

**Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.**

**How much does health insurance cost? Comparison of premiums
in administrative and survey data**

Jeff Larrimore and David Splinter

2018-030

Please cite this paper as:

Larrimore, Jeff, and David Splinter (2018). "How much does health insurance cost? Comparison of premiums in administrative and survey data," Finance and Economics Discussion Series 2018-030. Washington: Board of Governors of the Federal Reserve System, <https://doi.org/10.17016/FEDS.2018.030>.

NOTE: Staff working papers in the Finance and Economics Discussion Series (FEDS) are preliminary materials circulated to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors. References in publications to the Finance and Economics Discussion Series (other than acknowledgement) should be cleared with the author(s) to protect the tentative character of these papers.

How much does health insurance cost?

Comparison of premiums in administrative and survey data

Jeff Larrimore
Federal Reserve Board

David Splinter
Joint Committee on Taxation

April 26, 2018

Abstract

Using newly available administrative data from the Internal Revenue Service, this paper studies the distribution of employer-sponsored health insurance premiums. Previous estimates, in contrast, were almost exclusively from household surveys. After correcting for coverage limitations of the IRS data, we find that average premiums for employer-sponsored plans are roughly \$1000 higher in IRS records than in the Current Population Survey. The downward bias in the CPS is largely driven by underestimating of premiums among married workers and topcoding of high premiums.

The results and opinions expressed in this paper reflect the views of the authors and should not be attributed to the Federal Reserve Board. This paper embodies work undertaken for the staff of the Joint Committee on Taxation, but as members of both parties and both houses of Congress comprise the Joint Committee on Taxation, this work should not be construed to represent the position of any member of the Committee. We thank Priyanka Anand, Tom Barthold, Victoria Bryant, Jim Cilke, Mark Klee, Tricia McDermott, Ellen Merry, Kevin Pierce, and Claudia Sahm for helpful comments on earlier drafts of this paper.

Slow wage growth in the United States is a substantial concern, but existing statistics may not fully capture non-cash fringe benefits – especially employer-sponsored health insurance.¹ Until recently, no comprehensive source of administrative data on the value of workers’ health benefits was available. Instead, researchers had to rely on aggregate statistics or household surveys, such as the Medical Expenditure Panel Survey and the Census Bureau’s Current Population Survey. These survey-based data on health insurance premiums have been used widely by researchers seeking to understand inequality trends and income growth both at the median and at the upper and lower tails of the distribution (Burtless and Svaton 2010; Burkhauser, Larrimore, and Simon 2012; Congressional Budget Office 2018, and Piketty, Saez, and Zucman 2018).

Under the Affordable Care Act, employers with at least 250 employees must now report health insurance premiums to the IRS on Form W-2. We use these new administrative data to measure the distribution of premiums, and compare these estimates to those from the CPS. We find that the average cost of employer health plans in the two sources is similar for single workers. However, for married workers, both the mean and median cost is higher in the administrative data. We also document the extent to which topcoding of insurance amounts in the CPS suppresses the upper-tail of the health insurance distribution. As a result, our findings suggest that the CPS data understates the average level of these benefits.

Data

This paper uses W-2 records from IRS population data for 2015. Starting in 2011, employers have reported health insurance premiums on Form W-2 (in box DD), with increased reporting in subsequent years. This amount includes the combined premiums paid by both the employee and the

¹ Baicker and Chandra (2006) suggest that a 10 percent premium increase results in a 2.3 percent wage decrease. Kolstad and Kowalski (2016) observe that the compensating differential for jobs with health insurance is close to the full average cost of the insurance to the employer.

employer.² Although only firms with at least 250 employees were generally required to report health insurance premiums in 2015, many smaller firms also reported this information to the IRS. For our analysis, we assume that the distribution of premiums for a given firm size is similar for the reporting and non-reporting firms, and we adjust for the underreporting among small firms in the IRS data based on coverage counts by firm size in the CPS, as discussed further below. Other researchers have used these administrative health insurance data to study the effect of insurance coverage on labor market and family formation decisions among young adults (Heim, Lurie, and Simon, 2017, Forthcoming), but our work is the first to use the administrative data to estimate the full distribution of premiums.³

