Non-Timed Coding Challenge - UChicago - Gottlieb

Kristian Blais

2022-11-20

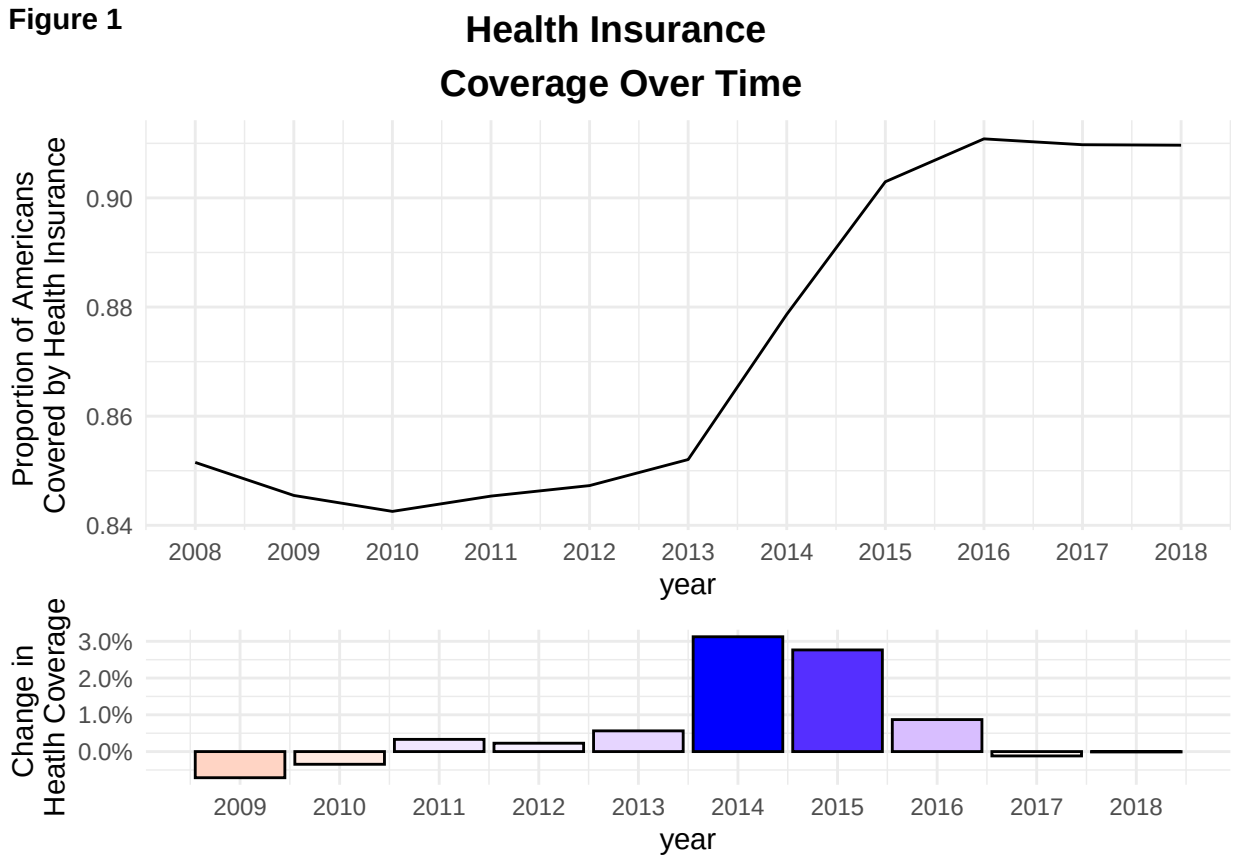
Analysis of Health Insurance Coverage in America from 2008 - 2018

Change in Overall Health Insurance Coverage Over Time

From 2008 to 2018, between 84% and 93% of all Americans were covered by some form of health insurance. For most of the sample period, health coverage is relatively steady. We do observe a small drop in coverage in 2008 and 2009, perhaps a result of people losing employment-coverage in the wake of the Great-Recession. However, this represents a decrease of less then 1% of the coverage rate for all Americans.

In fact, in most observations, year-over-year changes in the rate of coverage do not exceed $\pm$ 1%. However, there are two exceptions to this. There is a relatively large year-over-year increase in health insurance coverage between the years of 2013-2014, and 2014-2015, with both periods realizing an ~3% increase in health insurance coverage.

