Alain Brou **Prof Kevin Foster**

Ibrahima Dokoure

Econometric Project

THE HEALTH CARE COVERAGE IN THE UNITED STATES

INTRODUCTION

The health care system in the United States is more complex because there is not a single and national system like in the many other industrial countries such as Canada, France, Germany, … It is now composed of the Patient Protection and Affordable Care Act (ACA), known popularly as Obamacare that Democrats passed in 2010 with no Republican support and many other private insurance compagnies. Even with the competition between those private insurances and the ACA, the cost of the health care is still higher for many Americans, it continues to increase year to year and the coverage rate is lower among low incomes and older peoples and generally among minorities groups.

In this paper, our principal goal is to analyze the health care coverage in the United States and more precisely to find out with data the coverage rate among poor with low incomes, older people in minorities group like African Americans and Hispanics. Somme academic and peer reviewed journals are going to be used in the order to present the overview of the health care system in the United States and the coverage rate among minorities group before confirming those previous studies by our own research.

LITERATURE REVIEW

“Journal of Health Care for the Poor and Underserved” is an academic journal written by Brennan Virginia from the Johns Hopkins University Press. Brenna underlined that poor people and those with existing conditions are suffering during this period of covid-19 as they did in past because of many litigations of ACA and the lower coverage rates among minorities group in US, which is linked to the health care cost what she called “Poor and Underserved” people.

“Eliminating Waste in US Health Care” is an academic journal written by Andrew Hackbarth and Donald M. Berwick and was published in Jama.com. In this article, they pointed out that the US health care is too costly and its ratio of the GDP is too higher and that has negative impacts on other government programms. They proposed some solutions to reduce wastes resources in

the the US health care system and at same time to make it affordable to every patient (specially low income and minorities groups).

“Socioeconomic, Geospatial, and Geopolitical Disparities in Access to Health Care in the US 2011-2015” is an academic journal written by Samuel D. Towne Jr and was published in International Journal of Environmental Research and Public Health. Towner Jr studied impacts on forgone medical care rates by some sociodemographic factors such as education, sex, race/ethnicity, age; socioeconomic factors like income; geospatial factors such as rurality, census region; and geopolitical factors such as state Medicaid Expansion or not among minorities groups: Hispanics, African Americans, Indian Americans, … He did find out that all those factors may play roles in whether an individual seeks or not a medical care even though he or she needs it in last twelve months prior to his research because of the cost related to the health care in the country.

Brennan pointed out that there are underserved people in the US who suffered and continue to suffer during this Covid-19 crisis and past health crisis. Their health and well-being are impacted because of police violence and because they could not have access to an affordable health care. He put it, “The organizing construct for JHCPU—socially defined populations that are underserved in ways affecting their health and well- being—places every one of these three titanic issues of 2020 in our domain. Prior to, during and after these three crises, JHCPU has been publishing on health policy, racial/ ethnic health inequities (extremely wide during COVID-19), and violence as a public health issue affecting underserved populations. Sadly, this year we have all three threatening to swamp underserved populations: health policy will continue to be most strongly influenced by the ultimate shape of the ACA, which is still being litigated; racial/ ethnic inequities in the incidence of and mortality from COVID-19 persist; and violence—including police violence against minority group members and the cognitively disabled—remains in the spotlight because it continues to occur in chilling ways relevant for the country as a whole (Brennan 2020)”.

Andrew Hackbarth and Donald M. Berwick demonstrated that the current health care system in the US cost a lot even though the system is politicized by Republicans and Democarts or Conservatives and liberals or Red States and Blue and polarized. Also, they recommended to focus on revealing many wastes in its resources in the order to help other important economic sectors to grow and favorize competitiveness among them. They wrote, “NO MATTER HOW POLARIZED politics in the United States have become, nearly everyone agrees that health care costs are unsustainable. At almost 18% of the gross domestic product (GDP) in 2011, beaded for 20% by 2020,’’^ the nation’s increasing health care expenditures reduce the resources available for other worthy government programs, erode wages, and undermine the competitiveness of US industry. Although Medicare and Medicaid are often in the limelight, the health care cost

problem affects the private sector just as much as the public sector. Both need serious relief (Hackbarth 2012)".

