Spring 5-4-2021

Analysis of Loss of Work during the COVID-19 Pandemic in the United States

Mira Shah

ABSTRACT

Analysis of Loss of Work during the COVID-19 Pandemic in the United States

By

Mira A. Shah

May 4th, 2021

A multiple logistic regression was performed to predict the likelihood of work loss due to the COVID-19 pandemic in the United States based on predictors from the 2020 U.S. Census Bureau Household Pulse Survey. The nine predictors included week (time period), birth year, number of children, number of adults, sex, race, Hispanic origin, educational attainment, and marital status. The purpose of the study is to estimate if there is a difference in work loss status at the beginning of the pandemic and at the end of 2020. The results of the model showed that an increase in time since the start of the pandemic has led to increased odds of job loss. This makes sense because the unemployment rate has remained high. All predictors in the model were significant. Females, Hispanics, and Blacks have higher odds of job loss (0.7%, 29.9%, and 30.3%, respectively). Those who are younger and who have not graduated high school have higher odds for loss of work. Certain demographic groups are more likely to have a loss of work, and measures need to be taken to prevent this disparity. Also, because there was missing data in the survey results, multiple imputation was used to analyze 10% of the original sample. These results were not entirely comparable to the estimates using the entire original sample, but the multiple imputation procedure did show that the estimates were different.

Analysis of Loss of Work during the COVID-19 Pandemic in the United States

by

Mira A. Shah

B.S., MERCER UNIVERSITY

A Capstone Submitted to the Graduate Faculty
of Georgia State University in Partial Fulfillment
of the
Requirements for the Degree

MASTER OF PUBLIC HEALTH

ATLANTA, GEORGIA
30303

APPROVAL PAGE

Analysis of Loss of Work During the COVID-19 Pandemic in the United States

by

Mira A. Shah

Approved:

Dr. Ruiyan Luo
Committee Chair

Dr. Terri Pigott
Committee Member

April 23rd, 2021
Date

Author's Statement Page

In presenting this capstone as a partial fulfillment of the requirements for an advanced degree from Georgia State University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote from, to copy from, or to publish this capstone may be granted by the author or, in his/her absence, by the professor under whose direction it was written, or in his/her absence, by the Associate Dean, School of Public Health. Such quoting, copying, or publishing must be solely for scholarly purposes and will not involve potential financial gain. It is understood that any copying from or publication of this capstone which involves potential financial gain will not be allowed without written permission of the author.

Mira Shah

Table of Contents

List of Tables and Figures

Introduction

The ongoing pandemic caused by the coronavirus disease, or COVID-19, has led to many unfortunate circumstances throughout the world, including the United States. Not only has the pandemic led to over two million lives being lost worldwide, but it has also created many social and economic challenges (Schnirring, 2021). On March 11th, 2020, the World Health Organization declared the outbreak a pandemic (World Health Organization, 2020). Governments and agencies had to act quickly to control the spread and reduce the number of lives lost due to the virus. Public health officials advised individuals and their families to avoid contact with other individuals and to stay at home as much as possible (National Conference of State Legislatures, 2020). Many non-essential businesses were forced to close or voluntarily closed to protect their employees. Some individuals were able to continue their work through teleworking. Other individuals were not as fortunate, and their jobs were impacted by social distancing recommendations and the partial economic shutdown (Engemann, 2020). Those who have experienced a loss of work personally or within their household have been associated with having financial struggles such as having trouble paying bills, or rent or used money from their savings or borrowed money from someone they know (Parker, Minkin, & Bennett, 2020). There is very limited research about the effects of the pandemic because it is ongoing. This type of research is necessary because it can help policymakers understand how they can create policies and direct funds to help those affected by the pandemic. The aim of this study is to complement existing research. The research question is to predict the likelihood of loss of work within U.S. households at the beginning of the pandemic and the end of 2020.

Methods

Data Source

This study used secondary, national-level data from the 2020 Household Pulse Survey collected by the U.S. Census Bureau. The purpose of the survey was to understand and "measure the social and economic impacts" of the COVID-19 pandemic in a timely and efficient manner in order to aid recovery. This research was a collaborative effort between the Bureau of Labor Statistics, the National Center for Health Statistics, the U.S. Department of Agriculture's Economic Research Service, the National Center for Education Statistics, the Department of Housing and Urban Development, and the U.S. Census Bureau. The survey collected information about how education, employment, food security, health, housing, social security benefits, household spending, consumer spending, intention to receive the COVID-19 vaccination, and transportation have been affected by the pandemic (Fields, et al., Forthcoming). Data were collected weekly at the beginning of data collection, but the collection frequency was switched to a biweekly basis in order to collect more data during the survey periods.