We compare the distribution of premiums in the 2015 IRS data to those from the 2016 March Current Population Survey (CPS), which captures health insurance for the previous calendar year.⁴ The CPS determines health insurance premiums separately for employee and employer contributions. Employee contributions are the self-reported amount paid by respondents for a health insurance policy in their own name. The Census Bureau imputes the employer contributions based on characteristics of the firm and the worker (see Janicki, O'Hara, and Zawacki, 2013 for details on the imputation approach). Total premiums are the combination of the self-reported employee contributions and imputed employer contributions.

² In addition to traditional health insurance premiums, it also contains certain health flexible spending account contributions paid by the employer. It does not, however, contain the more common flexible spending account contributions paid by employees. Flexible spending plans are excluded in the CPS data, irrespective of who makes the contributions. Other related expenditures excluded from both datasets include health reimbursement account and health savings account contributions. Dental and vision insurance coverage are generally excluded as well, although reporting of these and multi-employer union plans is optional in the W-2 data. Premium amounts for self-insured employers should equal costs distributed among employees using an actuarial approach. For details on insurance that is included and excluded in the W-2 data, see Internal Revenue Service (2018). The reported amount, however, is likely not subject to the same level of data integrity standards as wages and therefore should not be considered as reliable as wage compensation reported on Form W-2.

³ IRS Statistics of Income presents premium values for select percentiles of the premium distribution, see www.irs.gov/statistics/soi-tax-stats-individual-information-return-form-w2-statistics

⁴ To compare U.S. residents in both the IRS and CPS data, Form W-2 data with addresses in foreign countries or in U.S. territories are removed.

In both datasets, health insurance premiums are measured at the individual level, and reflect the total payments made by the policyholder and their employer for all individuals insured through the policy. Premium payments for family plans will cover multiple individuals, although the IRS data do not specify the type of plan or how many people are covered. However, this issue can be partially addressed by separately analyzing premiums based on the marital status and parental status of the policyholder.⁵

Table 1 shows the number of individuals with employer-sponsored insurance in the two datasets, separated by firm size. Unsurprisingly, given that small firms are not required to report health insurance premiums on Form W-2 to the IRS, more workers report health insurance through an employer in the CPS data than appear in the IRS records. The share of workers with health insurance converges in the two datasets as the firm size increases – although it is always higher in the CPS than in the IRS data, which likely results from some firms of all sizes not yet reporting this information to the IRS. This would not impact the distribution of premiums if employers of all sizes provided insurance with similar premiums. However, we also observe that the average premium rises with firm size, which is consistent with the higher wages seen among larger firms.⁶ In subsequent results, we therefore reweight the IRS data – increasing the weight on small firms that are underrepresented due to the reporting requirements – so that the counts of individuals with insurance by firm size match those from the CPS data.^{7, 8}

⁵ In the CPS, 80 percent of married policyholders with children and 57 percent without children have multiple people on their employer-sponsored health insurance plan. Among single adults, 67 percent with children and 5 percent without children have multiple people on their plan, which can include siblings or other dependents.

⁶ In the IRS data, the average wage is \$42,300 for firms with fewer than 100 workers, whereas it is \$50,700 for firms with over 500 workers.

⁷ Firm size is measured in the CPS at the establishment level and in the IRS data by using employer identification numbers (EINs) matched to Form 941. This form is an employer-level quarterly tax form that lists the number of employees. The use of multiple EINs among large firms can result in missing links between individual level Form W-2 and business level Form 941. Workers for whom no firm size is available in the IRS data are therefore included with firms with over 1,000 workers when reweighting.