Figure 1



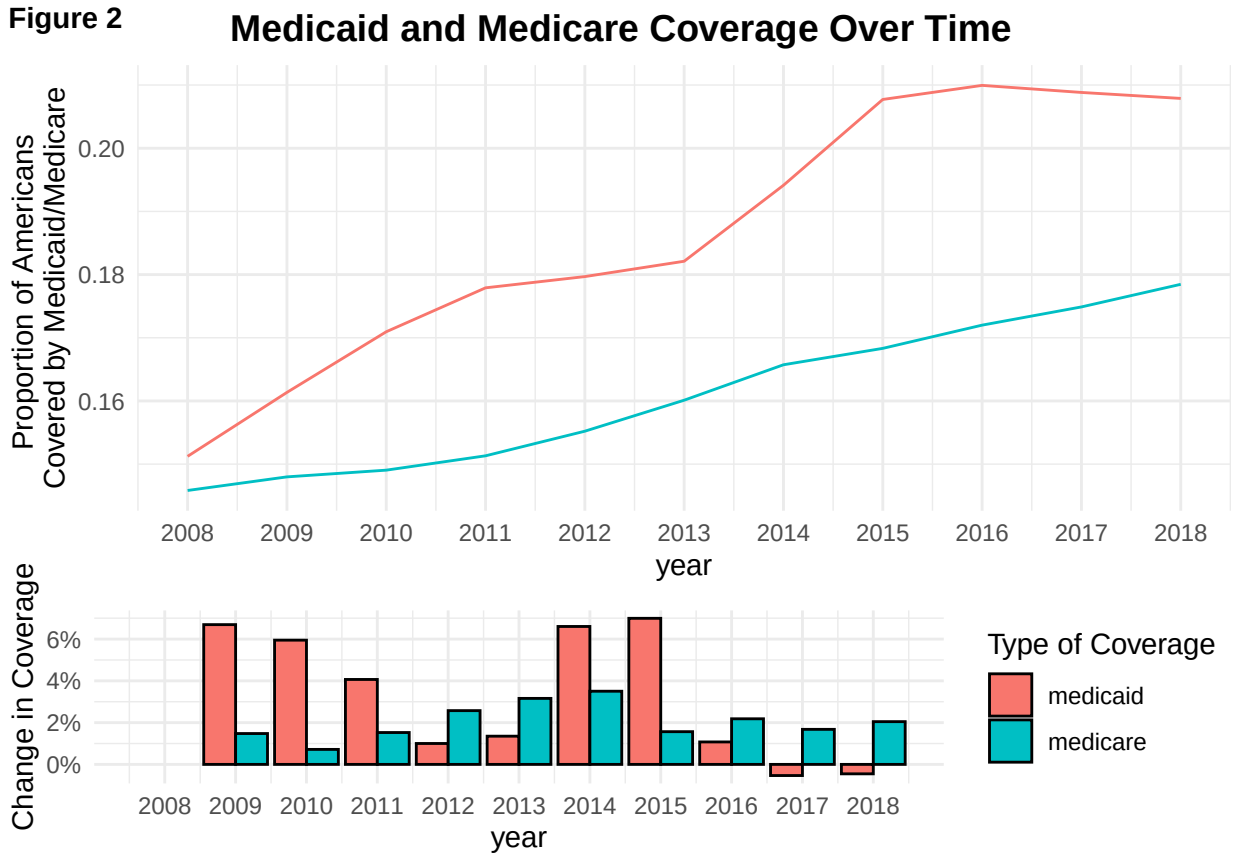
Change in Medicaid and Medicare Coverage over Time

As we can see in Figure 2, the proportion of Americans covered by Medicare is increasing for each year in our sample. Medicaid coverage is also increasing in every sample year, except for the last two years. That being said, coverage under Medicare and Medicaid is still relatively small, with either program covering less than 21% of the American population over the sample period.

We see large increases in Medicaid coverage between 2008-2011, and again between 2013-2014. As a result of the mandate of the Medicaid program - to provide health insurance to those with low incomes - the first large increase from 2008-2011 is most likely a result of the worsening socio-economic conditions brought in by the Great Recession. This relatively large increase in Medicaid coverage is perhaps masking the amount of people who lost private coverage between 2008 and 2010. The second large increase in Medicaid coverage corresponds to a similar jump in overall coverage over the same time period as shown in Figure 1.

