The Effect of Immigration Population on State Income Inequality

By Olivia Pan, The University of Texas at Austin

Abstract

This study investigates the relationship between the percentage of the immigration population and income inequality, measured by the Gini Index. A regression which considers many independent variables like median income, poverty rate, and education was conducted to examine the effect of the immigration population percentage on the Gini Index. The data collected covers all 51 states, and the results have passed all diagnostic tests. The results indicated that for every 1% increase in the immigration population, the Gini Index is predicted to increase by 0.152%, with a corresponding p-value of 0.00, which is statistically significant ($p < 0.05$). These findings suggest that an increase in the immigration population proportion is associated with higher income inequality in the United States. The government agencies can use the findings to implement further studies to explore policies that can help reduce income inequality.

Introduction

Income has always been an intrinsic factor to investigate in the Economics field. It provides the opportunity for scientists to study society's financial wellbeing and compare across countries. By using variance of income to study the overall distribution of wealth, we will be able to explain the range of the family income and understand the distribution of wealth. The economic terminology for variance in income is called income disparities.

Income disparities always draw attention from social activists and economists because of its social significance. Birdsall et al. (2012) argues that income inequality inhibits economic growth and slows poverty reduction, especially in developing countries. Also, Norris et al. (2015) claimed that higher inequality lowers growth by depriving the ability of lower-income households to stay healthy and accumulate physical and human capital. Other than that, inequality undermines the political process and inhibits collective decision-making at a societal level (Birdsall et al. 2012). Moreover, Smith et al. (1996) stated that income inequality within populations in the US has been an important determinant of population mortality: greater income inequality is associated with higher mortality from infant mortality, life expectancy at birth, and life expectancy at age 5. Unfortunately, the wealth gap in the US is widening: Keister et al. (2000) pointed out that the top 1% of wealth holders has owned an average of 30% of total household sector wealth in early 1920s, but the percentage of wealth held by top 1% has increased to 40% of total net worth and 50% of total financial assets in late 1980s and 1990s. It is urgent to find out potential causes for income inequality because we are facing worsening situations in health equality in which people's freedom to live out their life spans in the US is threatened. Also, without solving income inequality, poverty cannot be relieved, and there will be a constant, or worse, increasing amount of population in the US suffering from insufficient living conditions.

The wealth gap in the US is widening, so is the immigration population in the US. The immigration population in the US is huge and enlarging: Ward et al. (2023) found that there were approximately 45.3 million immigrants as of 2021, composing 13.6% of the total US population, and the share and number of immigrants are increasing rapidly. There is a large number of Americans who fear that the United States has more immigrants than the country can absorb and assimilate (Hirschman et al. 2015). Due to the astonishing growth in immigrant population, we are interested in the effect of immigrant population on income inequality. As a result, the paper tries to address the relationship between immigrant population and wealth inequality. Our hypothesis is: as the state-level immigrant population, which is the independent variable of interest, increases, the state-level wealth gap, which is the dependent variable of interest, increases in the US.

Literature Review

There is ambiguity on the effect of immigration on income inequality when looking at the published literature. First of all, Xu (2021) stated that foreign born population has a strong positive effect on state-level income inequality, and the state income inequality is driven primarily by low-skill immigrants. Also, Muhammad (2023) quoted that about 14.6% of immigrants were living below the poverty line in 2018, which is defined as an annual income of $25,701 for a family of four, compared to 11.8% of the national population. Looking at the other end of the wealth distribution, Keister (2017) argues that about 3% of households in the top 5 percent of wealth owners are of European or Canadian origin, and Cuban, Asian immigrants are well-represented among top wealth holders. For example, 0.65% of the top 1 percent are Mainland Chinese. Immigrants earning lower or higher income than the national mean income increases the income disparity in the US. Interestingly, Borjas proposed his theory that immigrants are self selected positively, which means the immigrants earn more than the average person in both home country and the US, when the inequality is greater in the US than the sending country, and negatively, meaning the immigrants earn less than the average person in both the home country and the US, when wealth inequality is less in the US (Berman et al. 2016). However, Liebig (2004) opposed Borjas and stated that according to other models based on the human capital theory of migration, migrants can be expected to be relatively skilled, earning the median income as other natives do. More research is necessary to estimate the effect immigration will bring to income inequality in the US.

Empirical Model

Linear regression is used to estimate the relationship between state income inequality and immigrant percentage. In this paper we estimate the following regression model:

$$Gini\ Index = a + \beta_1\, Pro100 + \beta_2\, Education + \beta_3\, MedianIncome + \beta_4\, PovertyRate + \beta_5 Region + \beta_6\, Black$$

In the model above, the dependent measure is the Gini Index which quantifies income inequality in each state.