⁸ Further justifying the reweighting, we compare aggregate values of employer-sponsored insurance premiums to National Income and Product Account (NIPA) estimates. The unadjusted IRS data observes \$572 billion of

Table 1. Employer-sponsored health insurance coverage and premiums by firm size

| | IRS | | | | CPS | | | |
|------------------------|-------------------------|----------------------|-------------------|---------------------|-------------------------|----------------------|-------------------|---------------------|
| | Policy-holders (thous.) | Share of workers (%) | Mean Premium (\$) | Median Premium (\$) | Policy-holders (thous.) | Share of workers (%) | Mean Premium (\$) | Median Premium (\$) |
| All firms | 60,041 | 40 | 9,415 | 7,600 | 78,744 | 53 | 7,891 | 7,595 |
| 1-9 workers | 738 | 6 | 7,423 | 5,600 | 5,254 | 24 | 7,303 | 7,041 |
| 10-49 workers | 1,902 | 10 | 6,925 | 5,200 | 8,612 | 37 | 7,256 | 6,766 |
| 50-99 workers | 1,662 | 19 | 7,212 | 5,600 | 5,716 | 53 | 7,450 | 6,963 |
| 100-499 workers | 8,939 | 41 | 8,926 | 7,000 | 11,589 | 61 | 7,705 | 7,479 |
| 500-999 workers | 5,458 | 56 | 9,861 | 7,800 | 5,224 | 63 | 8,001 | 7,744 |
| 1000+ workers | 36,107 | 57 | 9,647 | 7,800 | 42,349 | 65 | 8,189 | 7,939 |
| No firm size available | 5,235 | 33 | 10,019 | 8,200 | | | | |

Note: All results are for calendar year 2015 aggregated to the individual level, and include both the employer and employee contributions for health insurance. Individuals with less than \$2,000 in annual wages are excluded. To preserve confidentiality, premiums in W-2 data are rounded to the nearest \$200. Premiums are topcoded in the W-2 data at \$40,000 to limit the impact of inaccurate entries. This threshold binds for 0.02 percent of workers. Insurance coverage and values are based on all jobs, although classifications of firm size are based on the largest firm at which the individual worked.

Source: Authors' calculations using Census and IRS data.

Distribution of Health Insurance Premiums

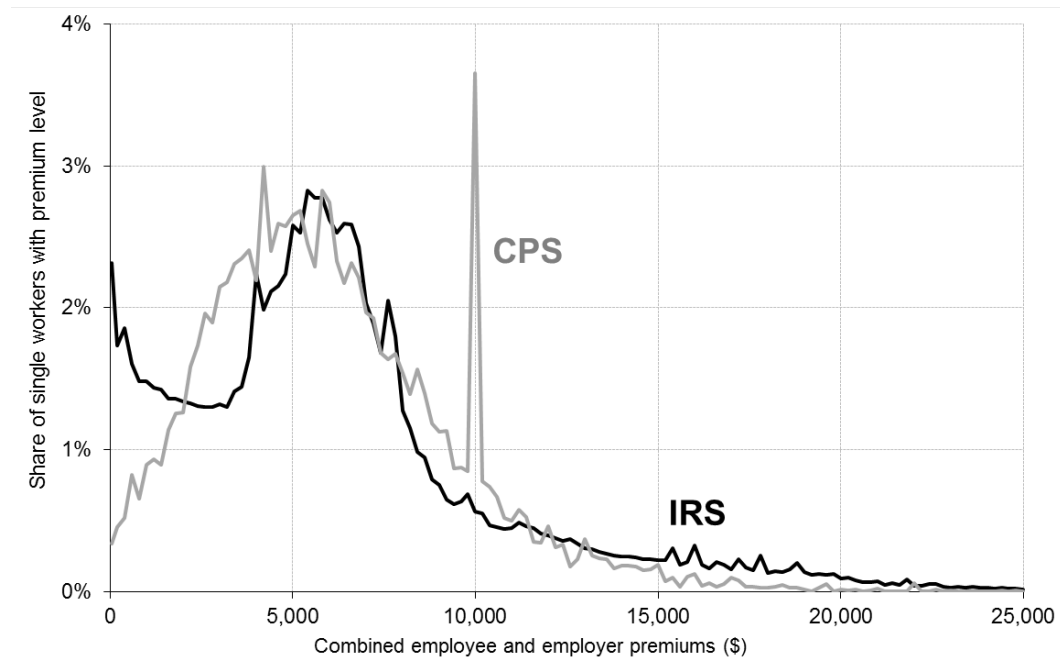
The distribution of premium for single workers from the two data sources, shown in Figures 1 and 2, are fairly similar with two exceptions. First, the IRS data finds more single workers with premiums under \$2,000 and fewer with premiums between \$2,000 and \$4,000. This likely reflects part-year workers who only worked and had insurance for a small portion of the year, which may be more accurately captured in the IRS data rather than through retrospective survey questions or imputations.⁹

