affordable Care Act.

Additional Insurance Options. For certain insurance options, CBO assigns coverage to some people in the sample during the data-preparation phase and does not allow that coverage to change during HISIM2 simulations. CBO does so because there are certain groups of people in the CPS who are less likely to change their health insurance coverage status in response to the types of policy proposals that HISIM2 is designed to simulate. Those groups include Medicare enrollees without an employment-based insurance offer, dual enrollees in Medicaid and Medicare, Medicare enrollees under age 65, foster children with Medicaid, people who are eligible for Medicaid because of a disability, people covered by military or veterans’ health care programs, people with alternative sources of coverage (including student health plans, the Indian Health Service, and foreign sources), policyholders or dependents with employment-based coverage under the Consolidated Omnibus Budget Reconciliation Act, adults with both Medicaid

and employment-based coverage, and dependents of policyholders with employment-based coverage who live outside the policyholder's home.³⁴

Health Insurance Coverage Targets

CBO relies on numerous sources of information in addition to the CPS to develop its own health insurance coverage targets, or total counts of people with a given type of coverage in the base year, broken out by certain characteristics.³⁵ In setting targets, the agency tries to match higher-quality sources, such as administrative enrollment data and data from surveys specifically designed to measure health insurance coverage (for example, the MEPS-HC and the NHIS). In addition, targets are created for health insurance choices that are not measured by the CPS, including the specific metal tier of insurance coverage chosen by an enrollee in the marketplaces.³⁶ One of the challenges of combining information from various data sources to define coverage targets is reconciling the targets so that they sum to the total of the relevant population.

For employment-based health insurance exclusive of military and veterans' health care programs, CBO creates targets for the total number of people covered. Separate targets are created for each combination of type of plan (high-deductible health plan, or HDHP; health maintenance organization, or HMO; and preferred provider organization, or PPO), type of coverage (single or family), and HIU size.³⁷ CBO uses MEPS-IC data and Office of Personnel Management data on the Federal Employees Health Benefits program to estimate the total number of employment-based health insurance policies. The agency combines that estimate with

³⁴ Medicare enrollees without an employment-based insurance offer may be either over or under age 65. Medicare enrollees over age 65 are included in HISIM2 because they may be able to cover their dependents on an employment-based plan if they are offered one. Most Medicare-eligible people under age 65 qualify for Medicare because they participate in the Social Security Disability Insurance program. People covered under military and veterans' health care programs include people with TRICARE coverage, military veterans who report using the Department of Veterans Affairs (VA) health system, and dependents of seriously disabled veterans who are enrolled in the VA's Civilian Health and Medical Program. Because the CPS does not collect representative information for active-duty service members, most of those with TRICARE coverage in the sample are dependents of active-service duty members, retired service members and their dependents, or reservists.

³⁵ Those characteristics include age, income, and HIU size. The breakouts vary by coverage type.

³⁶ Plans in the nongroup market are categorized into metal tiers on the basis of their actuarial value, which is the average percentage of total costs for covered benefits that a plan pays for. "Bronze" plans are those with an actuarial value of 60 percent, "silver" plans are those with an actuarial value of 70 percent, and "gold" plans are those with an actuarial value of 80 percent.

³⁷ HDHPs are high-deductible insurance plans that allow the use of a tax-preferred health savings account to cover expenses not paid by the plans. HMOs are insurance plans in which services obtained outside a specified network of providers are often not covered. PPOs tend to offer wider provider networks, cover services from providers outside of their network, and typically limit costs through cost-sharing arrangements and a deductible.

the average number of dependents per family policy according to the MEPS-HC to estimate the total number of people covered by employment-based insurance.

CBO uses administrative data to develop targets for three of the major categories of public coverage: Medicaid, CHIP, and Medicare. For Medicaid and CHIP, CBO uses CMS’s administrative data—specifically, estimates based on Form CMS-64—and data published by the California Department of Health Care Services.³⁸ CBO combines those estimates with historical data from the Medicaid Statistical Information System (MSIS) and Medicaid Analytic eXtract data to determine the shares of each type of enrollee by age and eligibility status.³⁹ CBO uses the Statistical Enrollment Data System to determine overall CHIP enrollment. CBO combines administrative data from CMS and the Social Security Administration to create a target for the number of people eligible for Medicare under age 65.⁴⁰

For the nongroup market, CBO combines data from CMS’s Medical Loss Ratio reports, data from the National Association of Insurance Commissioners, CMS’s quarterly effectuated enrollment reports, and the MEPS-HC to estimate the total number of people enrolled in the nongroup market in the base year.⁴¹ CBO supplements those sources with detailed data from CMS on effectuated marketplace enrollment through HealthCare.gov and state-level marketplace enrollment reports for New York and California in order to separately identify nongroup enrollment targets by HIU size, metal tier, subsidized versus unsubsidized enrollment (for marketplace enrollment), and age and income (for subsidized marketplace enrollment).

³⁸ Form CMS-64 is the form on which states report actual expenditures on Medicaid and CHIP to the CMS Medicaid Budget Expenditure System/State Children’s Health Insurance Budget and Expenditure System. CBO uses data from the California Department of Health Care Services to estimate enrollment for adults enrolled in California’s Medicaid program (“Medi-Cal”), but not made eligible because of the ACA. See California Department of Health Care Services, “Medi-Cal Certified Eligibles – Recent Trends” (accessed March 26, 2019), www.dhcs.ca.gov/dataandstats/statistics/Pages/Medi-Cal-Certified-EligiblesRecentTrends.aspx.

³⁹ The Medicaid Statistical Information System was a historical program in which states provided data from their claims processing systems to CMS. The final year for which those data are available to CBO was 2011. For more information on MSIS, see Centers for Medicare & Medicaid Services, “Medicaid Statistical Information System (MSIS)” (updated April 3, 2017), www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MSIS/. The Medicaid Analytic eXtract is compiled by CMS using data submitted to MSIS. The most recent year for which Medicaid Analytic eXtract files are available for all states and the District of Columbia is 2012. For additional information on the Medicaid Analytic eXtract data, see Centers for Medicare & Medicaid Services, “Medicaid Analytic eXtract (MAX) General Information” (updated March 2, 2018), www.cms.gov/research-statistics-data-and-systems/computer-data-and-systems/medicaiddatasourcesgeninfo/maxgeneralinformation.html.

⁴⁰ Most people under age 65 who are eligible for Medicare qualify because they receive Social Security Disability Insurance benefits.

⁴¹ CBO uses the MEPS-HC to estimate the percentage of nongroup enrollees who also have coverage through their employer. Doing so allows the agency to ensure the targets sum to the total population.

Nongroup enrollment outside of the marketplaces is estimated from the difference between total nongroup enrollment and marketplace enrollment.

Finally, CBO develops targets for the number of uninsured individuals.⁴² Measuring the number of uninsured people can be done in various ways: by counting people who are uninsured for the full year, those uninsured at any point during the year, and those uninsured at a single point in time during the year. In addition, various surveys—including the NHIS, the MEPS-HC, the CPS, the Survey of Income and Program Participation, and the American Community Survey—provide different estimates of the number of uninsured people in a given year (partly because they use different measures). CBO uses the NHIS to develop targets for the average annual number of uninsured people because its point-in-time measurement is well-suited to estimating the annual average and because CBO assessed it as providing the highest-quality measure. CBO uses the MEPS-HC to develop targets for the distribution of uninsured people by poverty level and HIU size because that survey has higher-quality data on income.

