Economic Outcomes of the COVID Recession in the State of New York

With Socioeconomic Variables

Jesseca Johnson
Utica University
June 30, 2022

Literature Review - Context

Generally, there are persisting economic disparities among different racial groups in the United States (Wilson, 2020). For example, in 2019, although there were gains across Black, Asian, Hispanic, and non-Hispanic White households' median incomes, Black household median incomes remained the lowest of those groups at just over $46,000, while Asian household median incomes rose to just over double that amount at more than $98,000, putting that group firmly at the top. According to data from the 2019 Survey of Consumer Finances, a typical Hispanic family's net worth is just one fifth that of the typical White family, and a typical Black family's is just one eighth (Bhutta et al., 2020).

Historically speaking, economic recessions have had greater negative effects on the financial security and wealth of nonwhite populations in the United States than on majority Americans (McKernan et al., 2014). During the Great Recession of 2007-2009, majority families lost 10 to 13% of their wealth on average, whereas Black families lost 31 to 34% and Hispanic families lost 44 to 50%. A study by Fazzari and Needler (2021) found that job losses impacted Asian, Black and Hispanic workers more than White workers in both the Great Recession and the COVID-19 recession.

Research Question

- Background: COVID-19 recession (March-April 2020)
- *Did regions with higher proportions of racial and ethnic minorities experience disproportionate economic distress during and after the COVID recession?*
- Data: EASI
- Limitations of data: only have 2019-2020; would prefer 2019-2021 or even 2017-2022
  - Correlation of 2019 and 2020 median HH income: 0.993334

What is a Zip+4?

- ZIP: Zone Improvement Plan
  - Developed in 1963 for USPS efficiency
  - First digit: group of states
- Last four added later for precision
- Usually represents geographic area
  - City block
  - Group of apartments
  - Individual PO box
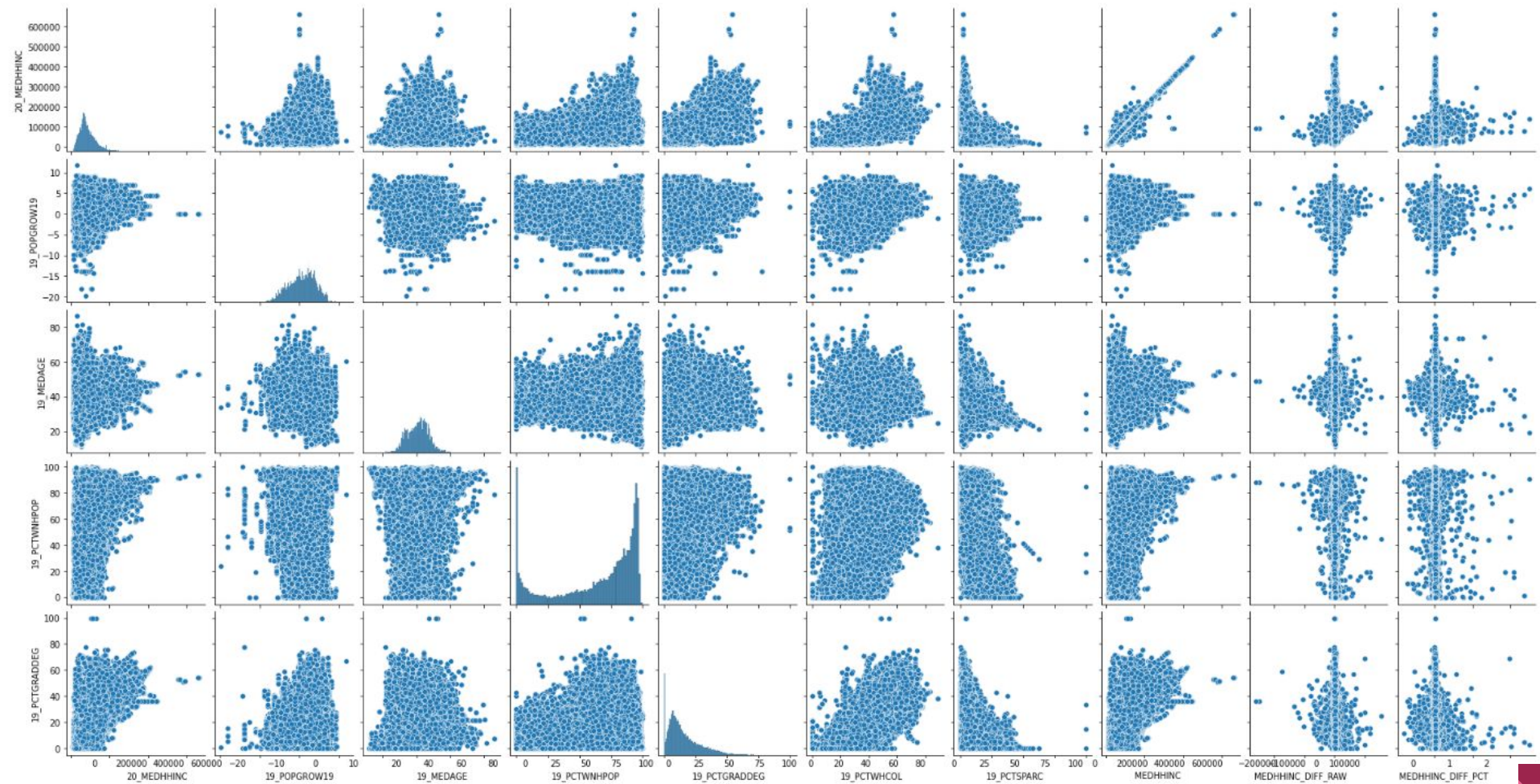- Some areas don't have a zip code