Medicare coverage is increasing at about half the rate that Medicaid coverage is also increasing over the same time period. Although, we do see the highest increase in Medicare coverage in the years of 2013 and 2014. This aligns with similar outsized increases in Medicaid coverage and overall coverage over the same time period. Most likely, the increases in Medicare and Medicaid coverage lead to the increase in overall coverage in 2013 and 2014. These increases could conceivably result from the expanded eligibility for Medicaid that began in January 2014.

Figure 2



Time Trend Differences Between Different Groups of People

Differences in Race

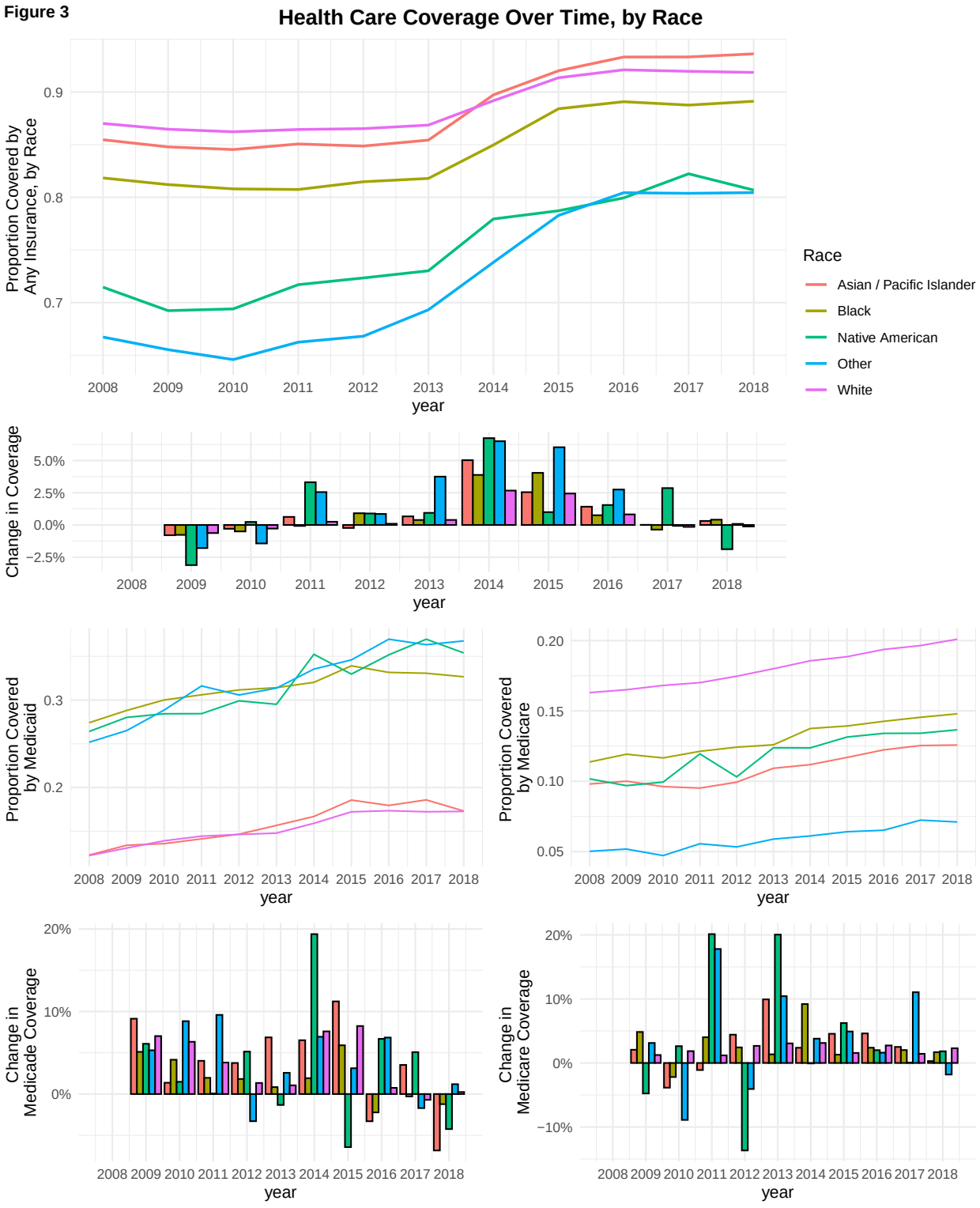
To simplify this analysis, I separate race into five broad categories: Asian / Pacific Islander, Black, Native American, White, and Other, with the white category containing both Hispanic and Non-Hispanic white

identifying people. If I were to have more time, I would define better racial categories for a more robust racial analysis.