The first week of data collection was April 23rd through May 5th, 2020, which was the first week that will be used in this study. The last week of data collection in 2020 was December 9th through December 21st, 2020 (Fields, et al., Forthcoming). This is the week that the first week will be compared to, so Week 1 and Week 21 will be compared. Data collection has continued into 2021 because the pandemic is ongoing.

Participants

The target population of this study is all adults over the age of 18 living in the U.S. In order to sample participants, the U.S. Census Bureau utilized a Master Address File to identify housing units. This file did not have contact information for the housing units, so the U.S. Census Bureau matched these addresses with email addresses and phone numbers from a Contact Frame. Addresses were randomly selected to participate in the survey. Email addresses and

phone numbers were rotated daily until all contact information was exhausted or the survey was complete for the housing unit. Participants received an email or a text with a link to participate in the survey (Fields, et al., Forthcoming). Participants were only interviewed once, and there was no way to track a single household over time.

Measures and Variables

The binary response of interest is work loss which is measured as if an individual or anyone in his or her household had experienced unemployment since March 13th, 2020. There were four numeric predictors: birth year, number of children, number of adults, and week. There were also five categorical predictors: sex, race, Hispanic origin, educational attainment, and marital status. Birth year was the age of the respondent, and it only included those who were 18 years and older at the time of the survey, so the birth year of 2002 was the greatest birth year that a respondent could include in their interview. The variable week is the time variable with two options: week 1 vs. week 21 of the survey. Race was categorized by White, Black, Asian, or any other race alone or race combination. Educational attainment was categorized by less than high school, some high school, high school graduate or equivalent, Associate degree, Bachelor's degree, or Graduate degree. Marital status was categorized by now married, widowed, divorced, separated, or never married.

Statistical Analysis

Data were downloaded from the U.S. Census Bureau's Household Pulse Survey Public Use Files. Variables were recoded as needed. The two separate weeks of data sets were combined by the common variable "week." This was done using the appending technique where the observations collected at week 21 were added to the same data set that contained the observations at week 1. An age variable was created by subtracting the birth year from the year 2020. Recoding was performed to make it easier to understand the analysis. Demographic

variables such as age, sex, race, Hispanic origin, educational attainment, marital status, number of children, and number of adults were edited using simple hot deck imputation by the U.S. Census Bureau before the data was released. Descriptive statistics were conducted to understand the demographic characteristics of the target population at each time point. A chi-square test of independence was conducted to determine if there was an association between each categorical predictor and the categorical outcome. An independent samples t-test was performed to compare the means of two independent groups for the continuous variables. There was an assumption of independent samples because those who were included once in the survey were not included again.

Multiple logistic regression was conducted to study the effect of the demographic variables and model the probability of an individual having a loss of work since the start of the pandemic (Proc Logistic). This model excluded any observation with missing values. Statistical analysis was performed using SAS 9.4 statistical software. An alpha level of <0.05 was considered statistically significant for statistical tests.

There were missing data for work loss status, so multiple imputation using the fully conditional method was implemented to address this issue of missing data. Work loss was the only variable with missing data, and there were 843 cases of missing work loss status in the full set of data. Due to the original sample size being so large, 10% of the sample (n = 14436) was selected using the simple random sampling method for the multiple imputation procedure (Proc Surveyselect). Ten imputed datasets were created using Proc MI, where a generalized logit distribution was assumed to impute the work loss variable, which was a nominal categorical variable. These 10 imputations were pooled together to identify the parameter estimates (Proc MIanalyze). The odds ratios for the imputed data set was calculated separately because it was not

produced in the output. The odds ratio was calculated with e to the power of the regression coefficient (e^{β}.) There were only 78 cases of missing work loss status in the 10% subset of the sample. A weighting procedure was applied within each state where data was collected from. This procedure was done in a four-step process to consider nonresponse and to account for the demographics of the individuals who were interviewed.

Results

After applying the weights to the sample, there were 143,514 observations included in this study. Table I shows the descriptive statistics for the demographic variables grouped by loss of work or not. All variables were statistically significant. For both groups, age ranged from 18 to 88 years old, the number of adults in the household ranged from 1 to 10, and the number of children ranged from 0 to 5. Among those who had a loss of work, those who considered their race as white (alone) had the highest percentage. The percentages for gender, race, ethnicity, educational attainment, and marital status are not equal, but this is not something to be concerned about since random sampling was done and the sample size is so large.