CRISP-DM Process

- Data understanding: EDA
- Data prep: Dropping pop zeroes, creating dependent variable
- Modeling: OLS, random forest, XGBoost
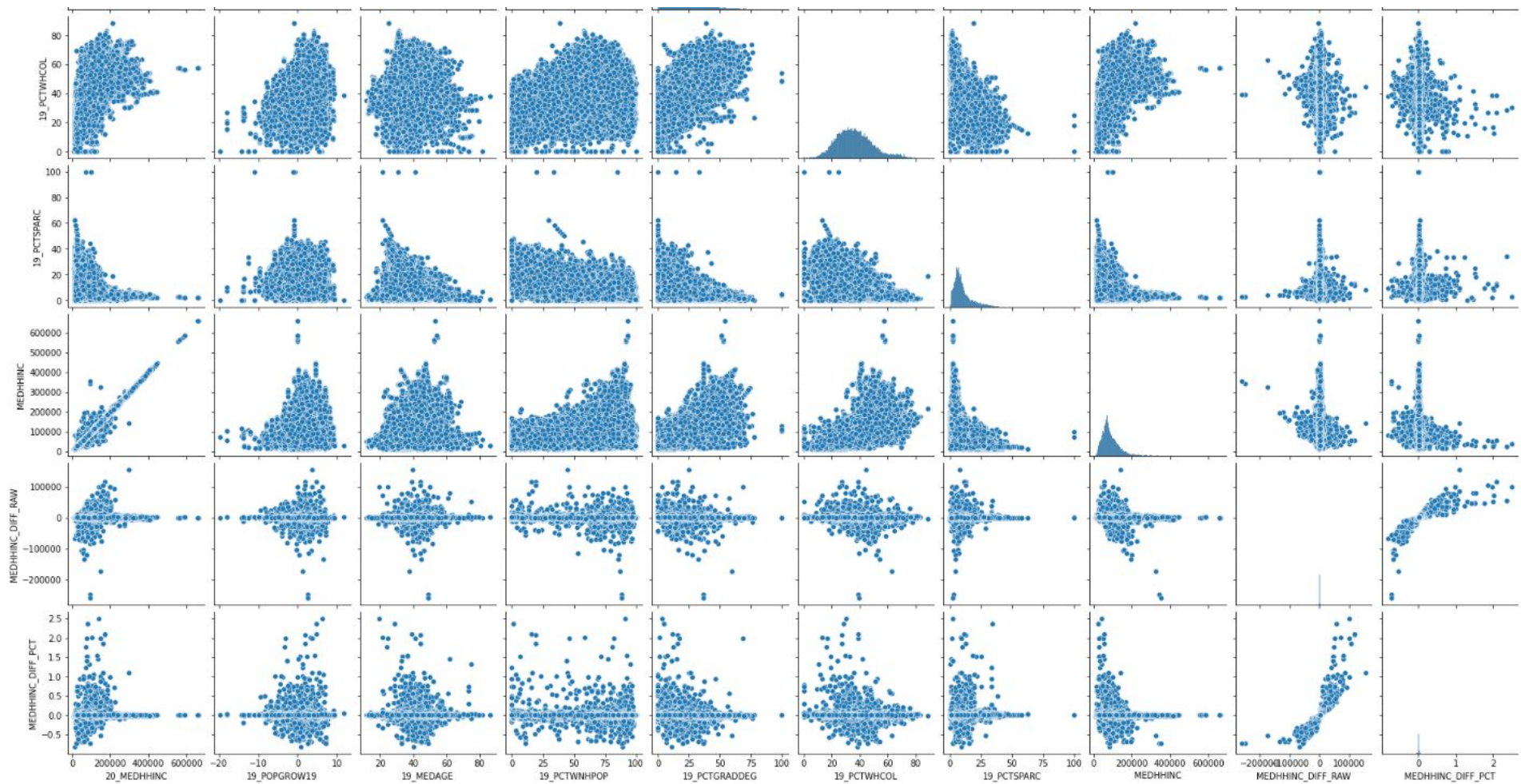- Evaluation
  - Accuracy
  - Shapley values