See Table A-2 in the appendix for an overview of the health insurance coverage targets, the data used to estimate them, and whether the decisions of people with that type of coverage are modeled by HISIM2. For additional discussion of the types of coverage that are not modeled by HISIM2, see the Additional Insurance Options section within [Health Insurance Coverage Options](#).

Health Spending

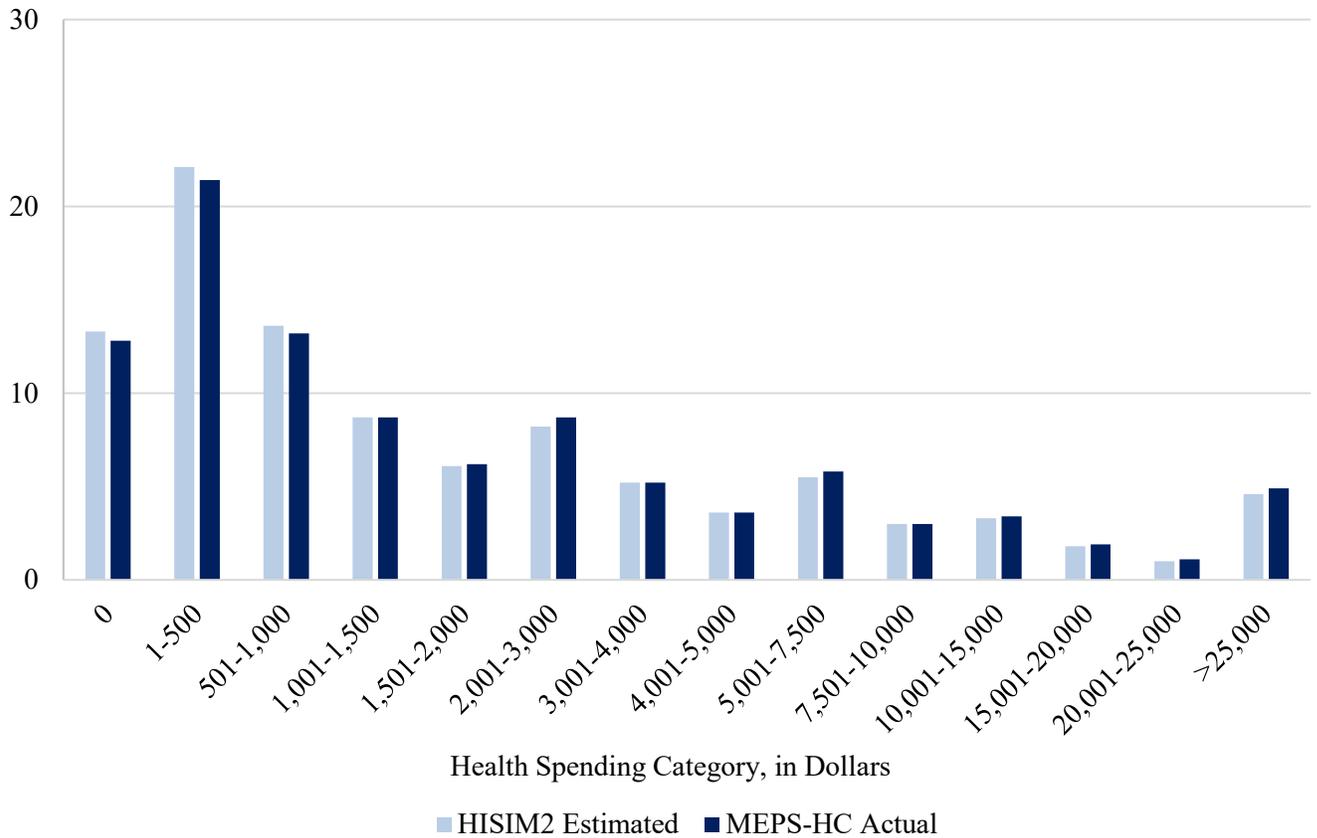
CBO imputes expected spending on health care services for each CPS respondent in order to model decisions about health insurance and estimate health insurance premiums. CBO does so because insurance premiums for a given policy generally reflect the amount an insurer expects to spend on claims for the HIUs covered by that policy, subject to market regulations. The imputed measure is meant to capture the amount a person would expect to spend on health care services, including both spending out-of-pocket and spending by an insurer, if he or she was enrolled in a typical private health insurance plan for the entire year. People without coverage, people with public insurance coverage, or people enrolled in private plans with different cost-sharing obligations may spend more or less in total on health care than the typical private plan enrollee. However, CBO's imputed spending measure is meant to serve as a reference point for the amount a typical person would desire to spend on health care independent of their actual coverage. The agency takes additional steps in its behavioral model to account for how health

⁴² CBO selected the sources for its uninsured targets based, in part, on the Health Insurance Coverage Data Council Workgroup, *Data Council Workgroup Report on DHHS Health Insurance Coverage Estimates: An Evaluation of Existing Capacity, Content, Comparability and Alignment* (Agency for Healthcare Research and Quality, working paper, forthcoming).

spending changes in response to insurance coverage provisions, often referred to as moral hazard.

Both the average amount an HIU expects to spend on health care in the upcoming year and the HIU's uncertainty about what it will spend on health care affect its demand for health insurance. Because the distribution of potential spending, and not just the average expected cost, influences both HIUs' behavior and insurers' costs, CBO imputes a discrete probability distribution of health spending to each individual. That probability distribution is modeled as 16 health spending groups with annual expenditures on health care ranging from zero to more than \$25,000. Consequently, the result of CBO's expected health care spending imputation is a set of 16 predicted probabilities, which estimate each CPS respondent's chances of having annual health care costs within the range of each spending group, and 16 spending amounts, which reflect the dollar amount each respondent would expect to spend if his or her spending fell within the range for a given group. CBO chose to use a discrete probability distribution with 16 health spending groups to preserve computational simplicity while still capturing a wide range of possible spending amounts. Figure 8 compares the average predicted probability of having total health care spending within each spending category for people with private health insurance in HISIM2 with the percentage of MEPS-HC respondents with actual reported health care spending in each group.

Figure 8.
People Under Age 65 With Private Health Insurance, 2015
 Percent



HISIM2 is CBO’s health insurance simulation model.