Table 2 shows the parameter estimates and odds ratios for the multiple logistic regression. When controlling for other variables, the odds of loss of work at Week 21 are 1.126 times the corresponding odds at Week 1. For every unit increase in age, it is expected that the odds of work loss decrease by 2.2%, given that all other variables in the model are held constant. Females have 0.7% higher odds of loss of work when compared to males. Hispanics have 29.9% higher odds of loss of work than non-Hispanics when all other variables are the same. Compared to Whites, Blacks have 30.3% higher odds of loss of work, Asians have 11.6% higher, and those who consider their race as other are 28.6% higher. Compared to those who have less than high school education, those with some high school education have 8.7% lower odds of loss of work, high school graduates have 20% lower, those with some college education have 20.3% lower,

Associate degree graduates have 19.7% lower, Bachelor's graduates have a 42.8% lower and Graduate degree graduates have 53.2% lower. Compared to those who are married currently, widowed individuals have 27.3% lower odds, divorced individuals have 34.8% higher, separated individuals have 68.8% higher, and single individuals have 1.8% higher.

Table 3 shows the parameter estimates of the multiple imputation analysis. Some of the variables in the imputed model are not significant as compared to the full sample, where all variables were significant. The estimates are different as well because only 10% of the data was used for the imputation. More about this is discussed in the limitations section. When controlling for other variables, the odds of loss of work at Week 21 are 1.180 times the corresponding odds at Week 1. For every unit increase in age, it is expected that the odds of work loss decrease by 2%, given that all other variables in the model are held constant. Females have 8.7% lower odds of loss of work when compared to males. Hispanics have 35.6% higher odds of loss of work than non-Hispanics when all other variables are the same. Compared to Whites, Blacks have 29.3% higher odds of loss of work, Asians have 6.2% lower, and those who consider their race as other are 22% higher. Compared to those who have less than high school education, those with some high school education have 24.9% lower odds of loss of work, high school graduates have 28% lower, those with some college education have 34.2% lower, Associate degree graduates have 32.8% lower, Bachelor's graduates have a 52.6% lower and Graduate degree graduates have 59.7% lower. Compared to those who are married currently, widowed individuals have 8.8% lower odds, divorced individuals have 50.4% higher, separated individuals have 75% higher, and single individuals have 12.3% higher.

Discussion

The outcome variable loss of work is an indirect measure of work loss because it asks if the participant themselves or anyone else in their household has had a loss of work. This has to

be kept in mind when trying the understand the meaning behind the results. Results for the model using the full sample are slightly different from the results using the subset sample. Both models showed that an increase in time since the start of the pandemic has led to an increased odd of loss of work. This makes sense because the unemployment rate has remained high. Figure 1 shows the unemployment rate from March 2019 to December 2020. The unemployment rate from March 2020 is higher than the rate from December 2020, but the December 2020 unemployment rate is still high when compared to that of December 2019 (U.S. Department of Labor, n.d.). The high unemployment rate could mean that there are not many jobs available that used to be before the pandemic, but a different study would need to be conducted to confirm this. In the summer of 2020, many states began reducing their restrictions, and businesses opened again. However, not everything was back to how it was before the pandemic began. Another finding from the results is that certain demographic groups have higher odds of loss of work due to the pandemic. For the model using the full sample, females, Hispanics and Blacks have a higher odds of job loss. The odds of job loss are only slightly higher for females than males, but it is still a difference that is important to note. There have been many reports of females voluntarily leaving the workforce to stay home. The model using the subset of data actually found that males have a higher odds of loss of work. However, this difference could be due to the smaller sample size. Job loss during the pandemic varies across racial and ethnic groups and genders, and policymakers need to be aware of this difference. Specific measures would need to be taken to prevent this from happening to minority groups and take steps to fix the issue. Policymakers need to ensure that certain segments of the population are not left behind during economic recovery from the pandemic. The confidence intervals for the odds ratios are the same as the point estimates because the standard error is quite small, while the sample size is very large.