Noisy data
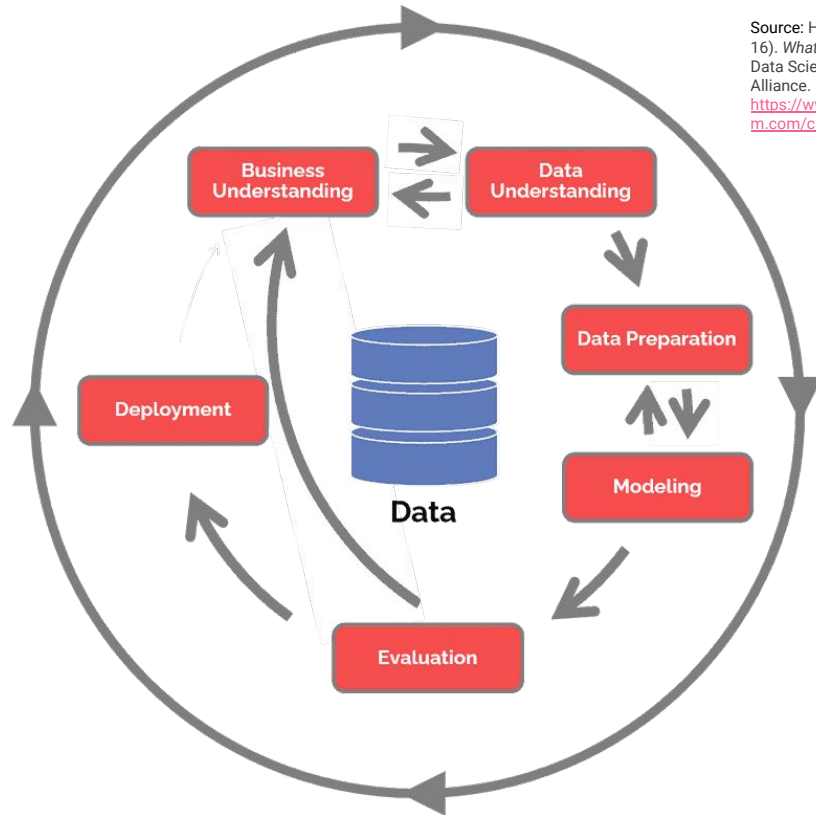
Chosen Dependent and Independent Variables

Dependent: Percent change in income from 2019-2020

Independent:

- Percent white non-Hispanic population
- Median age
- 2019 population growth
- Percent grad degree
- Percent white collar
- Percent single parents