premiums, which is below both the \$621 billion of premiums reported in the CPS and the \$672 billion in the NIPA estimates. After reweighting to reflect the undercount of insurance in the IRS data, we observe \$705 billion of premiums in the IRS data.

⁹ Additionally, the Affordable Care Act initially required that firms with at least 200 full-time employees must automatically enroll new full-time employees in their health insurance plan. Although this provision was never implemented (and the change was repealed as part of the Bipartisan Budget Act of 2015), workers who were automatically enrolled, but opted out within the allowable window, may also have a small amount of premiums reported. For details, see www.shrm.org/resourcesandtools/hr-topics/benefits/pages/healthplan-auto-enroll.aspx

A second difference occurs at the \$10,000 topcoding threshold in the CPS data. This threshold is only weakly binding, since it applies to the employer share of premiums and not the portion paid by the worker. Nevertheless, nearly 4 percent of single workers in the CPS have premiums at this threshold. As a result, the IRS data shows more single workers with premiums over \$10,000 than does the CPS.

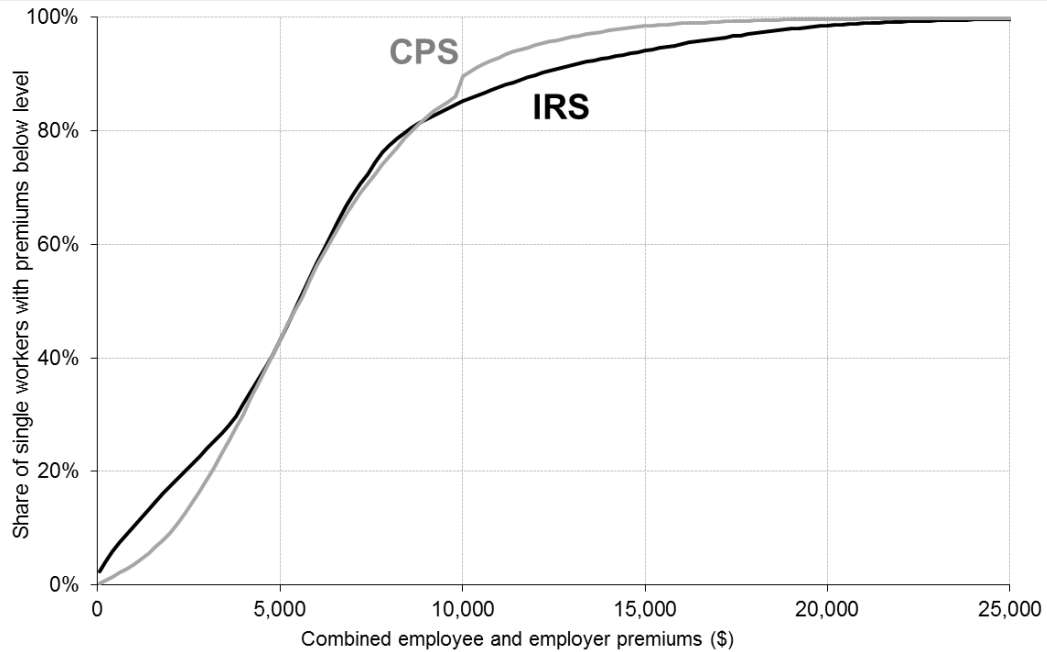
Figure 1. Distribution of premiums for employer-sponsored insurance among single workers



Note: See Table 1. Reflecting that not all firms are required to report insurance premiums to the IRS, the counts of workers in the IRS data with employer-sponsored insurance are reweighted based on the firm size to match the overall count of workers by firm size with employer-sponsored insurance from the CPS. Shares are per \$200 range.