Furthermore, they think that ACA need some serious examinations in the order to reduce its burden costs on Government; however, they recommended that it should be done very carefully because vulnerable people such as old people who are beneficiaries of Medicaid and low incomes people in general may see their health care negatively impacted. They continue, "The cost reductions in the ACA are necessary and prudent, but if other initiatives to cut spending are taken too far or too fast, they become risky. Vulnerable Medicaid beneficiaries and seniors covered by Medicare with marginal incomes may find important care services out of reach, either because they cannot afford the new cost-sharing, because clinicians and hospitals have withdrawn from local markets, or both (Hackbarth 2012)".

Also, they argued that both strategies are needed: a new reconstruction of the current health care system in the country in the order to eliminate many of its wasted resources and to make it affordable to everyone. They argued, "Nonetheless, if the United States is to reconstruct a health care industry that is both affordable and relentlessly focused on meeting the needs of every single patient and family, waste reduction (that is, the removal of non-value-added practices in all their forms) is the best strategy by far (Hackbarth 2012).

Towne Jr. seems to be the one who studied seriously the health care coverage rates among minorities groups in the US and came up with very interesting results among targets of this research paper. He identified serious obstacles that could prevent some Americans to forgone needed health care even though they may have what it is called "pre-existence conditions" that are animating every political debate in this country about the building of the eventual and national health care system. He put it, "Individuals forgoing needed medical care due to barriers associated with cost are at risk of missing needed care that may be necessary for the prevention or maintenance of a chronic condition among other things. Thus, continued monitoring of factors associated with forgone medical care, especially among vulnerable populations, is critical. National survey data (2011–2015) for non-institutionalized adults residing in the USA were utilized to assess forgone medical care, defined as not seeking medical care when the individual thought it was necessary because of cost in the past 12 months (Towne Jr. 2017)".

He used the National survey data (2011–2015) for non-institutionalized adults residing in the USA for his econometric study. The first interesting factor that he found to be one of those that may explained causes of the low health care coverage rates among minorities groups is socioeconomic factors such as incomes that are linked to economic cycles. He wrote, "Socioeconomic factors such as income may influence whether an individual seeks medical care when they believe it necessary. Further, differences in forgone medical care have been shown

throughout times of economic downturns with particular gaps in access to medical care facing minority adults. These gaps in access to needed care have been shown to be different across social determinants of health (Towne Jr. 2017)".

Furthermore, he used the World Organization's Framework for Action and Social Determinants of Health to prove that the health care coverage rates among minorities groups and their lower frequencies to seek health care when they needed them or to see doctors are very low around the world. He is not an American problem; it is a global one. He wrote, "The World Health Organization's Framework for Action on the Social Determinants of Health postulates that health and health-related outcomes can be influenced by the structural and social determinants of health and health-related outcomes. Specifically, race/ethnicity, sex, education, and individual-level socioeconomic factors (e.g., income) serve as social determinants of health in this framework".

He concluded that there are other determinants beyond the socioeconomic factors that may prevent minorities groups composed of African Americans, Hispanics, American Asians, Native Americans, … to seek health care or to see doctors such as some sociodemographic factors such as education, sex, race/ethnicity, age; socioeconomic factors like income; geospatial factors such as rurality. He put it, "In addition to these individual-level factors, the structural determinants of health may include socioeconomic measures (e.g., macroeconomic policies) of larger geographic areas such as countries or states and public policies, including but not limited to health policies. Area-level characteristics that may impact access to medical care may include region, state socioeconomic variables based on income/poverty, and geopolitical characteristics, including Medicaid policies, which are based on research into early evaluations of Medicaid expansion that included a smaller geospatial scope (i.e., three US states with expansion relative to comparisons of non-expansion states). (Towne Jr. 2017)". For the Data of his research, Towne Jr. used the Behavioral Risk Factor Surveillance System (BRFSS) as the primary dataset used in his current analyses. The 2011 (n = 506,467), 2012 (n = 475,687), 2013 (n = 491,773), 2014 (n = 464,664), and 2015 (n = 441,456) BRFSS public use files were used. The BRFSS is an annual telephone survey of non-institutionalized adults.