Used two different samples of 20% each

Descriptive Stats: Fun Facts

	20_MEDHHINC	19_POPGROW19	19_MEDAGE	19_PCTWNHPOP	19_PCTGRADDEG	19_PCTWHCOL	19_PCTSPARC	MEDHHINC	MEDHHINC_DIFF_RAW	MEDHHINC_DIFF_PCT
count	1762317.00000	1762317.00000	1762317.00000	1762317.00000	1762317.00000	1762317.00000	1762317.00000	1762317.00000	1762317.00000	1762317.00000
mean	88023.99267	0.60932	40.96090	63.86083	15.51462	36.24046	10.30384	87453.60965	570.38302	0.00738
std	45614.64663	3.36928	7.11057	32.72703	12.27937	12.71037	7.25154	45450.10249	2240.09742	0.02936
min	9321.00000	-50.00000	0.00000	0.00000	0.00000	0.00000	0.00000	9366.00000	-259659.00000	-0.91175
25%	58246.00000	-1.76000	35.80000	42.89000	6.47000	27.13000	5.43000	57685.00000	113.00000	0.00151
50%	77853.00000	0.84000	41.40000	77.24000	11.98000	35.24000	8.29000	77362.00000	505.00000	0.00677
75%	109929.00000	3.11000	45.80000	91.04000	21.77000	43.93000	12.95000	109275.00000	952.00000	0.01239
max	660453.00000	11.76000	86.50000	100.00000	100.00000	88.71000	100.00000	661615.00000	155068.00000	6.00221

- Median age: 41.4 years
- Median white non-Hispanic pop.: 77.24%
- Median change in household income: .68%

- The most fortunate ZIP+4 increased their earnings by 6x
- The least fortunate lost over 91% of their income

Comparison of Feature Importance

OLS Regression

OLS Regression Results

Dep. Variable:	MEDHHINC_DIFF_PCT	R-squared (uncentered):	0.069
Model:	OLS	Adj. R-squared (uncentered):	0.069
Method:	Least Squares	F-statistic:	4369.
Date:	Wed, 29 Jun 2022	Prob (F-statistic):	0.00
Time:	01:14:27	Log-Likelihood:	7.5797e+05
No. Observations:	352463	AIC:	-1.516e+06
Df Residuals:	352457	BIC:	-1.516e+06
Df Model:	6		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
19_POPGROW19	0.0001	1.81e-05	8.244	0.000	0.000	0.000
19_MEDAGE	0.0003	5.31e-06	58.560	0.000	0.000	0.000
19_PCTWNHPOP	-5.243e-05	2.25e-06	-23.253	0.000	-5.68e-05	-4.8e-05
19_PCTGRADDEG	-8.583e-05	6.09e-06	-14.098	0.000	-9.78e-05	-7.39e-05
19_PCTWHCOL	-2.799e-05	5.53e-06	-5.067	0.000	-3.88e-05	-1.72e-05
19_PCTSPARC	1.245e-05	7.29e-06	1.707	0.088	-1.85e-06	2.67e-05

Omnibus:	844728.444	Durbin-Watson:	1.998
Prob(Omnibus):	0.000	Jarque-Bera (JB):	36908686119.365
Skew:	24.139	Prob(JB):	0.00
Kurtosis:	1587.571	Cond. No.	35.0