Source: Authors' calculations of Census and IRS data.

Figure 2. Cumulative distribution of premiums for employer-sponsored insurance among single workers

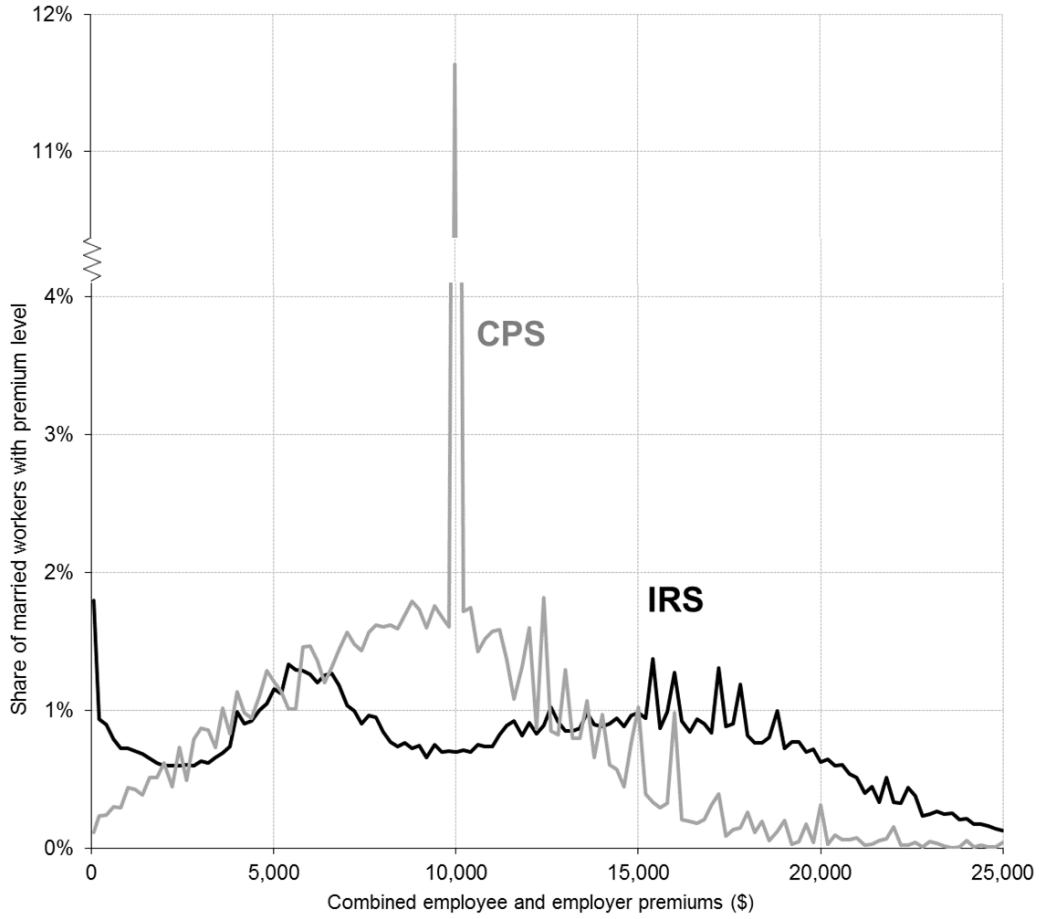


Note: See Figure 1.

Source: Authors' calculations of Census and IRS data.