Independent variables include Pro100 which is the percentage of foreign born population in each state, Education which captures proportion of the population with a bachelor's degree or higher in each state, MedianIncome which is the median income in households in dollars in every state, PovertyRate which captures the proportion of population with income less than

$14880 for 1 person, $18990 for 2 people, $23280 for 3 people, $29950 for 4 people. Black variable which is percentage of population being black in each state. By including the black variable in the regression, we are able to investigate if race population will affect the income inequality in the area. The independent variable, Region, is controlled; South is excluded for collinearity.

Data

The study uses publicly available state-level data from 2020 to 2023 because of limited availability of data of the same year for all variables. Data sources for the 2021 Gini Index is available at: Gini Coefficient 2021 | World Population Review. The Gini index is calculated as the ratio of the area between the perfect equality line and the Lorenz curve (A) divided by the total area under the perfect equality line (A+B). The Gini Index is the percentage form of the Gini Coefficient, and it expresses the expected absolute gap between people's incomes relative to the mean income of the population.

The Pro100, which is the percentage of foreign born population in thousands divided by population in thousands in each state in 2022, has sources: Statistia | Percentage of foreign born population in the US in 2022 by state.

The Education variable in 2022, which denoted the percentage of population having bachelor's degree or above, uses data from the Federal Reserve Bank of St. Louis, and it is available at: Educational Attainment or higher by State in 2022 | FRED | St. Louis Fed.

The MedianIncome in each state from 2023, in units of households, is obtained from Median Annual Household Income in 2023 | KFF.

The PovertyRate variable, which takes average from 2020 to 2022, has data from Mapped: Poverty Rate by US State.

The Region variable has source at Census Regions and Divisions of the United States.

Finally, the Black variable, which is the proportion of black people in each state in year 2023, has dataset at Black Demographics | Black Population by State.

Table 1 below presents descriptive statistics.

Table 1. Descriptive Statistics

Variable	Mean (st. dev)	Minimum value	Maximum value

Gini Index, %	47.039 (1.897)	43	53
Pro100, %	9.487 (6.107)	1.555	26.557
Education, %	35.241 (6.937)	24.8	65.4
MedianIncome ($)	69439.840 (11272.450)	48716	90203
PovertyRate, %	11.175 (2.917)	7.1	18.2
Black, %	12.137 (10.538)	1	44
West	0.255 (0.440)	0	1
South	0.333 (0.476)	0	1
MidWest	0.235 (0.428)	0	1
Northeast	0.176 (0.385)	0	1

Table 1 above shows a moderate variance in Gini Index across states. The Gini Index in the US varied from 43% in New Hampshire to 53% in the District of Columbia with an average of 47.039%. Pro100, which is the proportion of foreign born population, varied from 1.555% in West Virginia to 26.557% in California with an average of 9.487%. Education varied from 24.8% in Mississippi to 65.4% in the District of Columbia with an average college degree attainment proportion of 35.241%. The MedianIncome variable ranges from $48716 in Mississippi to $90203 in Maryland, with an average of $63439.840 in the US. The PovertyRate ranges from 7.1% in New Hampshire to 18.2% in New Mexico with an average of 11.175%. The

Black variable ranges from 1% in Idaho to 44% in the District of Columbia with an average of 12.137%.

Empirical Results

Regression results in Table 2 below show that immigrant proportion is an important determinant of state income disparity level. Higher immigration population proportion in a state will lead to higher Gini Index: when the proportion of foreign born population in a state increases by 1%, the Gini Index is expected to increase by 0.152% (p-value < 0.01).

Table 2. Regression Results

```
. reg giniindex pro100 education medianincome povertyrate west south midwest northeast black
note: south omitted because of collinearity.

      Source |       SS           df       MS      Number of obs   =        51
-------------+----------------------------------   F(8, 42)        =     14.83
       Model |  132.886599         8  16.6108248   Prob > F        =    0.0000
    Residual |    47.03497        42  1.11988024   R-squared       =    0.7386
-------------+----------------------------------   Adj R-squared   =    0.6888
       Total |  179.921569        50  3.59843137   Root MSE        =    1.0582

------------------------------------------------------------------------------
   giniindex | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
      pro100 |   .1521941   .0362479     4.20   0.000     .0790429    .2253454
   education |   .1763036   .0434776     4.06   0.000     .0885622    .2640449
medianincome |  -.0001099   .0000385    -2.85   0.007    -.0001877   -.0000321
 povertyrate |   .3221709   .1020953     3.16   0.003     .1161343    .5282076
        west |   .4681331    .621177     0.75   0.455    -.7854528    1.721719
       south |          0  (omitted)
     midwest |   .3391587   .5520245     0.61   0.542    -.7748719    1.453189
   northeast |   1.180807   .6205814     1.90   0.064    -.0715768    2.433191
       black |   .0368004   .0251379     1.46   0.151    -.0139299    .0875307
       _cons |   42.55898   2.687014    15.84   0.000     37.13637    47.98159
------------------------------------------------------------------------------
```