Random Forest
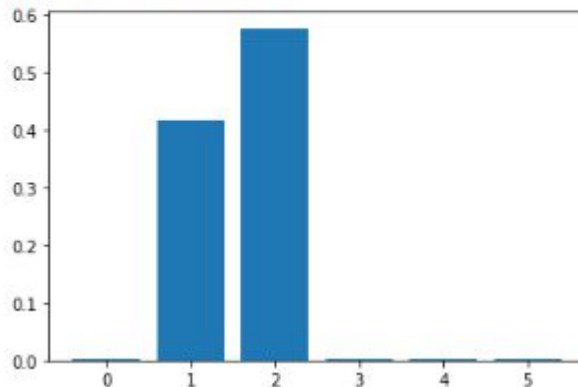
```
Feature: 0, Score: 0.00187
Feature: 1, Score: 0.41629
Feature: 2, Score: 0.57701
Feature: 3, Score: 0.00178
Feature: 4, Score: 0.00156
Feature: 5, Score: 0.00149
```
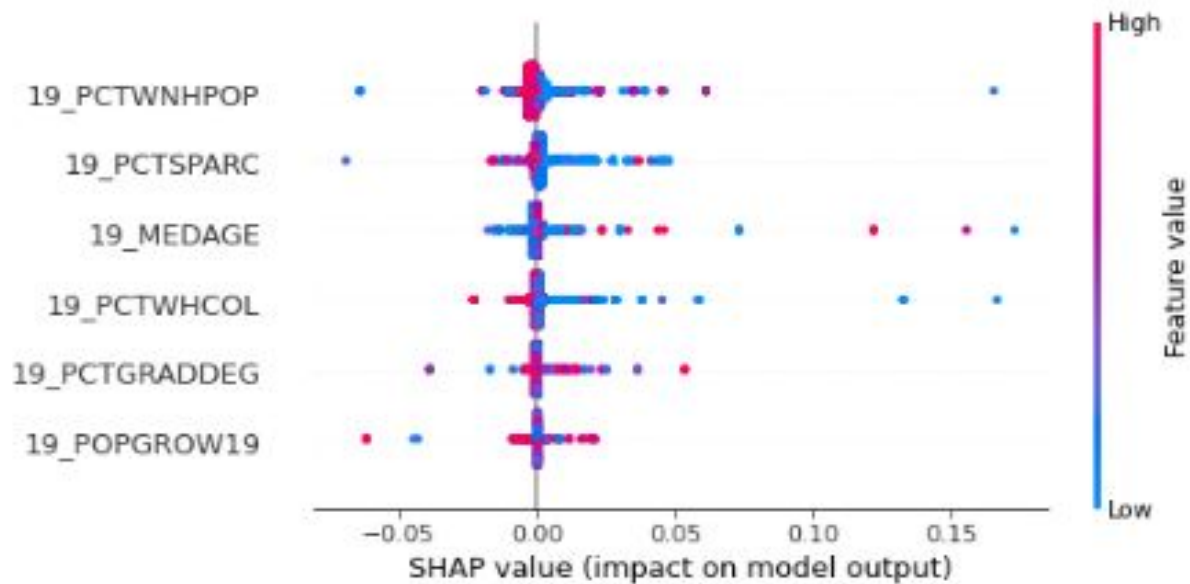

Median age and white non-Hispanic pop

Stepwise

```
['19_PCTWHCOL',
 '19_PCTWNHPOP',
 '19_PCTSPARC',
 '19_MEDAGE',
 '19_PCTGRADDEG',
 '19_POPGROW19']
```

XGBoost Shapley values

Conclusion

- The racial factor does appear to explain a little bit of the research question
- However, it appears that the majority race actually lost more income
- It would be great to have more years of data
- Age appeared to be a protective factor
- However, the high yearly income correlation makes it hard to tell

Speculation

- It is possible that majority households were in a better position to begin with, and thus had more to lose
- It may be more industry-dependent
- Age may be a protective factor because older people are more established, have a better network, and are eligible for social security

References

- Bhutta, N., Chang, A. C., Dettling, L. J., & Hsu, J. W. (2020, September 28). *Disparities in Wealth by Race and Ethnicity in the 2019 Survey of Consumer Finances*. Board of Governors of the Federal Reserve System - FEDS Notes. https://www.federalreserve.gov/econres/notes/feds-notes/disparities-in-wealth-by-race-and-ethnicity-in-the-2019-survey-of-consumer-finances-20200928.htm
- Hotz, N. (2022, April 16). *What is CRISP DM?* Data Science Process Alliance. https://www.datascience-pm.com/crisp-dm-2/
- Fazzari, S. M. & Needler, E. (2021, March 31). *US Employment Inequality in the Great Recession and the COVID-19 Pandemic*. Institute for New Economic Thinking. https://www.ineteconomics.org/uploads/papers/WP_154-Fazzari-Needler-Emp.pdf
- McKernan, M., Ratcliffe, C., Steuerle, E., & Zhang, S. (2014, April). *Impact of the Great Recession and Beyond: Disparities in Wealth Building by Generation and Race*. Urban Institute. https://www.urban.org/sites/default/files/alfresco/publication-pdfs/413102-Impact-of-the-Great-Recession-and-Beyond.PDF
- United States Zip Codes. https://www.unitedstateszipcodes.org/
- Wilson, V. (2020, September 16). Racial disparities in income and poverty remain largely unchanged amid strong income growth in 2019. *Economic Policy Institute Working Economics Blog*. https://www.epi.org/blog/racial-disparities-in-income-and-poverty-remain-largely-unchanged-amid-strong-income-growth-in-2019/