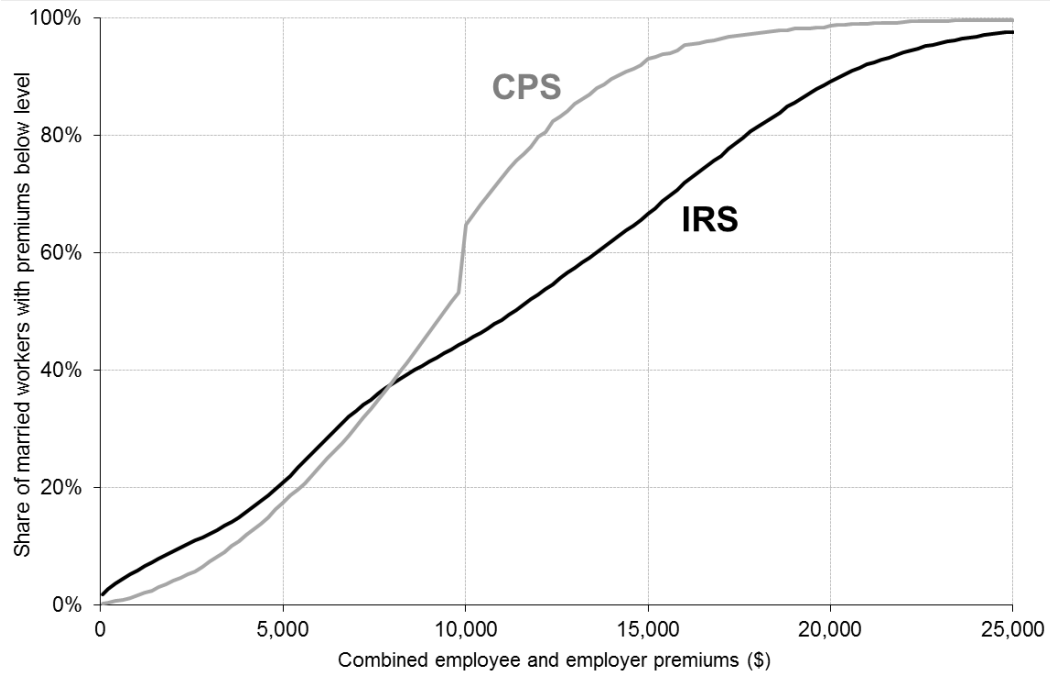
Figures 3 and 4 turn to the distributions of premiums for married workers. Here, much larger differences between the two datasets emerge. Most notably, the Census topcoding binds more frequently, reflecting the fact that family health insurance plans are more expensive on average than individual plans. Nearly twelve percent of insured married workers in the CPS data have total premiums at the \$10,000 topcode threshold and 24 percent are at the topcode level for the employer share of contributions, even if their combined employer and employee premiums are above this level. As a result, while one-third of married policyholders in the IRS data have plans that cost over \$15,000, just 7 percent of married policy holders in the CPS data have plans this expensive.

Figure 3. Distribution of premiums paid for employer-sponsored insurance among married workers



Note: See Figure 1. Shares are per \$200 range.
Source: Authors' calculations of Census and IRS data.

Figure 4. Cumulative distribution of premiums for employer-sponsored insurance among married workers



Note: See Figure 1.

Source: Authors' calculations of Census and IRS data.

While topcoding is a substantial source of the divergence between the datasets for married workers, it cannot fully explain the different distributions. Only 15 percent of policies for married workers have premiums between \$6,500 and \$9,900 in the IRS data, whereas 27 percent in the CPS data fall within this range. Since this is below the topcode threshold, the different frequencies in this range cannot be the result of CPS topcoding. Also, similar to that seen among single workers, the IRS data observes additional married workers with premiums under \$2,500, likely reflecting part-year workers who may be better captured in the IRS data.

Table 2 provides summary statistics for the health insurance premiums in the two datasets, separated by marital status and the presence of children or dependents. For the entire population, the median premium is approximately \$600 lower in the IRS data than in the CPS. But, because the CPS data

largely misses the upper tail of the distribution, the average premium is just over \$1,000 higher in the IRS administrative data.