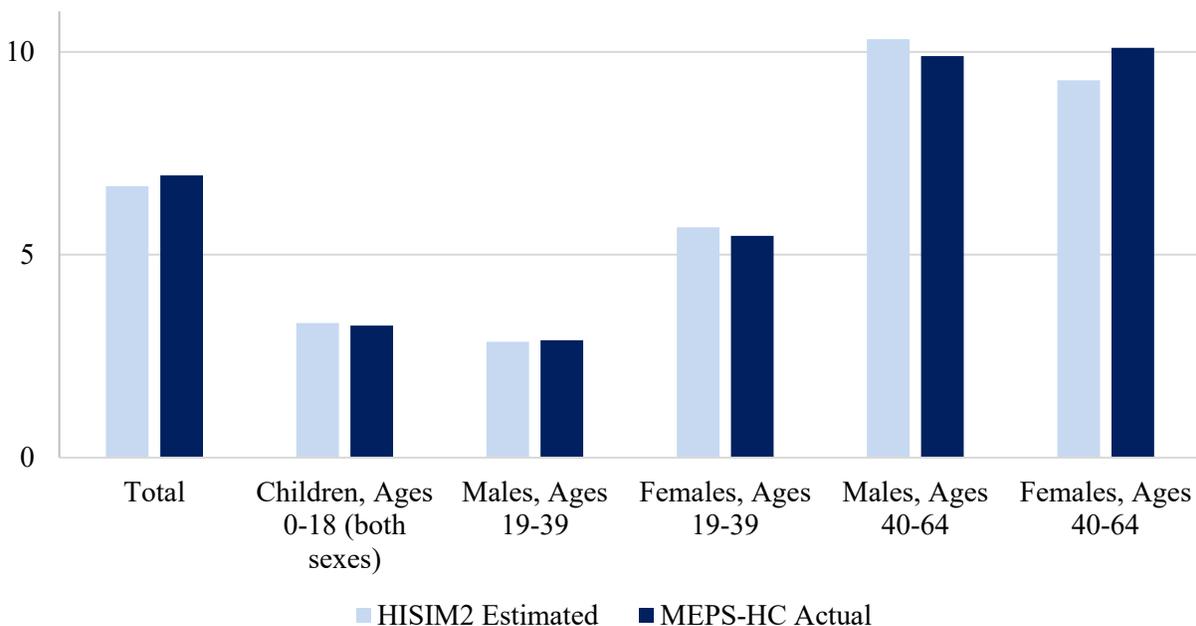
Small differences remain between HISIM2 and MEPS-HC data owing to differences in the covariates between the samples. The MEPS-HC data is adjusted to match the National Health Expenditure Accounts.

MEPS-HC = Medical Expenditure Panel Survey—Household Component.

Because the CPS does not include a measure of expected or actual health care spending, CBO uses an analysis of spending data in the MEPS-HC to impute expected health care spending. In CBO’s view, the MEPS-HC is the best source for individual-level data on health care spending by people with various sources of insurance coverage; however, it is known to undercount total health expenditures when compared with more reliable measures of those aggregates,

particularly the NHEA.⁴³ To correct for that undercount, CBO adjusts the MEPS-HC data used to impute health care spending by increasing the sample weights of underrepresented populations and by scaling reported spending amounts by payer and service categories to match NHEA aggregates. Figure 9 presents a comparison of average expected health care spending for the HISIM2 sample of people under age 65 with private health insurance coverage with average actual spending in the MEPS-HC by age and sex.

Figure 9.
Average Health Care Spending for People Under Age 65 With Private Health Insurance Coverage, 2015
 Thousands of 2015 Dollars



HISIM2 is CBO’s health insurance simulation model.

Small differences remain between HISIM2 and MEPS-HC data owing to differences in the covariates between the samples. The MEPS-HC data is adjusted to match the National Health Expenditure Accounts.

MEPS-HC = Medical Expenditure Panel Survey—Household Component.

⁴³ The undercounting of aggregate health spending in the MEPS-HC compared with the NHEA largely results from four causes: undercounting of Medicaid enrollees, undercounting of high-cost cases due to sampling difficulties and differential attrition, underreporting of expenditures by all people, and exclusion of institutionalized individuals. To correct for those causes of undercounting, CBO used the methods presented in Didem Bernard, Thomas M. Selden, and Yuruy O. Pylypchuk, *Aligning the Medical Expenditure Panel Survey to Aggregate U.S. Benchmarks, 2010*, Working Paper 15002 (Agency for Healthcare Research and Quality, January 2015), https://meps.ahrq.gov/data_files/publications/workingpapers/wp_15002.pdf.

After adjusting the MEPS-HC data, CBO estimates a person's total health care spending as a function of his or her characteristics. The person's characteristics included in the regressions comprise age, sex, race, self-reported health, disability, educational attainment, family size, income as a percentage of the federal poverty level, and health spending in the prior year.⁴⁴ Because the imputation is meant to represent a measure of expected health spending for a person in a typical private plan, CBO only includes MEPS-HC respondents who report 12 months of private health insurance coverage in the sample for those models and does not include any measures of insurance status in the regression models themselves. CBO estimates several models by age and sex groups because their health spending distributions differ substantially. To more accurately impute spending in the highest-spending group, CBO also estimates a regression model of spending conditional on spending more than \$25,000 in the sample year.

On the basis of the relationships estimated using the regression models and the MEPS-HC data, CBO generates predicted probabilities of having health care costs that fall within each of the 16 spending groups for each person in the CPS. Next, for each observation, the agency randomly imputes dollar spending amounts conditional on being in each health spending group. To adjust for geographic variation, those dollar amounts are multiplied by factors that vary by the CPS respondent's state of residence. The factors are calculated such that state-level variation in CBO's average imputed health spending measures matches the variation in state-level private health spending averages calculated using NHEA historical data.

CBO's health spending imputation assigns each person in the CPS a probability distribution of health spending for the next year. Those imputed spending distributions are used to create premiums, as described in the [Employment-Based Health Insurance Premiums](#) section, and to calculate out-of-pocket costs and their potential variance for each individual for each insurance option he or she may choose from during a simulation. The mean expected health spending for the HISIM2 population under age 65 is about \$6,700 in the base year (2015). That mean is not a measure of actual health care spending. Rather, it is a measure, based on the available data, of what people might expect to spend if they enrolled in a typical private health insurance plan for the entire year.

Synthetic Firms

Because most people under age 65 obtain health coverage through their or their family member's employer, simulating firms' decisions to offer coverage to some of their employees and their responses to various policy proposals is an important component of HISIM2. A firm's decisions are influenced by the characteristics of its workforce. However, CPS data do not provide any