For his dependent Variable, he used the Forgone medical care as the primary outcome of interest in the current study. The survey item included 'Was there a time in the past 12 months when you needed to see a doctor but could not because of cost?' This was coded as yes vs. no (Towne Jr. 2017).

For independent variables, "the BRFSS was used to measure all individual-level variables. Income was coded as Do not know/Not sure/Missing, Less than $15,000, $15,000 to less than $25,000, $25,000 to less than $35,000, $35,000 to less than $50,000, and $50,000 or more. Sex was coded as male or female. Education was coded as Did not graduate High School, Graduated

High School, Attended College or Technical School, or Graduated from College or Technical School. Race and ethnicity were included as non-Hispanic White, non-Hispanic Black or African American, non-Hispanic American Indian or Alaska Native, non-Hispanic Asian, and Hispanic. An 'other' category was included, defined as individuals reporting any of the following classifications: Native Hawaiian or Other Pacific Islander, no preferred race, Multiracial but preferred race not asked, do not know/Not sure, or Refused. Those adults categorized as 'other' in terms of race/ethnicity are reported (Towne Jr. 2017).

For the results, he presented them, "the distribution of outcomes of interest by key characteristics. Overall, nearly 20% of adults between the ages 18 to 64 reported a time when they did not seek health care services because of cost in the past 12 months (forgone medical care). This dropped to 15% in 2015. The rates were highest for adults with the lowest incomes (less than $15,000 annually) and among adults with the lowest education levels (without a high school degree or equivalent). The rates of forgone medical care were also highest for Hispanic adults, followed by Black or African American and American Indian or Alaska Native adults (not including those categorized as 'other') at nearly twice the rates of White or Asian adults. Further, individuals in rural areas had higher rates of forgone medical care than those in urban areas. The rates of forgone medical care were highest in the South (Towne Jr. 2017).

METHODS

Data were obtained from the Behavioral Risk Factor Surveillance System (BRFSS), an annual survey fielded by Centers for Disease Control and Prevention. BRFSS uses random-digit-dial, disproportionate stratified sampling design, and is administered over telephone to a representative sample of the U.S. population aged >=18 years living in households. The coverage ranged from 87% to 98% across states and was lower in the South, for minorities, and for the poor, because of their lower telephone coverage. Details of the BRFSS survey methodology are published elsewhere32–34 BRFSS data 2019 is collected and analyzed in 2020.

MEASURES

The health coverage was measured by the question:
Do you have any kind of health care coverage, including health insurance, prepaid plans such as HMOs, or government plans such as Medicare, or Indian Health Service? Then yes or no were the elements. This question is part of the BRFSS optional module on race.

The question of race was measured by the question: Preferred race category? Response options included White, Black or African American, Asian, American Indian or Alaska Native (henceforth referred to as Native American).

Hispanic ethnicity was assessed independently in a question that preceded the question measuring racial background. Respondents were asked: Are you Hispanic or Latino? Response options were yes and no. Following the U.S. Office of Management and Budget, Hispanic ethnicity was treated as independent of race, meaning that members of any racial group could be Hispanic or non-Hispanic, and Hispanics could be members of any racial group.