Figure 3 summarizes the time trends of different types of health insurance for all five racial groups. We see that those who identify as White and Asian / Pacific Islander have the highest rates of coverage (~90% are covered), followed by those who identify as Black, Native American, and "Other" races, respectively. Although all races saw a net increase in coverage over the sample period, Native Americans, and those in in the "Other" race category saw the largest gains in health coverage.

Stark differences start to emerge when we isolate Medicaid/Medicare coverage between each group. Medicaid, a program intended to provide coverage for low-income people, covers approximately 30% of the Black and Native American population in the sample, over double the coverage when compared to White and Asian populations. Conversely, Medicare coverage is highest among the White population at ~90%, being about 5 percentage points higher than the Black and Native American Populations. Interestingly, the Asian / Pacific Islander population is more in line with the Medicare coverage rates of Black and Native American populations, despite trending alongside the White population with Medicaid and overall coverage.

Figure 3



Differences by State

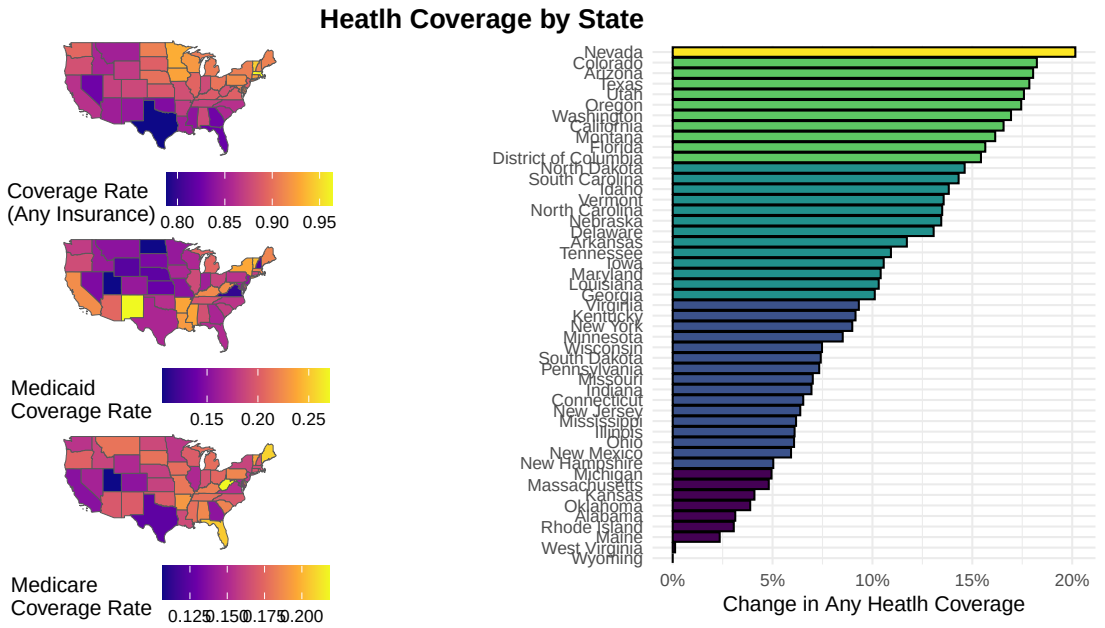
From Figure 4, we can see the average coverage rate for each state, averaged over our sample period. Non-continental states have been removed for readability, but should be included for future research. There

is much heterogeneity between the coverage rates, with the highest coverage states covering 15 percentage points more of their population compared with the lowest coverage states. Looking at the data we can also see emerging regional trends in health coverage, with the Mid-West and New England having relatively higher rates of health insurance coverage than the South. I am not sure how to pinpoint these regional differences. They could be caused by a heterogeneity in employment, with some states more inclined to have a larger proportion of employers who provide healthcare. Alternatively, these differences could track alongside racial patterns as seen in Figure 3. However, this is just conjecture and would require more analysis to provide a robust answer.