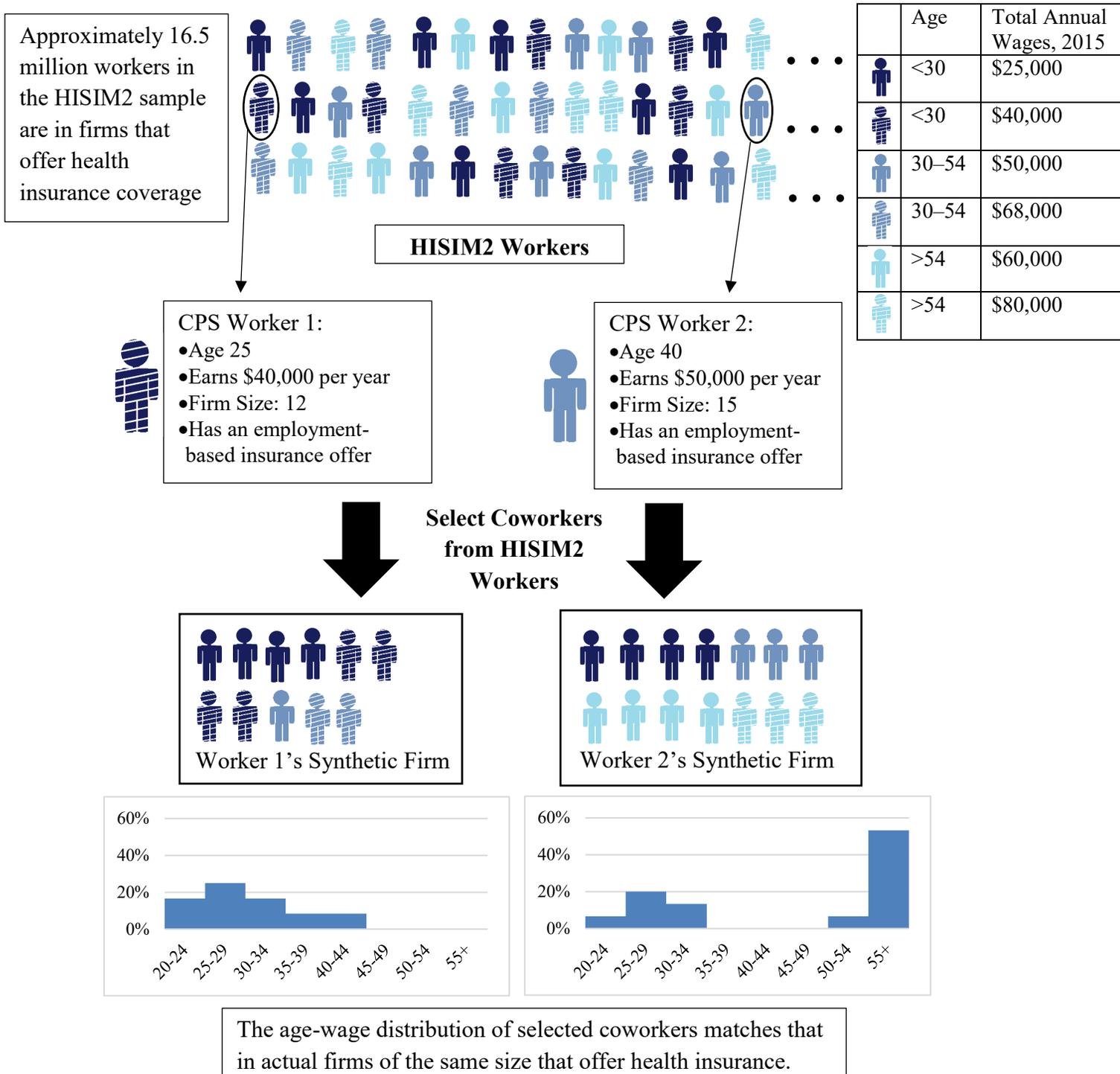
⁴⁴ Although the CPS includes most of the same measures used in the MEPS-HC analysis, it does not include a measure of health care spending in the prior year. Consequently, CBO assigns prior-year health care spending to each CPS respondent using a cold-deck imputation from the MEPS-HC that is based on all of the other variables used in the health spending regression models.

information on respondents' coworkers. Therefore, CBO builds a synthetic firm for each respondent consisting of an imputed set of coworkers that match a realistic distribution of coworker characteristics. That practice allows CBO to better model the availability and price of employment-based health insurance for individual workers.

To build synthetic firms, CBO assigns coworkers to most of the workers in the CPS data.⁴⁵ To do so, CBO and JCT conducted a joint analysis of how workforces vary from firm to firm using administrative tax data. That analysis first classifies firms by size (that is, by number of employees) and by whether they offer employment-based health insurance. Firms are then clustered into groups of firms with similar employee age and wage distributions. CBO then randomly assigns each CPS worker to a specific cluster of firms based on the probability that someone with their age and earnings works at firms like those in each cluster. Finally, CBO randomly assigns coworkers to each CPS worker according to the average employee age and earnings distribution in the cluster of firms to which they were assigned. Figure 10 illustrates the process CBO uses to build synthetic firms for two example workers.

⁴⁵ CBO does not build synthetic firms for workers under the age of 18 and most self-employed workers.

Figure 10.
Illustrative Examples of the Development of Synthetic Firms for Two Workers



HISIM2 is CBO's health insurance simulation model.

CPS = Current Population Survey.

CBO assigns individuals to a cluster of firms to mimic the real-world variation between firms.⁴⁶ For example, in the real world, young, low-earning workers can work at firms with very different distributions of coworkers from the average distribution of all young, low-earning employees. If CBO assigned coworkers to each young, low-earning worker on the basis of an average distribution across all firms, then all such workers would end up with similar firms. Instead, CBO randomly selects a firm cluster that is reasonable for each worker. Those clusters differ from one another in ways that mirror differences among firms in the real world. For example, one cluster might consist of firms with mostly older, higher-earning coworkers (for instance, a law firm), and another might consist of firms with mostly young, low-earning workers (for example, a food service job).

CBO also ensures variation in other measures that might affect the estimated impact of policies. The state or states in which a firm's employees live can affect employees' health insurance coverage options and costs as a result of variation in nongroup options, Medicaid eligibility requirements, and the overall cost of health care. Therefore, on the basis of an additional analysis of administrative data, about two-thirds of synthetic firms include coworkers sampled from a single state (to the greatest extent possible), while the remaining one-third includes coworkers sampled from the United States as a whole. As a result of that process, HISIM2's synthetic firms' decisions vary by region and for workers with similar ages and earnings profiles, reflecting the diversity among actual firms.

Employment-Based Health Insurance Premiums

The premiums charged for employment-based health insurance are another important factor in determining whether a worker's firm decides to offer health insurance to some of its' employees and whether that worker's HIU decides to enroll in that insurance. To approximate the variety of employment-based insurance options typically available, CBO generates premiums for six different employment-based plans per firm. Those six plans comprise three types of plans—HDHP, HMO, and PPO—for two types of coverage, single and family.⁴⁷ To ensure that those premiums reflect the expected health spending of the firm's employees and each plan's cost-sharing characteristics, CBO imputes health insurance premiums using a multistep process that is intended to be a simplified representation of how premiums are actually constructed by firms and group insurers.

Assigning Plans' Cost-Sharing Characteristics. First, for each synthetic firm, CBO assigns cost-sharing characteristics to each of the six plans. CBO defines four variables to capture variation in plan design in a simplified format that abstracts from which services are subject to

⁴⁶ Similar firms are clustered together to mask confidential, identifiable data.

⁴⁷ CBO constructs premiums for all three types of plans for every firm. Those premiums are then used to model whether the firm offers insurance to its workers and which plan types it offers.

the deductible. Those variables consist of a deductible, a pre-deductible coinsurance rate, a postdeductible coinsurance rate, and an out-of-pocket maximum. Firms are assigned the average deductible and out-of-pocket maximum for each plan type for that firm's size and state based on information from MEPS-IC data. The pre- and postdeductible coinsurance rates are imputed to match average actuarial values for each type of plan. The pre-deductible coinsurance rate is meant to approximate certain pre-deductible benefits (benefits that are not subject to the deductible), such as preventive care. The pre-deductible coinsurance rate is zero for benefits that are not subject to the deductible and greater than zero for benefits that are subject to the deductible. The postdeductible coinsurance rate is intended to serve as a simplified representation of how costs incurred between the deductible and the out-of-pocket maximum are split between the policyholder and insurer.