The multidimensional concept of SES is typically measured by indicators such as income, education, and employment status. These dimensions are inter-related, but each captures a different type of health-relevant resources.38 The variables included in the study were level of school attended (less than high school, high school graduate, some college, Bachelor &master's degree); annual household income from all sources measured in eight categories and recoded to better understanding into eight categories between less than $10,000 and More than $75,000. Finally, the Gender and Age were used to catch the frame of our data and comprehend if there is different interest of coverage depending on age and gender. The Marital status as a component of the SES, is measured by the question: What is you your marital status? the categorical responses have been turned into dummy variables for the purpose of knowing how one influences the dependent variable.

HYPOTHESIS

Hypothesis 1: racial discrimination is more likely to negatively affect the health insurance coverage

Hypothesis 2: Higher education is related to higher likelihood of getting health insurance

Hypothesis 3: low income is less likely to get health insurance coverage

STATISTICAL ANALYSIS

Analyses were conducted using R, version 4.0. After calculating descriptive and bivariate statistics, multiple logistic regression models of determinants of health coverage care were estimated. Independent variables included race, Hispanic ethnicity, income, education, Marital Status, age, gender, employment status, earning a wage, gender, age. Curvilinear effects of age were examined. Models were estimated for the built sample called dataforanalysis. Robust estimators to account for deviations from normality were used and the random forest were used at last to better perform our regression.

RESULTS

To test our hypothesis, we constructed a multiple linear probability model, then a multiple logit regression and a random Forest models. In the objective to run those different regression, we first standardize our model such as:

sobj<-standardize (HCOV~X_AGE80+I(X_AGE80^2) + female+ Black+ White+ Asian+ HISPANIC+ (Black + White +Asian +HISPANIC) *female+Married+ Divorced+ Single+ NOHS+ HS + Somecol+ BMdegr +(female+ Black+ White+ Asian+ HISPANIC)*Lessthan25000 + Student+ Sefemploy+ Hired + Lessthan25000+Lessthan50000+Morethan75000, dat_train, family = binomial)once the model standardized under this form , we proceed to our regression.

The first regression

model_lpm1<-lm (sobj$formula, data = sobj$data)

Table1: Multiple linear Regression of health coverage in New York State results

coefficients	Estimated	St. Error	t.value	Pr(>|t|)
Intercept	0.8381162	0.0610193	13.735	< 2e-16 ***
X_AGE80	-0.0633673	0.0409861	-1.546	0.122159
I_X_AGE80.pow.2	0.0847852	0.0398001	2.130	0.033203 *
Femlae1	0.0111377	0.0175150	0.636	0.524878
Black1	0.0392870	0.0160401	2.449	0.014352 *
White1	0.0374437	0.0135655	2.760	0.005800 **
Asian1	-0.0124863	0.0217185	-0.575	0.565378
HISPANIC1	-0.0516758	0.0120839	-4.276	1.94e-05 ***
Married1	0.0248251	0.0071092	3.492	0.000484 ***
Divorced1	0.0259671	0.0097037	2.676	0.007478 **
Single1	0.0274699	0.0075241	3.651	0.000264 ***
NOHS1	-0.0803764	0.0270644	-2.970	0.002996 **
HS1	0.0099261	0.0262449	0.378	0.705291
Somecol1	0.0263258	0.0262451	1.003	0.315879
BMdegr1	0.0325463	0.0261993	1.242	0.214207
Lessthan250001	-0.0084319	0.0285507	-0.295	0.767754
Student1	0.0303473	0.0117222	2.589	0.009661 **

Sefemploy1	-0.0225435	0.0085629	-2.633	0.008500 **
Hired1	0.0167477	0.0059820	2.800	0.005137 **
Lessthan500001	-0.0137909	0.0086866	-1.588	0.112447
Morethan750001	0.0142790	0.0057136	2.499	0.012486*
female1: Black1	-0.0013863	0.0090772	-0.153	0.878621
female1: White1	-0.0066750	0.0073841	-0.904	0.366062
female1: Asian1	-0.0196057	0.0115892	-1.692	0.090769.
female1: HISPANIC1	0.0122996	0.0068814	1.787	0.073945.
female1: Lessthan250001	-0.0005769	0.0091261	-0.063	0.949602
Black1:Lessthan250001	0.0189148	0.0159763	1.184	0.236504
White1:Lessthan250001	0.0019210	0.0135643	0.142	0.887387
Asian1:Lessthan250001	-0.0128632	0.0216321	-0.595	0.552117
HISPANIC1:Lessthan250001	-0.0075768	0.0119335	-0.635	0.525514