Limitations

There are several limitations of this study. The first is that this publicly available data set did not provide any form of an identification variable for the observations. A longitudinal study tracking the participants for a more extended period might have resulted in better data with more meaningful information. As with any large national survey, nonresponse bias was also an issue because it could have created higher variances due to a smaller sample size. The number of people who reported a loss of work could be different from the true proportion of those who had a loss of work. For the data that was provided, the researchers used hot deck imputation for the demographic variables, which was needed for weighting purposes (Fields, et al., Forthcoming). The researchers explained how they proceeded to do this, but it could have been better to provide an original data set with missing values and allow others to choose how to impute the missing data on their own. This method is also outdated, and the theory behind used it is not well-developed. Another limitation of this study was that the multiple imputation procedure performed was only done so using 10% of the data, and this did not consider the weighting variable.

Conclusion

The COVID-19 pandemic has caused a strain on the U.S. workforce and job market. The pandemic has impacted everyone, but it has done so unequally. This study found that there are inequalities in job loss during the COVID-19 pandemic, where some groups of individuals have higher odds for job loss than others. Those who have lost their jobs are at risk for losing their health insurance coverage and may even find it difficult to find a new job during the pandemic. Prolonged loss of work is a major concern for some individuals because it can lead to many other implications such as food insecurity, physical and mental health issues, financial struggles, and more. The pandemic has already been so difficult for everyone, so it is important that the

government and officials can do what they can to return to normal prior to the pandemic. Future research could study whether the several stimulus payments from the government helped reduce some of the financial burdens of loss of work and loss of income.

References

Engemann, K. (2020, October 14). *How Has the COVID-19 Pandemic Affected the U.S. Labor Market?* Retrieved from Federal Reserve Bank of St. Louis: https://www.stlouisfed.org/open-vault/2020/october/how-covid19-pandemic-has-affected-labor-market

Fields, J. F., Hunter-Childs, J., Tersine, A., Sisson, J., Parker, E., Velkoff, V., . . . Shin, H. (Forthcoming). *Design and Operation of the 2020 Household Pulse Survey.* Retrieved from U.S. Census Bureau: https://www2.census.gov/programs-surveys/demo/technical-documentation/hhp/2020_HPS_Background.pdf

National Conference of State Legislatures. (2020, May 13). *COVID-19: Impact on Employment and Labor*. Retrieved from National Conference of State Legislatures: https://www.ncsl.org/research/labor-and-employment/covid-19-impact-on-employment-and-labor.aspx

Parker, K., Minkin, R., & Bennett, J. (2020, September 24). *Economic Fallout From COVID-19 Continues To Hit Lower-Income Americans the Hardest*. Retrieved from Pew Research Center: https://www.pewresearch.org/social-trends/2020/09/24/economic-fallout-from-covid-19-continues-to-hit-lower-income-americans-the-hardest/#:~:text=Fully%2015%25%20of%20adults%20report,they%20are%20currently%20not%20employed.

Schnirring, L. (2021, January 14). *Global COVID deaths near 2 million as hospitals struggle*. Retrieved from Center for Infectious Disease Research and Policy: https://www.cidrap.umn.edu/news-perspective/2021/01/global-covid-deaths-near-2-million-hospitals-struggle

Sparks, C., & Saenz, R. (2020, August 11). *The Inequities of Job Loss and Recovery Amid the COVID-19 Pandemic*. Retrieved from University of New Hampshire: Carsey School of Public Policy: https://carsey.unh.edu/publication/inequities-job-loss-recovery-amid-COVID-pandemic

U.S. Department of Labor. (n.d.). Retrieved from U.S. Bureau of Labor Statistics: https://www.bls.gov/

UCLA: Institute for Digital Research and Education. (n.d.). Retrieved from Multiple Imputation In SAS Part 1: https://stats.idre.ucla.edu/sas/seminars/multiple-imputation-in-sas/mi_new_1/

World Health Organization. (2020). *Timeline: WHO's COVID-19 response*. Retrieved from World Health Organization: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/interactive-timeline?gclid=CjwKCAiAo5qABhBdEiwAOtGmbrn9N2imdcoTT8lKjk-LMiU2H--1m3C8ZAqVraMtyMpWk--vBTcFgxoCIfwQAvD_BwE#!