Other important determinants of Gini Index include Education, MedianIncome, and PovertyRate. When the percentage of people with bachelor's degree in a state increases by 1%, the Gini Index is expected to increase by 0.176% (p-value < 0.01); when the median income of family in a state increases by 1 dollar, the percentage of Gini Index is expected to decrease by 0.0001% (p-value < 0.01); when the poverty rate in the state increases by 1%, the Gini Index is expected to increase by 0.322% (p-value < 0.01). Because the p-value of West (0.455), MidWest (0.542), NorthEast (0.064), Black (0.151) are greater than the significance level of 0.01, they are not significant in the regression, meaning that Region variable and Black variable are not able to explain the Gini Index in the regression.

Additional Diagnostic Tests

Heteroskedasticity Test - Breusch Pagan Test: Success

```
. reg sls pro100 education medianincome povertyrate west south midwest northeast black
note: south omitted because of collinearity.

      Source |       SS           df       MS      Number of obs   =        51
-------------+----------------------------------   F(8, 42)        =      0.00
       Model |           0         8           0   Prob > F        =    1.0000
    Residual |  47.0349704        42  1.11988025   R-squared       =    0.0000
-------------+----------------------------------   Adj R-squared   =   -0.1905
       Total |  47.0349704        50  .940699407   Root MSE        =    1.0582

------------------------------------------------------------------------------
         sls | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
      pro100 |  -1.02e-09   .0362479    -0.00   1.000    -.0731513    .0731513
   education |  -5.97e-10   .0434776    -0.00   1.000    -.0877414    .0877414
medianincome |   6.67e-13   .0000385     0.00   1.000    -.0000778    .0000778
 povertyrate |   2.48e-09   .1020953     0.00   1.000    -.2060366    .2060366
        west |  -2.05e-08    .621177    -0.00   1.000    -1.253586    1.253586
       south |          0  (omitted)
     midwest |  -1.71e-08   .5520245    -0.00   1.000    -1.114031    1.114031
   northeast |  -9.59e-09   .6205814    -0.00   1.000    -1.252384    1.252384
       black |  -9.28e-10   .0251379    -0.00   1.000    -.0507303    .0507303
       _cons |  -1.64e-08   2.687014    -0.00   1.000    -5.422613    5.422613
------------------------------------------------------------------------------
```

With the null hypothesis that the coefficients on pro100, education, medianincome, povertyrate, west, south, midwest, northeast, black are 0, we obtained the P-value of 1.00, which is bigger than the significance level of 0.01, so we failed to reject homoscedasticity. We conclude that there is likely homoscedasticity in the residuals.

Multicollinearity-VIF: Success

```
. vif
```

Variable	VIF	1/VIF
medianincome	8.42	0.118736
education	4.06	0.246207
povertyrate	3.96	0.252585
west	3.34	0.299629
black	3.13	0.319199
northeast	2.55	0.392331
midwest	2.50	0.400478
pro100	2.19	0.457108
Mean VIF	3.77	

The next test, VIF (Variance Inflator Test), applied above is for checking collinearity. The mean VIF for the model is 3.77, which is a value lower than 10. Since there is no variable with VIF over 10, the model does pass the test, and there is no multicollinearity in the variables.

Normality: Success

```
. sktest giniindex pro100 education medianincome povertyrate west south midwest northeast black
```

Skewness and kurtosis tests for normality

Variable	Obs	Pr(skewness)	Pr(kurtosis)	Joint test Adj chi2(2)	Joint test Prob>chi2
giniindex	51	0.1000	0.1774	4.57	0.1019
pro100	51	0.0067	0.6736	6.83	0.0328
education	51	0.0000	0.0002	22.57	0.0000
medianincome	51	0.4047	0.1764	2.66	0.2643
povertyrate	51	0.0292	0.6792	4.88	0.0873
west	51	0.0017	0.1983	9.64	0.0081
south	51	0.0327	0.0000	34.09	0.0000
midwest	51	0.0007	0.6427	9.81	0.0074
northeast	51	0.0000	0.1238	15.52	0.0004
black	51	0.0010	0.1495	10.60	0.0050

For this normality test to pass, the dependent variable should be normally distributed. In this model, the Prob>chi2 value for ginicoefficient is 0.1019, which is bigger than 5%, so we fail to reject the null hypothesis, meaning the distribution for ginicoefficient is normally distributed. The normality test is passed.