Table 2. Health insurance premiums by marital status and presence of dependents

| | IRS | | | CPS | | |
|-------------------|-------------------------|-------------------|---------------------|-------------------------|-------------------|---------------------|
| | Policy-holders (thous.) | Mean premium (\$) | Median premium (\$) | Policy-holders (thous.) | Mean premium (\$) | Median premium (\$) |
| All | 78,744 | 8,958 | 7,000 | 78,744 | 7,891 | 7,595 |
| Single | 34,836 | 6,242 | 5,600 | 33,475 | 6,102 | 5,659 |
| With children | 7,368 | 7,395 | 6,000 | 6,413 | 8,152 | 8,336 |
| Without children | 27,467 | 5,932 | 5,600 | 27,061 | 5,616 | 5,248 |
| Married | 40,459 | 11,510 | 11,400 | 45,269 | 9,214 | 9,517 |
| With children | 20,683 | 12,598 | 13,200 | 24,667 | 10,071 | 9,999 |
| Without children | 19,777 | 10,371 | 9,600 | 20,601 | 8,188 | 8,261 |
| Non-filers | 3,449 | 6,458 | 5,000 | | | |

Note: See Figure 1. To preserve confidentiality, premiums in W-2 data are rounded to the nearest \$200.

Source: Authors' calculations of Census and IRS data.

Considering only single workers, both mean and median premiums in the two datasets are within \$100 of one another – which matches the relatively similar cumulative distribution function for single workers in Figure 2. On average, however, the CPS data overestimates the premiums of single workers with children and underestimates the premiums of single workers without children, relative to that reported to the IRS.¹⁰ Consistent with Figures 3 and 4, more substantial differences emerge for the mean and median premiums among married workers. The average premium for married workers in the

¹⁰ The Medical Expenditure Panel Survey (MEPS) Employer Survey observes that in 2015, the mean premium paid for a single plan from private employers was \$5,963 (median \$5,800), for a self-plus-one plan \$11,800 (median \$12,000), and for a family plan (\$17,322 (median \$17,000). These higher estimates for family and self-plus-one plans than observed in the CPS supports the idea that the CPS underestimates premiums, although the results are not directly comparable to those in Table 1 since the MEPS data is based on the number of individuals on the plan rather than the overall family structure. The gap between the CPS and IRS data is generally not attributable to proxy survey responses, as the distribution of premiums (available upon request from the authors) are very similar for insured individuals who answer the CPS themselves and those for whom another household member completes the survey.

CPS data is \$2,300 (20 percent) lower than that reported to the IRS on Form W-2, and the median premium is \$1,900 (17 percent) lower.

Relationship between Health Insurance Premiums and Wage Incomes

The rising values of fringe benefits, such as health insurance, may have offset potential wage gains for middle-income workers. To understand how administrative and survey-based premiums vary with wages, Table 3 documents the share of workers who receive insurance through their job and average premiums for each quintile of the wage distribution. For the middle quintile of workers, average premiums were about \$650 higher in the IRS data. These additional health insurance benefits suggest that combined total compensation is approximately 2 percent more for these middle-income workers than is observed in the unadjusted CPS data.

Table 3: Health insurance coverage and premiums by policyholder wages

| Wage quintile | IRS | | | CPS | | |
|---------------|-----------------|--------------------|-------------------|-----------------|--------------------|-------------------|
| | Mean wages (\$) | Policy-holders (%) | Mean premium (\$) | Mean wages (\$) | Policy-holders (%) | Mean premium (\$) |
| Bottom | 8,015 | 9 | 2,847 | 9,081 | 15 | 4,841 |
| Second | 19,493 | 34 | 5,068 | 23,077 | 14 | 5,595 |
| Middle | 32,553 | 62 | 7,371 | 36,819 | 62 | 6,625 |
| Fourth | 51,131 | 75 | 9,394 | 56,254 | 71 | 8,080 |
| Top | 132,177 | 82 | 12,012 | 129,502 | 75 | 10,624 |

Note: See Figure 1. Wage quintiles and mean wages are based on all workers with annual wages of at least \$2,000, including non-policyholders. Mean premiums are only among policyholders. Reflecting that not all firms are required to report premiums to the IRS, the share of workers in the IRS data with insurance are reweighted based on the firm size to match the overall count of workers by firm size with employer-sponsored insurance from the CPS.