Calculating Insurers' Average Cost per Policy. Next, given the characteristics of plans offered by each firm, CBO calculates the insurers' expected costs for each plan for each of the firm's eligible workers. For single plans, CBO uses the plan's cost-sharing characteristics to calculate the out-of-pocket costs for each worker for each of the 16 health spending groups. The cost to an insurer for a given worker is that worker's health care spending minus the worker's out-of-pocket costs. The insurer's total expected cost for the given worker under a given plan is calculated by summing that particular worker's cost over all 16 health spending groups, weighted by his or her imputed probability of being in each group. In some cases, insurers and firms—especially firms with more than 100 employees—set premiums on the basis of actual health care spending for their policyholders in the prior year, a method known as experience rating. To approximate experience rating, CBO calculates the insurer's estimated cost for each worker under each plan using the worker's imputed prior-year health spending and the plan's cost-sharing characteristics. For each worker, CBO then calculates a blended average of the insurer's expected spending and the insurer's prior-year spending.⁴⁸

For family plans, CBO uses an analogous method, aggregating HIU spending into 16 HIU health spending categories and then calculating a blended average of expected and prior-year spending for each eligible worker's HIU. After calculating those blended average costs for each worker (or each worker's HIU for family plans), CBO calculates the mean firm-level cost to insurers for each of the six plans by averaging the worker-level expected costs over all eligible workers and weighting that average by each eligible worker's predicted probability of taking up employment-based coverage.⁴⁹ The result is the average health care cost to insurers for each of the six representative plans at each synthetic firm.

⁴⁸ The blending factor used in that calculation varies by firm size. It is meant to capture the full or partial insurance often used by small firms and the inherent risk spreading of large firms.

⁴⁹ Only eligible workers in HIUs with more than one member are used to calculate the mean cost to insurers for family plans.

Imputing Premium Load. The final step in calculating premiums is estimating the premium load. A premium load is a multiplicative factor used to account for costs other than health care costs, which can include insurer profit, risk-bearing costs, administrative costs, and cross-subsidization of retiree plans by active enrollees. To impute the premium load, CBO first assigns each synthetic firm an average load for that firm’s size and industry category. Next, to ensure sufficient variation in the load factor, and therefore premiums, CBO applies limited random variation to each firm’s assigned average.⁵⁰ For each firm, the average cost per policy for each insurer’s six representative health plans is then multiplied by the firm’s premium load to calculate six plan premiums.

Applying Small-Group Premium Rating Rules. After calculating a set of premiums based on the expected health costs of firms’ workers, CBO then applies the ACA’s small-group rating rules for firms that are subject to them.⁵¹ Under those rules, health insurance premiums can vary only by enrollees’ family size, age, tobacco use, and geographic location, and the amount premiums may vary by those characteristics is limited by a set of rating factors, which may differ by state.⁵² To determine the premiums for firms in the small-group market, CBO assigns the state rating factor for each employee (for single plans) and each employee’s HIU (for family plans). Those rating factors are used to calculate the average (single and family) rating factors for each small firm. Those firm-level rating factors are multiplied by the average state premium for small-group firms (calculated before applying rating rules using the method detailed above) to calculate the firm’s premiums for each plan.

Imputing Employees’ Premium Shares and Contributions to Health Savings and Reimbursement Accounts. After calculating premiums and applying small-group rating rules, CBO imputes the share of the premium paid by the employees and whether employees may pay for premiums on a pretax basis. The employees’ share of the premium for single plans is randomly imputed using MEPS-IC data on the distribution of employees’ premium shares by firm size, the percentage of the firm’s workers earning low wages, and the percentile of the firm’s premium for single plans.⁵³ CBO then uses additional MEPS-IC data on the relationship between the employer’s share of premiums for single plans and for family plans to randomly assign a premium share for family plans. Additional MEPS-IC data are used to randomly impute

⁵⁰ The amount of random variation applied to firms of a given size and industry depends on the fractions of firms in that group that self-insure and offer coverage to retirees. Through consultations with actuaries, CBO identified those two characteristics as driving much of the interfirm variation in premium load.

⁵¹ States determine the definition of small-group employers—typically as firms with either 50 or fewer employees or 100 or fewer employees.

⁵² For more information on the small-group rating rules, see Sec. 2701 of the Patient Protection and Affordable Care Act, Public Law 111-148 (codified at 42 U.S.C. § 300gg (2019)).

⁵³ The 2015 MEPS-IC defines low-wage workers as those earning less than \$11.50 per hour.

workers' eligibility to pay premiums with pretax income on the basis of each worker's imputed single-plan premium share and firm size.

Finally, CBO imputes whether and how much an employer would contribute to each worker's health savings account (HSA) or health reimbursement account (HRA) if that firm offered an HDHP. HRAs and HSAs represent costs to employers and benefits to workers, so those imputed values influence both employers' decisions and the utilities of those choices for employers and workers. The imputation of whether a firm contributes to one of the two accounts uses data from the EHBS and is random, based on the firm's size.⁵⁴ For those firms that do contribute, that contribution is randomly imputed to match data from the EHBS.

The premiums, employees' premium shares, and contributions to HRAs and HSAs are used to calculate the costs to firms of offering each of the three different types of plans and to model firms' decisions about whether to offer employment-based insurance and, if so, what plan type to offer their employees. Similarly, the measures, as well as plans' cost-sharing characteristics, are used to calculate the costs and benefits of plan choices for HIUs that are eligible for employment-based coverage.

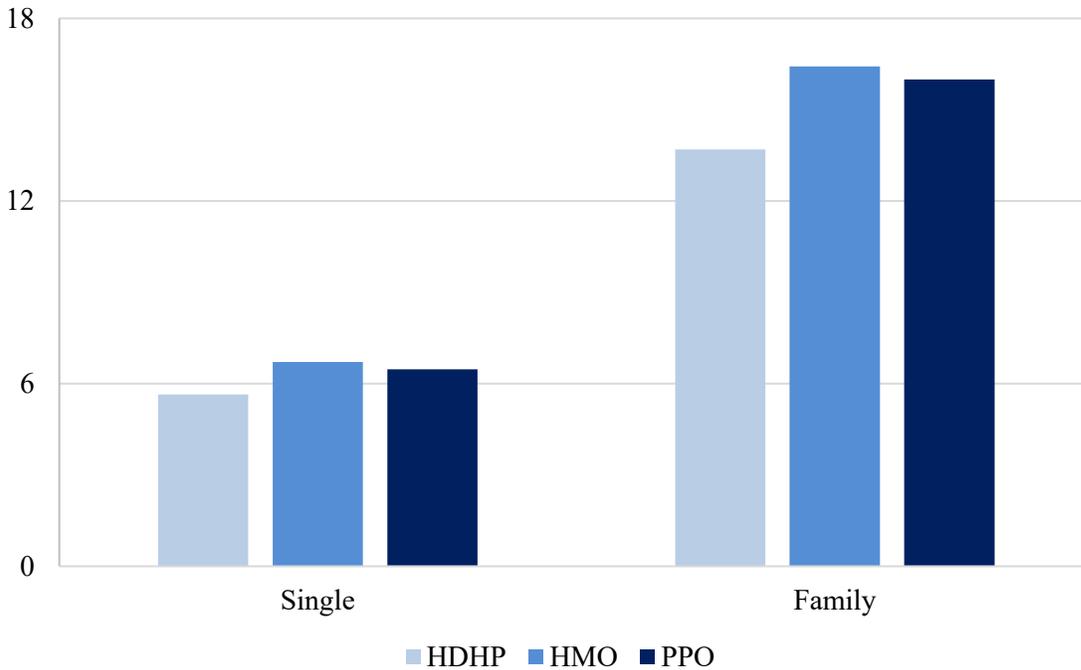
Using that method, CBO determined that the average premium for employment-based coverage in 2015 is about \$6,300 for single plans and about \$15,400 for family plans in HISIM2, compared with \$6,200 and \$15,600, respectively, in the MEPS-IC. In the same year, the average share of premiums covered by employers in HISIM2 is about 78 percent for single plans and 72 percent for family plans. Figure 11 displays the average HISIM2 premiums for single and family plans, by type of plan.