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3031 on 4455 degrees of freedom

Multiple R-squared: 0.1255, Adjusted R-squared: 0.1198

F-statistic: 22.05 on 29 and 4455 DF, p-value: < 2.2e-16

Based on the multivariate linear table, the curvilinear effect of Age, Black, White, Hispanic, marital status, no high school level, Morethan750001 as income ,employment status and the composite effect between female and Hispanic, female and Asian are statistically significant to some extent. The curvilinear effect of Age has a positive effect on getting a health insurance. This means, when one in society grows older, it tends to secure a better health. The race in the business of getting health insurance matters. We saw three of the four variables describing the race are significant. However, their influence differs after the ethnicity. White (0.037) and Black (0.039) have a positive influence on the dependent variable health insurance. But this is not the case of Hispanic (-0.05) with a higher and negative influence. This could be the fact of the immigration status which leads many to hide. The female Asian (-0.019) has a negative influence on having health insurance. Yet the Hispanic female has a positive relation with the

access to health insurance. This could be the fat that Hispanics female get many children. Concerning the education attainment, we have only NOHS no high school as significant. This has a lot to say.

Now we examined the regression to the light of the Logit model. The results are show in the table 2

Table2: Multiple linear Regression of health coverage in New York State results

coefficients	Estimated	St. Error	t.value	Pr(>|t|)
Intercept	2.230947	0.529482	4.213	2.51e-05 ***
X_AGE80	-0.835849	0.439941	-1.900	0.057444.
I_X_AGE80.pow.2	1.085693	0.436404	2.488	0.012853 *
Femlae1	0.081530	0.163863	0.498	0.618799
Black1	0.249202	0.150525	1.656	0.097813.
White1	0.260634	0.113111	2.304	0.021210 *
Asian1	-0.192035	0.179775	-1.068	0.285432
HISPANIC1	-0.363016	0.104522	-3.473	0.000514 ***
Married1	0.224456	0.069464	3.231	0.001232 **
Divorced1	0.225548	0.104204	2.164	0.030428 *
Single1	0.230002	0.072816	3.159	0.001585 **
NOHS1	-0.396022	0.229354	-1.727	0.084224.
HS1	0.065928	0.226209	0.291	0.770709
Somecol1	0.231031	0.228064	1.013	0.311057
BMdegr1	0.332131	0.228281	1.455	0.145691
Lessthan250001	-0.023437	0.240679	-0.097	0.922427
Student1	0.319434	0.129484	2.467	0.013626 *
Sefemploy1	-0.167721	0.080801	-2.076	0.037919 *
Hired1	0.199133	0.061842	3.220	0.001282 **
Lessthan500001	-0.150568	0.086445	-1.742	0.081547.
Morethan750001	0.240339	0.074290	3.235	0.001216 **
female1: Black1	0.018714	0.087893	0.213	0.831394
female1: White1	-0.006371	0.068816	-0.093	0.926236

female1: Asian1	-0.184267	0.110499	-1.668	0.095396.
female1: HISPANIC1	0.008827	0.063550	0.139	0.889524
female1: Lessthan250001	-0.013157 Black1:Lessthan250001 White1:Lessthan250001 Asian1:Lessthan250001 - HISPANIC1:Lessthan250001	0.084692	-0.155	0.876547
Black1:Lessthan250001	0.157573	0.149491	1.054	0.291854
White1:Lessthan250001	-0.029872	0.113087	-0.264	0.791666
Asian1:Lessthan250001	0.079420	0.177927	-0.446	0.655335
HISPANIC1:Lessthan250001	-0.031168	0.103091	-0.302	0.762392