Table 1. Marginal Association Between Study Characteristics and Status of Loss of Work

Variable		Loss of Work	No Loss of Work	P-Value
Age M(SD)		44.57 (15.49)	51.44 (17.73)	<.0001
Number of Adults M(SD)		2.94 (1.46)	2.48 (1.36)	<.0001
Number of Children M(SD)		0.87 (1.20)	0.66 (1.07)	<.0001
Gender %	Male	48.42	48.24	<.0001
	Female	51.58	51.76	
Ethnicity %	Non-Hispanic	78.95	86.94	<.0001
	Hispanic	21.05	13.06	
Race %	White, alone	72.50	78.93	<.0001
	Black, alone	14.35	10.62	
	Asian, alone	5.76	5.46	
	Other	7.39	5.00	
Education %	Less than high school	2.90	1.71	<.0001
	Some high school	6.62	4.23	
	High school graduate, or equivalent	32.94	30.02	
	Some college	22.97	19.34	
	Associate degree	9.91	8.98	
	Bachelor's degree	15.18	19.17	
	Graduate degree	9.49	16.54	
Marital Status %	Now married	50.72	59.38	<.0001
	Widowed	2.53	5.58	
	Divorced	12.75	11.89	
	Separated	3.13	1.71	
	Never married	30.88	21.44	

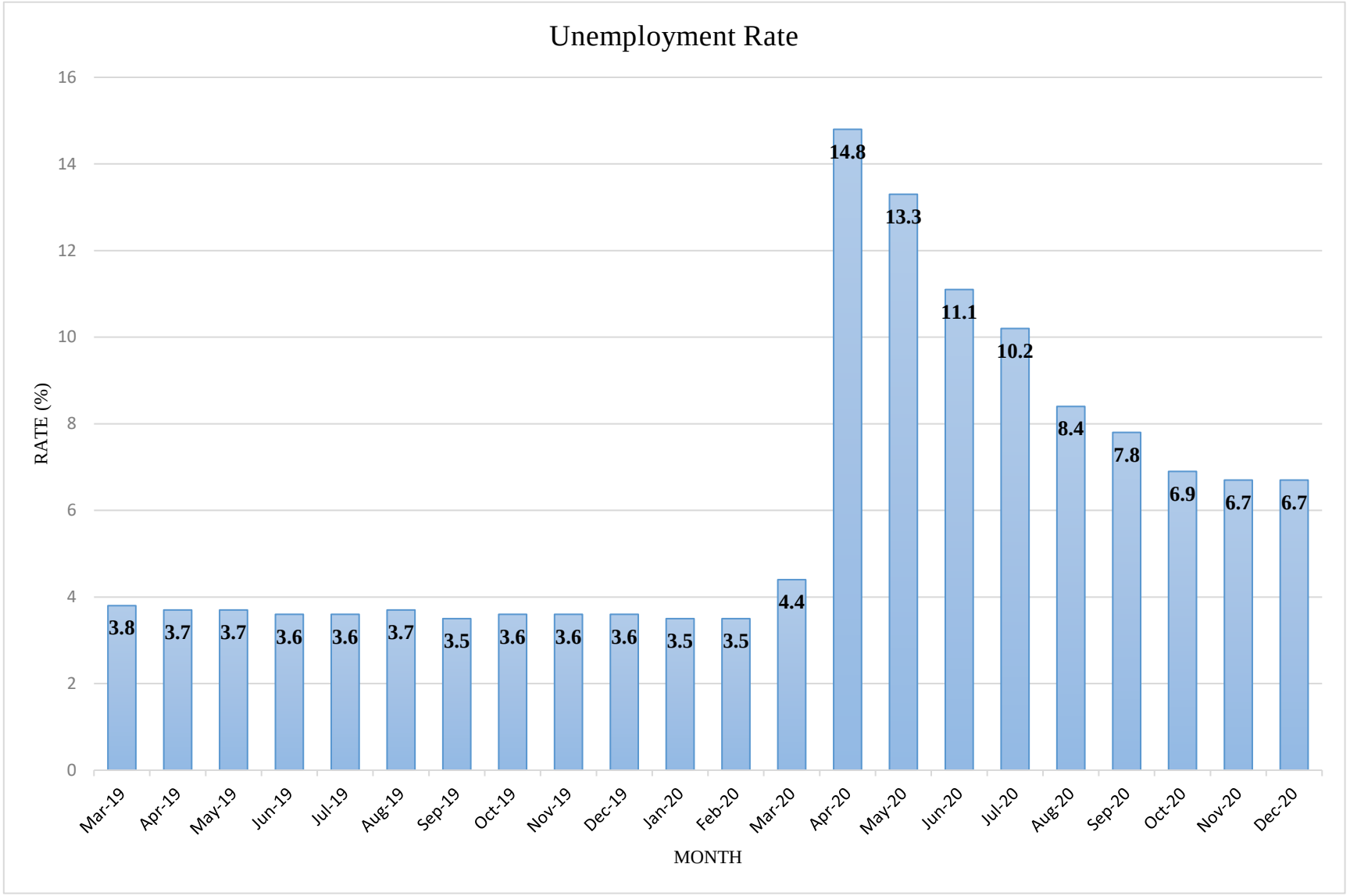
Table 2. Multiple Logistic Regression Estimates