⁵⁴ HISIM2 does not distinguish between HSAs and HRAs. It incorporates the assumption that HIUs use funds from either account type to reduce any out-of-pocket costs up to the total contribution amount for an HRA or HSA.

Figure 11.

Average Premiums for Employment-Based Coverage, 2015

Thousands of 2015 Dollars



HDHPs are high-deductible insurance plans that allow the use of a tax-preferred health savings account to cover expenses not paid by the plans. HMOs are insurance plans in which services obtained outside a specified network of providers are often not covered. PPOs tend to offer wider provider networks, cover services from providers outside of their network, and typically limit costs through cost-sharing arrangements and a deductible.

HDHP = high-deductible health plan; HMO = health maintenance organization; PPO = preferred provider organization.

Nongroup Premiums

CBO uses comprehensive data on past and current premiums offered in the nongroup market to determine nongroup premium offers. The data are drawn from the Robert Wood Johnson Foundation’s HIX Compare database for 2015 through 2019. CBO determines the average premium for each metal tier in each state. The premiums are determined on a state-by-state basis because there are large differences between states in their nongroup premiums. CBO also assigns plan characteristics (such as deductibles, copayments, and coinsurance amounts) using the HIX Compare database. Those are likewise specified by state and by metal tier.

Projecting Data to the End of the Baseline Budget Projection Period

After adjusting the base-year data for input into HISIM2 (as described in the previous section), CBO develops the input data for each year through the end of the period for the baseline budget

projections. CBO takes a number of steps to project population characteristics from the base-year data to the end of the 10-year baseline budget period. (The current base year is 2015, and CBO's baseline projections extend through 2029, the final year in this year's budget period.)

CBO employs two main approaches to project population characteristics of the base-year data: applying nominal growth factors and inflation to variables measured in dollars and adjusting sample weights. The agency first inflates income, health spending, and employers' contributions to premiums using different rates that are meant to capture the unique growth patterns of those variables. Then, CBO uses an optimization routine to adjust the sample weights for people in the CPS sample over the projection period to simultaneously match projections of population characteristics, including population growth and changes in patterns of employment. Table A-3 in the appendix provides an overview of the major adjustments CBO makes to project variables through the budget period and the reasons behind them.

Income Growth

As mentioned earlier, income in HISIM2 is grouped into measures of wage, self-employment and property, and other income, which CBO expects to grow at different rates. In addition, CBO expects the growth of individual wage, self-employment, and property income to vary by filing status and income level. Therefore, CBO calculates different income growth rates before adjusting the weights for people in the sample. Overall income growth is consistent with CBO's macroeconomic forecast of economic growth and projections of employment.

CBO expects wage growth to vary by age, tax-filing status, and wage levels. To calculate the nominal income growth factors, CBO divides TFUs into groups, first based on whether the unit files an income tax return. Tax filers are grouped into eight categories based on age (under and over age 65) and wages (1st–90th percentiles, 91st–95th percentiles, 96th–99th percentiles, and the top 1 percent). Nonfilers are grouped into two categories by age (under and over 65). Those groupings were informed by the historical differences in annual nominal income growth for workers with higher wages and by the availability of more complete information on income for tax filers than nonfilers. Income is categorized into three types (wage, self-employment and property, and other income), allowing for each of the three types to grow at different rates.

The growth factors are calculated for each type of income for each of the eight groups of filers, taking into account that wages grow faster over time for people with higher wages. CBO calculates the self-employment and property income and other income growth factors based on expected per capita growth in CBO's baseline. CBO's income growth projections for each income source are allowed to differ, consistent with historical experience, but the factors are normalized so that their average growth by broad category equals the average in the

macroeconomic forecast.⁵⁵ For nonfilers, CBO calculates the wage growth factors based on the per capita growth using data from the SOI matched to the CPS. The growth factors for property income and other income are smaller for nonfilers than for filers because the composition of property and other income is more heavily weighted toward slower growing income sources for nonfilers.

Health Spending and Cost Sharing Over Time

CBO also calculates a premium growth factor for private insurance because health care costs are generally expected to grow faster than most types of income and the economy.⁵⁶ The premium growth factor is estimated as a function of growth in personal disposable income and of lagged premium growth using historical health spending data from the NHEA. CBO then applies that growth factor to base-year measures related to health care spending, including expected health spending for all health insurance units, employment-based insurance premiums, employment-based and nongroup insurance plan out-of-pocket maximums, deductibles for nongroup insurance plans, and HRA and HSA contributions. CBO inflates deductibles for employment-based insurance plans at a slightly faster rate than premiums because deductibles have historically grown at faster rates than premiums.

Employers' contributions to premiums have their own historical pattern: CBO projects changes in employers' contributions to premiums using historical MEPS-IC data starting in 2004, which show that employers' contributions are declining at a roughly constant rate. The agency uses that information to project the share of premiums contributed by employers for each year in the projection period.

Unlike employment-based premiums, nongroup premiums are determined during the simulation stage of HISIM2. As described in the [Nongroup Premiums](#) section, actual average state-level premiums by metal tier for marketplace and nonmarketplace plans are used through 2019. In subsequent years, the premium is determined by the expected health spending of individuals who enroll in the nongroup market and the health care premium growth factor described above.

⁵⁵ See Congressional Budget Office, "An Overview of CBO's Microsimulation Tax Model" (June 2018), www.cbo.gov/publication/54096.

⁵⁶ See Jessica Banthin, Deputy Assistant Director, Health, Retirement, and Long-Term Analysis Division, Congressional Budget Office, "Projections of Federal Spending on Major Health Care Programs" (presentation for the Alliance for Health Policy Summit on Health Care Costs in America, Washington, D.C., May 24, 2018), www.cbo.gov/publication/53887.

Adjustments to Person Weights

CBO adjusts the sample weights of CPS respondents to match its projections of growth in the U.S. population. CBO changes how much sample weight an individual in the survey represents in each year of the projection period to ensure the appropriate distributions of certain characteristics, including age, race, sex, employment status, earnings, and immigration status. The target values are based on data for recent years and CBO's most recent 10-year budget and economic projections.⁵⁷

Additional Adjustments

CBO makes additional adjustments to the input data for future years to incorporate certain changes in tax law and expectations regarding states' expansions of Medicaid.