Furthermore, there is also a large difference in Medicare and Medicaid coverage by state. States such as West Virginia and Florida have a relatively higher proportion of citizens registered with Medicare, which could be due to their relatively larger aging populations who are becoming eligible for Medicare. On the other hand, New Mexico has a disproportionately large fraction of their population covered under Medicaid. This is a little odd, as even though New Mexico does have the third largest poverty rate in the nation, it's Medicaid coverage rate is still noticeably higher than Mississippi and Louisiana; two states with similar poverty rates to New Mexico. This discrepancy would also be an interesting avenue for further research, especially comparing the racial makeups of these states.

Over the period, all states saw an increase in the proportion of citizens covered under some form of health insurance. Nevada saw the largest increase out of any state, with Wyoming and West Virginia realizing almost no change in the coverage rate. Given more time, it would be interesting to investigate why Nevada specifically has gained so much coverage over this 10-year period.

Figure 4



Regressions

Finally, I will perform some simple regressions to understand how different factors effect the rate of health coverage. Since having coverage is a binary, I will be using a Logit regression to estimate the explanatory power. This is a very preliminary regression and should be taken with a grain of salt, as I will be ignoring the time series component of the data, which can lead to issues in auto-correlation.

As we can see in the regression table below, you are more likely to have health insurance and medicaid, and less likely to be on medicare if you are a woman. Similarly, those who are single are more likely to have medicare and medicaid coverage, but less likely to have any insurance then they're married counterparts.

Given more time, this type of regression would be a valuable avenue for further research. Time trends, robust

errors, a larger domain of explanatory variables, and the addition of interaction terms could provide insight into the nature of health insurance coverage in America.

Table 1:

	Dependent variable:		
	coverage_status		
	(Any)	(Medicaid)	(Medicare)
SEX = Female	0.275^{***}	0.143^{***}	-0.029^{***}
	(0.004)	(0.003)	(0.005)
age	0.028^{***}	0.005^{***}	0.239^{***}
	(0.0001)	(0.0001)	(0.0004)
MARST = Spouse Absent	-1.130^{***}	1.293^{***}	0.198^{***}
	(0.011)	(0.011)	(0.018)
MARST = Seperated	-1.240^{***}	1.540^{***}	0.503^{***}
	(0.011)	(0.011)	(0.018)
MARST = Divorced	-0.901^{***}	1.093^{***}	0.334^{***}
	(0.006)	(0.006)	(0.007)
MARST = Widowed	-0.029^{**}	1.144^{***}	0.723^{***}
	(0.013)	(0.007)	(0.010)
MARST = Single	-0.431^{***}	1.328^{***}	1.176^{***}
	(0.005)	(0.005)	(0.008)
Income	0.00002^{***}	-0.00003^{***}	-0.00001^{***}
	(0.000)	(0.000)	(0.000)
Constant	0.969^{***}	-2.902^{***}	-15.475^{***}
	(0.008)	(0.007)	(0.023)
Observations	3,429,162	3,429,162	3,429,162
Log Likelihood	−1,098,084.000	−1,467,020.000	−572,687.600
Akaike Inf. Crit.	2,196,186.000	2,934,058.000	1,145,393.000
Note:			*p<0.1; **p<0.05; ***p<0.01

Conclusion

There are many interesting research questions one could persue with this data. One of the pieces of information I believe would be the most insightful would be understanding how the types of employers in a state effect the rate of health insurance coverage. For example, states that attract energy industries may have a different healthcare coverage composition as compared to states that attract technology industries. Additionally, smaller geographic bounds, such as counties, PUMAS (Public Use Microdata Areas), or Census Tracts could provide insightful intricacies that may be lost in aggregation.

If I had more time I would like to look at race with more detail, as the "Other" race category was exhibiting interesting trends that were different from the rest of the racial groups. By unaggregating them, we could

perhaps find valuable insight into those who are not properly covered, and why. Furthermore, I would like to have spent more time looking at the intersection of different populations (e.g. Health Coverage by Race and State and Income Level).