Variable		Parameter Estimate (Standard Error)	P-Value	Odds Ratio (95% Confidence Interval)
Intercept		0.517 (0.000835)	<.0001	
Week	21	0.119 (0.000188)	<.0001	1.126 (1.126, 1.127)
	1	Reference		
Age		-0.022 (7.117E-6)	<.0001	0.978 (0.978, 0.978)
Number of Adults		0.212 (0.000070)	<.0001	1.236 (1.235, 1.236)
Number of Children		0.039 (0.000089)	<.0001	1.039 (1.039, 1.039)
Gender	Female	0.006 (0.000191)	<.0001	1.007 (1.006, 1.007)
	Male	Reference		
Ethnicity	Hispanic	0.261 (0.000265)	<.0001	1.299 (1.298, 1.299)
	Non-Hispanic	Reference		
Race	Black, alone	0.265 (0.000293)	<.0001	1.303 (1.302, 1.304)
	Asian, alone	0.110 (0.000416)	<.0001	1.116 (1.115, 1.117)
	Other	0.251 (0.000401)	<.0001	1.286 (1.285, 1.287)
	White, alone	Reference		
Education	Some high school	-0.091 (0.000758)	<.0001	0.913 (0.912, 0.915)
	High school graduate, or equivalent	-0.223 (0.000667)	<.0001	0.800 (0.799, 0.801)
	Some college	-0.227 (0.000680)	<.0001	0.797 (0.796, 0.798)
	Associate degree	-0.220 (0.000715)	<.0001	0.803 (0.802, 0.804)
	Bachelor’s degree	-0.549 (0.000690)	<.0001	0.578 (0.577, 0.578)
	Graduate degree	-0.760 (0.000704)	<.0001	0.468 (0.467, 0.468)
	Less than high school	Reference		
Marital Status	Widowed	-0.319 (0.000538)	<.0001	0.727 (0.726, 0.728)
	Divorced	0.299 (0.000300)	<.0001	1.348 (1.348, 1.349)
	Separated	0.524 (0.000639)	<.0001	1.688 (1.686, 1.690)
	Never married	0.018 (0.000262)	<.0001	1.018 (1.018, 1.019)
	Now married	Reference		

Table 3. Parameter Estimates of Multiple Imputation

Variable		Parameter Estimate (Standard Error)	P-Value	Odds Ratio Estimate
Intercept		0.284 (0.108)	0.0084	
Week	21	0.083 (0.018)	<.0001	1.180
	1	Reference		
Age		-0.020 (0.001)	<.0001	0.980
Number of Adults		0.347 (0.018)	<.0001	1.415
Number of Children		0.035 (0.018)	0.0505	1.036
Gender	Female	-0.046 (0.019)	0.0140	0.913
	Male	Reference		
Ethnicity	Hispanic	0.152 (0.032)	<.0001	1.356
	Non-Hispanic	Reference		
Race	Black, alone	0.159 (0.056)	0.0042	1.293
	Asian, alone	-0.162 (0.069)	0.0190	0.938
	Other	0.101 (0.065)	0.1229	1.220
	White, alone	Reference		
Education	Some high school	0.165 (0.135)	0.2230	0.751
	High school graduate, or equivalent	0.054 (0.060)	0.3701	0.72
	Some college	0.032 (0.053)	0.5521	0.658
	Associate degree	0.053 (0.062)	0.3895	0.672
	Bachelor's degree	-0.296 (0.052)	<.0001	0.474
	Graduate degree	-0.459 (0.055)	<.0001	0.403
	Less than high school	Reference		
Marital Status	Widowed	-0.286 (0.082)	0.0005	0.912
	Divorced	0.214 (0.048)	<.0001	1.504
	Separated	0.345 (0.099)	0.0005	1.75
	Never married	-0.078 (0.050)	0.1163	1.123
	Now married	Reference		

Figure 1. Unemployment Rate in the U.S. from March 2019 to December 2020



*Data from the U.S. Bureau of Labor Statistics