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 3262.5 on 4484 degrees of freedom

Residual deviance: 2782.9 on 4455 degrees of freedom

AIC: 2842.9

The variables which explained the dependent variables in the linear regression are almost the same as the logit regression. However, the logit regression yields more variables that influence the health insurance. the main interest is to know if there is any discrimination among health insurance coverage. The variable Black and white influence positively the health insurance coverage with the respective coefficients of 0.24 and 0.26. Hispanic race has a negative coefficient (-0.36). Race once again turns out to be consequent and significant in the health insurance coverage in New York. The marital status under the variable Married (0.22), Divorced (0.22) and Single (0.23) have positive coefficients. The education level under the High school plays a negative effect of getting health insurance in New York City. This draws us to comprehend that the population needs to be taught on the importance of getting any health insurance coverage. this point is shown to be statistically verified by the coefficient of the student variable (0.31) and is significant with alpha =5%. The Self-employed has a negative coefficient (-0.16), this means that the self-employed tend not to be insured to avoid expenses because knowing not the importance of health insurance. The Hired variable yields a positive influence on health insurance access. This is verified by the fact that workers in NY are

proposed by their employer the possibility to get health insurance. The income matters in the end, the people getting between $35000 and $49000 dollars have a negative coefficient (-0.15) and is statistically significant at 90%. On contrary, the income that is more than$75000 is highly significant and has a positive coefficient (0.24). This is an answer of one of our questions. The female Asian is statistically significant at 90% confidence interval. But the coefficient is negative (- 0.18) which shows that Asian female is less likely to get health insurance in New York. Yet, the variable associated such as race and Income has shown no significance to prove any discrimination for any race or low incomes.

LIMITATIONS

We found that there is multicollinearity in the variable of education such as High school level or Bachdegr and female. However, we have maintained these variables in the interest of our hypothesis to comprehend if there is any racial discrimination intervening in the access to health insurance. Which we understand could undermine the accuracy of our predictions.

DISCUSSION

This study makes several important contributions to the understanding health insurance coverage among individuals who have low income and faced discriminations in American Society. First, racial background clearly matters, with blacks and White Americans more likely to getting health insurance coverage while Hispanics are less likely to get health insurance even after adjusting for sociodemographic factors. Second, socioeconomic disadvantages, including lower level of education, lower income, and lower education, are linked to decreased health insurance coverage. Another funding is the high effect of student on the access to health insurance coverage. This is the fruit of lecture and awareness of disease and pandemic developed by students, which leads them to get any health insurance that could remedy in case of illness. The negative influence of self-employed people on having a health insurance coverage also draws the attention. This could be motivated by the fact that they do not have an Agency that oversees them as to offer the possibility to get health insurance coverage. As a recommendation, the governor of New York to encourage those in self employment should create an agency which could bring to them the awareness of the need to get insured. Also, the regression also showed us that the poor family such as less $25000 dollars or medium class are less likely to get insured. this could be related to the price because those with the highest income such as more than $75000 are more likely to get a health insurance. So, as previous studies proved it, the Socioeconomics determinants are significant to individuals and households to get health insurance, even in New York.

CONCLUSION

This study suggests that socioeconomic status and race are important social determinants of anyone getting health insurance in New York. Though the first two hypothesis has not been confirmed, the third one has. Also, the likelihood of getting insured comes from student, hired people and income more than $75000 dollars. this suggest that the sensibilization and education could lead to the decrease of noninsured people. the government should also create agency special for the self-employed people to encourage them to have a health coverage. Also, a point brought by this research is about Hispanic that have a negative coefficient regression on health insurance due to immigration issues. Policymakers should assist by policy those who do not have legal papers to get insured.