Ramsey Test:

```
. ovtest

Ramsey RESET test for omitted variables
Omitted: Powers of fitted values of ginicoefficient

H0: Model has no omitted variables

F(3, 40) =    0.85
Prob > F = 0.4755
```

The Ramsey test helps analyze whether nonlinear combinations of the fitted values may help explain the dependent variable, ginicoefficient. Since the p-value, 0.4755, is bigger than the significance level 0.05, we fail to reject the null hypothesis that there is no omitted variable bias in the regression, and we reach the conclusion that there is no omitted variable bias here in the regression.

Conclusions and Policy Implications

Empirical results show that increasing immigration will cause a bigger wealth gap in the US. Our regression has passed the Breuch Pagan test, VIF test, Normality test, and Ramsey test. This positive relationship between immigration and income inequality may be caused by the potential fact that immigrant families have incomes that fall on the extreme sides of the metrics, meaning immigrant families are likely to have incomes that are higher or lower than the median income in the US. More research is needed to find out the reasons. The result is consistent with the hypothesis proposed, and our estimates are consistent with what Xu (2021) stated, that foreign born populations have a strong and positive effect on state-level inequality. Therefore, suggesting the government to implement free job skill training for immigrant families with incomes that are lower than the median income should be considered an important target for social activists. By implementing job skill training that is free, immigrants from a lower income background will have an opportunity to gain skills and get paid more. By improving the lower bound of the income scale, the income inequality should be somehow relieved. Also, offering job training programs will bring more job opportunities to other qualified job-seekers, which will decrease the unemployment rate. When immigrant families are employed and paid more, the overall disposable income in immigrant families increases, and GDP on a national scale, which is partially based on consumption, will increase. However, the sustainability of the program will be questioned because running the program will be costly, and the government has to adjust its spending to save enough funds for the program to run. In the long run, with the development of technology, the skills being taught in the program might not bring a positive effect on the wage, which nullifies the purpose of the program.

Limitations

Results of this study are not without limitations. First, the dataset being investigated is cross sectional, meaning that there is time series data in the dataset. As a result, difference-in-difference tests and autocorrelation tests are not able to be implemented. Also, the income, education, and language abilities of immigrants are not taken into account in the study, so we are not able to study the effect of immigrants of different levels of skills and education on the overall income inequality in the US. Other than that, we collected data from the year 2020 to 2023, when Covid occurred. The overall income level in the US has been altered because many workers lost their jobs, and people’s living styles have changed. Picking another set of data from years before the pandemic will be more considerable. Lastly, there is a potential possibility that income inequality in the US drives immigration. More research is needed to further explore the relationship between immigration population and income inequality.

Reference

Muhammad, D. A. (2023, August 22). *Racial wealth snapshot: Immigration and the racial wealth divide " NCRC*. National Community Reinvestment Coalition. https://ncrc.org/racial-wealth-snapshot-immigration-and-the-racial-wealth-divide/

Berman, Y. (2016). *Explore scientific, technical, and medical research on ScienceDirect*. ScienceDirect.com | Science, health and medical journals, full text articles and books. https://www.sciencedirect.com/

Keister, L. A., & Aronson, B. (2017). *Immigrants in the One Percent: The national origin of top wealth owners*. PLOS ONE. https://journals.plos.org/plosone/article?id=10.1371%2Fjournal.pone.0172876

Birdsall, N. (2012, September 28). *Why inequality matters: Some economic issues: Ethics & International Affairs*. Cambridge Core. https://www.cambridge.org/core/journals/ethics-and-international-affairs/article/abs/why-inequality-matters-some-economic-issues/6B4D6DC238306AF50F5003BADE816E64

Smith, G. D. (1996, April 20). *Income inequality and mortality: Why are they related?*. The BMJ. https://www.bmj.com/content/312/7037/987.short

Hirschman, C. (2014). *Immigration to the United States: Recent trends and future prospects*. Malaysian journal of economic studies : journal of the Malaysian Economic Association and the Faculty of Economics and Administration, University of Malaya. https://pmc.ncbi.nlm.nih.gov/articles/PMC4302270/

Ward, N. (2023). *Frequently Requested Statistics on Immigrants and Immigration in the United States*. The Italians in America. https://www.americacallsitaly.org/emigrazioni/italian.htm

Liebig, T. (2004). *Migration, Self-Selection and Income Inequality: An International Analysis*. Wiley Online Library. https://www.researchgate.net/publication/228751641_Export_performance_measurement_An_evaluation_of_the_empirical_research_in_the_literature

Keister, L. A., & Moller, S. (2000, August 1). *Wealth inequality in the United States*. Annual Review of Sociology. https://www.annualreviews.org/content/journals/10.1146/annurev.soc.26.1.63