States' Expansions of Eligibility for Medicaid. Since the base year of 2015, some states have expanded eligibility for Medicaid under the ACA. CBO expects that if current federal laws remained in place, even more states would expand eligibility for the program and more people would enroll in the program in states that have already done so. Most of the increase in enrollment during that period would stem from the additional states' expansions of eligibility for the program, CBO estimates. Therefore, CBO creates income eligibility thresholds for Medicaid and CHIP for each year during the projection period to reflect states' expansions that have occurred since 2015 and the expectation of expansions in eligibility in additional states in future years. Using those thresholds, CBO calculates eligibility for each person in a given HIU and creates Medicaid and CHIP eligibility variables for each individual for each year of the projection period.

Changes to Marginal Tax Rates. As mentioned earlier, CBO uses the National Bureau of Economic Research's TAXSIM model to produce estimates of the AGI and MTR for each tax-filing unit in the base year of input data. CBO uses TAXSIM to produce those estimates for all years in the projection period. Currently, however, TAXSIM incorporates tax law only through 2023. For years beyond 2023, CBO's estimates of AGIs and MTRs incorporate the assumption that 2023 tax law remains in effect. For 2026 and later years, however, CBO corrects federal MTRs produced by TAXSIM to adjust for the expiration of provisions in the 2017 tax act (Public Law 115-97).⁵⁸ Those adjustments to tax rates in later years are based on CBO's baseline projections and JCT's predictions of the distributional effects of tax law. MTRs in 2026 and later

⁵⁷ See Congressional Budget Office, *The 2018 Long-Term Budget Outlook* (June 2018), www.cbo.gov/publication/53919. For more information about CBO's budget and economic projections, see Congressional Budget Office "Outlook for the Budget and the Economy" (accessed March 2019), www.cbo.gov/topics/budget/outlook-budget-and-economy.

⁵⁸ For more information, see Congressional Budget Office, "How the 2017 Tax Act Affects CBO's Projections," *CBO Blog* (April 20, 2018), www.cbo.gov/publication/53787.

are expected to be higher because provisions of the 2017 tax act that lower individual income tax rates will expire on December 31, 2025.

Final Data Set

This report describes how CBO developed the input data for the agency’s new version of its health insurance simulation model, HISIM2, which underlies the agency’s spring 2019 baseline projections. The final data set is used to model health insurance decisions made by HIUs and employers. The process described in this paper is updated annually to improve the model’s accuracy and to incorporate new data. CBO has also published information about the assumptions, equations, and processes underlying its simulations of health insurance decisions.⁵⁹

Once HISIM2 simulates health insurance decisions and produces output in the form of coverage distributions, CBO takes several additional steps to produce the baseline projections. HISIM2’s coverage projections are adjusted slightly using CBO’s Medicaid enrollment model and other models that analyze aspects of current law that are simplified in HISIM2. Spending for Medicaid, CHIP, and the Basic Health Plan is estimated using models specific to those programs. The net costs of federal subsidies for work-related coverage and coverage through the nongroup market, as well as taxes and penalties related to that coverage, are estimated using the JCT’s tax models. All of those models are used together to develop CBO’s baseline and cost estimates.

⁵⁹ See Congressional Budget Office, “HISIM2—The Health Insurance Simulation Model Used in Preparing CBO’s Spring 2019 Baseline” (April 2019), www.cbo.gov/publication/55097.

Appendix

For ease of reference, the appendix includes three tables that summarize the discussion in the main body of the paper. Table A-1 provides information on the Congressional Budget Office’s key edits and additions to the Current Population Survey data. Table A-2 lists other sources of data used by CBO to determine health insurance coverage targets. Table A-3 summarizes how CBO projects key groups of variables through the budget projection period.

Table A-1.
Overview of Edits and Additions to the CPS Base Data

Variable	Add or Edit	Primary Reason for Modifying Data	Relation to CPS	Sources of Data and Methods Other Than the CPS
HIU	Add	To simulate health insurance decisionmaking	Constructed using CPS relationship variables	NA
TFU	Add	To determine subsidy eligibility and tax liability	Constructed using CPS relationship and income variables ^a	NA
Age	Edit	To adjust the age reported at time of survey to match health insurance coverage reported for the prior year	Based on CPS age variable	NA
Lawful Presence of Noncitizens	Add	To define health insurance options and subsidy eligibility	Modeled using CPS variables on occupation, government assistance, length of time since arriving in the United States, and family relationships	Methods drawn from Borjas (2017) ^b
Firm Size	Edit	To better match CPS data to data from a more accurate source ^c	Based on CPS firm size variable	<i>MEPS-IC data</i>
Employment-Based Health Insurance Offer Status	Edit	To better match CPS data to data from a more accurate source ^c	Based on CPS offer status variable	<i>MEPS-HC and MEPS-IC data</i>
Employment-Based Health Insurance Eligibility	Edit	To account for workers at firms that do not offer health insurance and to better match CPS data to data from a more accurate source ^c	Based on CPS offer status variable	Imputed offer status and <i>MEPS-HC</i>
Wage Income	Edit	To better match CPS data to data from a more accurate source ^d	Based on CPS wage variable	<i>IRS administrative tax data and data from information returns</i>

Variable	Add or Edit	Primary Reason for Modifying Data	Relation to CPS	Sources of Data and Methods Other Than the CPS
Self-Employment Income	Edit	To better match CPS data to data from a more accurate source ^d	Based on CPS self-employment income variable	<i>Statistics of Income</i>
Capital Gains Income	Add	To better match CPS data to data from a more accurate source ^d	NA	<i>Statistics of Income</i>
Other Income	Edit	To better match CPS data to data from a more accurate source ^d	Based on CPS income variables	<i>IRS administrative tax data and data from information returns</i>
Medicaid/CHIP Eligibility	Add	To determine health insurance coverage options	Constructed using CPS income, age, and family relationship variables	Data from a survey conducted by the Kaiser Family Foundation with the Georgetown University Center for Children and Families ^e
Public Insurance Coverage^f	Edit	To better match CPS data to more accurate sources of information on average annual enrollment in health coverage of the population	Based on CPS coverage variables	MEPS-HC, <i>CMS administrative data</i> , NHIS
Health Spending	Add	To model HIUs' choice of coverage	Imputed using CPS age, sex, race, health status, disability, education, and income	MEPS-HC, NHEA
Coworker and Firm Characteristics	Add	To create synthetic firms and model employers' health insurance decisions	Uses CPS respondents as coworkers	<i>IRS administrative tax data on employers and data from information returns for employees</i>
Employment-Based Health Insurance Premiums	Add	To estimate the cost of health insurance plan options for employers and HIUs	NA	<i>MEPS-IC</i>
Employees' Premium Shares and HRA/HSA Contributions	Add	To help determine the cost to employers and employees of various plans	NA	<i>MEPS-IC</i> , EHBS
Nongroup Premiums and Plan Characteristics	Add	To define the options available to people through the nongroup market	NA	Robert Wood Johnson Foundation HIX Compare database

Data from sources in italics are confidential.

ACA = Affordable Care Act; CHIP = Children's Health Insurance Program; CMS = Centers for Medicare & Medicaid Services; CPS = Current Population Survey; EHBS=Employer Health Benefits Survey; HIU = health insurance unit; HRA = health reimbursement account; HSA = health savings account; MEPS-HC = Medical Expenditure Panel Survey—Household Component; MEPS-IC = Medical Expenditure Panel Survey—Insurance Component; MSIS = Medicaid Statistical Information System; NA = not applicable; NHEA = National Health Expenditure Accounts; NHIS = National Health Interview Survey; TFU = tax-filing unit.