Source: Authors' calculations of Census and IRS data.

Nevertheless, while including health insurance boosts median compensation, most of the benefits are earned by the top half of the distribution and its inclusion increases distribution-wide inequality (Kaestner and Lubotsky, 2016). Table 3 also shows that the insurance-wage gradient is greater in the IRS data than in the CPS – both when considering the share with insurance and the

average premiums among those who are insured. Administrative estimates of health insurance premiums would therefore imply more inequality in total compensation than the survey estimates.

Conclusion

Administrative data on health insurance premiums provides valuable information on the cost of employer-sponsored health insurance benefits. Using initial data from these IRS records, we observe that the Current Population Survey data understates the upper tail of the premium distribution and, as a result, understates average premiums. This suggests that total compensation may be higher than previously believed, also implying that employer-sponsored health insurance benefits may represent a larger share of employee compensation. For middle-income workers, this understatement represents approximately 2 percent of total compensation. However, we also observe that the underestimate is concentrated among higher income workers, thereby increasing the overall level of compensation inequality relative to that observed in the CPS data.

References

- Baicker, Katherine and Amitabh Chandra (2006). "The Labor Market Effect of Rising Health Insurance Premiums." *Journal of Labor Economics* 24(3): 609-634.
- Burtless, Gary, and Svaton (2010). "Health Care, Health Insurance, and the Distribution of American Incomes." *Frontiers in Health Policy Research* 13(1): 1-39.
- Burkhauser, Richard V., Jeff Larrimore, and Kosali I. Simon (2012). "A Second Opinion on the Economic Health of the American Middle Class and Why it Matters in Gauging the Impact of Government Policy." *National Tax Journal* 65(1): 7-32.
- Congressional Budget Office (2016). "The Distribution of Household Income, 2014." Available at: <https://www.cbo.gov/publication/53597>.
- Heim, Bradley, Ithai Lurie, and Kosali I. Simon (2017). "Did the Affordable Care Act Young Adult Provision Affect Labor Market Outcomes? Analysis using Tax Data." *Industrial and Labor Relations Review*. Available At: <https://doi.org/10.1177/0019793917744176>
- Heim, Bradley, Ithai Lurie, and Kosali I. Simon (Forthcoming). "The impact of the Affordable Care Act Young Adult Provision on Childbearing, Marriage, and Tax Filing Behavior." *Demography*.
- Internal Revenue Service (2018). "Form W-2 Reporting of Employer-Sponsored Health Coverage." Available At: <https://www.irs.gov/affordable-care-act/form-w-2-reporting-of-employer-sponsored-health-coverage>
- Janicki, Hubert, Brett O'Hara, and Alice Zawacki (2013). "Comparing Methods for Imputing Employer Health Insurance Contributions in the Current Population Survey." SEHSD Working Paper 2013-23.
- Kaestner, Robert, and Darren Lubotsky (2016). "Health Insurance and Income Inequality." *Journal of Economic Literature* 30(2): 53-78.
- Kolstad, Jonathan T. and Amanda E. Kowalski (2016). "Mandate-based health reform and the labor market: Evidence from the Massachusetts reform." *Journal of Health Economics* 47: 81-106.
- Piketty, Tomas, Emmanuel Saez, and Gabriel Zucman (2018). "Distributional National Accounts: Methods and Estimates for the United States." *Quarterly Journal of Economics* 133(2): 553-609.
- Romer, Christina and Mark Duggan (2010). "Exploring the Link between Rising Health insurance Premiums and Stagnant Wages." Available at: <https://obamawhitehouse.archives.gov/blog/2010/03/12/exploring-link-between-rising-health-insurance-premiums-and-stagnant-wages>