^a CBO does not use the CPS TFU variable; rather, it creates a new TFU variable using the household relationships and tax-filing status variables from the CPS. CBO also uses Statistics of Income data to adjust TFUs to ensure they match that more accurate data source.

^b For more information, see George J. Borjas, “The Labor Supply of Undocumented Immigrants,” *Labour Economics*, vol. 46 (June 2017) pp. 1–13, <https://doi.org/10.1016/j.labeco.2017.02.004>.

^c The MEPS-IC is an employer-based survey, meaning that employers provide the responses. In CBO’s assessment, employers are more likely than employees to accurately report firm size and information about insurance offers.

^d CBO considers administrative tax data—composed of Statistics of Income data and data from information returns (such as Form W-2s)—to be of better quality for estimating income than CPS income variables, which rely on self-reporting.

^e See Kaiser Family Foundation, “Annual Updates on Eligibility Rules, Enrollment and Renewal Procedures, and Cost-Sharing Practices in Medicaid and CHIP” (various years), www.kff.org/medicaid/report/annual-updates-on-eligibility-rules-enrollment-and/.

^f Other public insurance coverage includes Medicare coverage, Medicare and Medicaid dual enrollment, and coverage through military and veterans’ health care programs.

Table A-2.
Sources of Health Insurance Coverage Targets

Health Insurance Coverage Category		Can the Choice of Coverage Change Through the Projection Period?	Data Source for 2015 Targets
Employment-Based		Yes ^a	MEPS-IC, MEPS-HC, OPM
Medicaid	Disabled	No	Form CMS-64, California Department of Health Care Services data, MSIS, Medicaid Analytic eXtract data
	Nondisabled	Yes ^b	Form CMS-64, California Department of Health Care Services data, MSIS, Medicaid Analytic eXtract data
CHIP		Yes	Statistics Enrollment Data System
Nongroup Purchased Through the Marketplaces		Yes ^c	CMS Medical Loss Ratio data, CMS quarterly effectuated enrollment reports, <i>effectuated marketplace enrollment data from HealthCare.gov</i> , MEPS-HC, covered California data, New York State Department of Health data
Nongroup Purchased Outside the Marketplaces		Yes ^c	CMS Medical Loss Ratio data, <i>effectuated marketplace enrollment data from HealthCare.gov</i> , MEPS-HC
Medicare	Over age 65 with employment-based insurance offer	Yes	MEPS-HC
	Over age 65 without employment-based insurance offer	No	Not targeted: Enrollment counts are based on adjusted CPS
	Under age 65 ^d	No	Administrative data from CMS and the <i>Social Security Administration</i>
Military and Veterans' Health Care Programs ^e		No	Not targeted: Enrollment counts are based on adjusted CPS
Medicaid and Medicare Dual Enrollment		No	Not targeted: Enrollment counts are based on adjusted CPS
Other ^f		No	Not targeted: Enrollment counts are based on adjusted CPS
Uninsured		Yes	<i>MEPS-HC</i> ^g and NHIS

Data from sources in italics are confidential.

CHIP = Children's Health Insurance Program; CMS = Centers for Medicare & Medicaid Services; MEPS-HC = Medical Expenditure Panel Survey—Household Component; MEPS-IC = Medical Expenditure Panel Survey—Insurance Component; MSIS = Medicaid Statistical Information System; NHIS = National Health Interview Survey; OPM = Office of Personnel Management; VA = Department of Veterans Affairs.

^a Most people with employment-based health insurance are able to choose their coverage throughout the projection period. Exceptions include people with coverage under the Consolidated Omnibus Budget Reconciliation Act, nonelderly retirees, people with employment-based coverage under a policyholder that lives outside their home, and people who report both employment-based coverage and public insurance coverage.

^b Most people who qualify for Medicaid for reasons other than a disability are able to choose their coverage throughout the projection period. Exceptions include foster children with Medicaid and adults with both employment-based coverage and Medicaid.

^c Most people with nongroup coverage are able to choose their coverage throughout the projection period. However, if a child with nongroup coverage is in a household in which no adults can choose their coverage throughout the projection period, then that child is assigned nongroup coverage throughout the period.

^d Most Medicare-eligible people under age 65 qualify for Medicare because they participate in the Social Security Disability Insurance program.

^e People covered under military and veterans' health care programs include people with TRICARE coverage, military veterans who report using the VA health system, and dependents of seriously disabled veterans who are enrolled in the VA's Civilian Health and Medical Program. Because the Current Population Survey includes few active-duty service members, most of those with TRICARE coverage in the sample are dependents of active-duty members, retired service members and their dependents, or reservists.

^f Includes people with other kinds of insurance, such as student health plans, coverage provided by the Indian Health Service, and coverage from foreign sources.

^g The MEPS-HC data used to measure the number of uninsured people include both publicly available variables and a restricted-access measure of income as a percentage of the federal poverty level.

Table A-3.
Projecting Variables Through the End of the Baseline Period

Variables	Process for Projecting	Reason
Income	Apply growth factors for wage, self-employment and property, and other income and adjust respondents' sample weights	Different types of income are expected to grow at different rates.
Health Spending ^a	Forecast premium growth using a regression model combining NHEA data on historical health spending and population data	Historically, health spending has grown faster than overall inflation.
Health Insurance Cost Sharing for Employment-Based Insurance	Use historical trends reflected in MEPS-IC data to project employers' shares of premiums and deductibles	Historically, cost sharing has shifted slowly over time from employers to employees, and deductibles for employment-based insurance have grown slightly faster than premiums.
Demographic Characteristics ^b	Adjust respondents' sample weights using CBO's demographic baseline projections	The size and demographic characteristics of the population have changed over time and are projected to continue changing in CBO's baseline.
Medicaid Eligibility	Incorporate the assumption that states will continue to expand eligibility for Medicaid under the ACA ^c	Historically, states have adopted optional Medicaid expansions offered by the federal government over time. CBO expects states to respond similarly in this case.
Marginal Tax Rates	Use TAXSIM to apply past tax law for the 2015–2017 period and current tax law for the 2018–2029 period, then adjust MTRs for 2026 through 2029 because of the expiration of some provisions of the Tax Act of 2017	The law includes changes to tax rates over time.

ACA = Affordable Care Act; MEPS-IC = Medical Expenditure Panel Survey—Insurance Component; MTR = marginal tax rate; NHEA = National Health Expenditure Accounts; TAXSIM = National Bureau of Economic Research's TAXSIM model.

^a Health spending includes expected health spending, employment-based insurance and nongroup premiums, employment-based and nongroup insurance plan out-of-pocket maximums, deductibles for nongroup insurance plans, and contributions to health reimbursement accounts and health savings accounts.

^b Demographic characteristics include age, sex, race, employment, earnings, and immigration status.

^c Under the ACA, states can expand eligibility for Medicaid to adults under age 65 whose income is no more than 138 percent of the federal poverty level. Generally, most people who became eligible under the law are low-income adults without children.