Asset Allocation: Variations among Racial Groups

De Hua Wang[a], Muhammad Arham Malik[b]

[a]1009770650 dehua.wang@mail.utoronto.ca

[b]1007009620 arham.malik@mail.utoronto.ca

ECO 375: Applied Econometrics I

University of Toronto

Department of Economics

Research Paper

30th November 2022

Abstract

Income and wealth inequality has been growing substantially in the past few decades. According to the distributional financial accounts of the Federal Reserve of the United States, the difference in the share of total wealth held by the top 1% and bottom 90% has widened drastically. Similar inequality and wealth gap are persistently rising in racial groups as well, where certain racial groups tend to be more wealthy than others. While many researchers have studied income inequality, this paper investigates the role of racial factors as reasons for the difference in asset allocation. Our evidence and analysis suggest that non-white people allocate and invest their resources much differently than white people, which can be seen in their equity share as a percentage of total assets. Given the same resources, we find an economically and statistically significant difference in asset allocation among both racial groups.

1 Introduction

Income and wealth inequality has grown uninterruptedly since 1980. According to the distributional financial accounts of the Federal Reserve, the share of total wealth held by the top 1% of the U.S. population increased from 22.7% in 1989 to 32.3% in 2022. On the other side, the share of the bottom 90% decreased from 40.1% to 30.6%. These results are striking because of their effect on economic growth, and social and political stability. This has caused concerns among researchers, policymakers, and politicians. Endless factors are the causes of this inequality, and numerous studies have been conducted. Our paper focuses on asset allocation as a driver of inequality and how asset allocation differs among different racial groups.

The paper we took inspiration from is "Wealth Stratification and Portfolio Choice" by Boulware and Kuttner. The authors explore the possible role of asset allocation as a contributor to the widening racial wealth gap of the past decade. They define a new metric of wealth inequality and analyze to which extent households' portfolio allocations depend on the type of assets and racial groups. We build upon their analysis and investigate the role of racial groups as an explanatory variable of equity share, holding other variables constant.

The "widening racial wealth gap over the past few decades, despite an increase in GDP and the lower unemployment rate" *(Boulware and Kenneth 2020)* is a concern for economists, and policy makers, as the existence of such disparities leads to further inequality in the economy. We aim to explore the role of assets in the form of equity share as a plausible contributor to the widening racial wealth gap.

While conducting our analysis, we pay attention to the rigorous theoretical framework behind the econometric analysis. Our analysis starts with a simple linear regression using only racial groups as a regressor. Then we include education and income as control variables to expand to multiple linear regression.

2 The Context and Data

The data we use for our model and paper comes from the Survey of Consumer Finances collected in 2007. The data contains information on various aspects of a respondee, such as their racial group, income level, education, family size, wealth proportions, etc. Each respondent in our data is a person from the United States currently employed and working.

The data reports different variables, all measured differently, which is why we have tried to report them in a way that makes it easier to understand (*Table 1*). Here our descriptive statistics do contain significant outliers. Still, we do not violate *LSA 3* [1] as we have finite kurtosis. Nevertheless, significant outliers can be seen in our initial results as our mean income is drastically affected due to the existence of a few very rich people. For better readability, we have included the median (50th Percentile) for such instances. Since the error term in our data is very likely to be heteroskedastic, all the regressions in the paper are conducted using heteroscedasticity-robust standard errors. This is common practice given their validity, whether or not the errors are heteroscedastic.

Some key pointers for our analysis later are as follows:

- We notice education [educ] goes from -1 to 14, where 12 represents bachelor, 14 represents Ph.D. and higher, while -1 represents not going to school.
- Equity share lies between 0-100 as it represents a percentage.
- All other variables are measured in thousands of dollars.
- We include weights in our model for our data to represent the population better.

Our descriptive statistics table (*Table 1*) gives us a basic understanding of how our key explanatory variables work. As mentioned before, the mean income, which is highly affected by outliers, is 1.31 million dollars for the 22,085 respondents. The high standard deviation indicates the existence of outliers with large incomes that skew our data. As

[1] LSA 3 mainly states the existence of large outliers in explanatory variables (X) and outcome variables (Y) are rare which in our case is true as we have a few individuals with large income. Mathematically we say that X and Y have non-zero finite fourth moments

a result, we report the median income for a better understanding of our model, being \$83000. Education which is represented on a scale of -1 to 14, is on average $9.9 \approx 10$, which means that our average respondent has graduated high school and not pursued further education. The standard deviation in this context is also pretty high, mainly because many of our respondents have a bachelor's degree or more. In contrast, many did not attend high school in general. Equity share, which represents the percent invested in equity for a respondent, is 13.1% on average, while wage income and debt are approximately \$44,000 and \$46,000 respectively on average as per the median.

3 Regression analysis

3.1 Simple Linear Regression

Using our data from the survey extract from 2007, we analyzed the effect of race on equity share. Here race is a dummy variable representing 1 if the respondent to our survey is non-white and zero if they are white. At the same time, equityshare is the percentage of stocks held in various accounts as a percentage of total assets. Our simple model holds for *LSA 2*[2]: as all respondents were independently chosen and not based on any education level, geographic differences, etc.

In our initial analysis, we ran a simple linear regression on our outcome variable (equityshare) using our explanatory variable "race".

$$equityshare_i = \beta_0 + \beta_1 race_i + \mu_i \tag{1}$$

To test our hypothesis of whether there is a difference in average equityshare among racial groups, we conduct an analysis on the following to see if the coeffecient of our regressor race is different:

$$H_0 : \beta_1 \ = 0 \ vs. \ H_1 : \ \beta_1 \neq 0 \tag{2}$$

The results were intriguing yet not surprising and can be seen in *Table 2 Spec (1)*:

[2]LSA 2 states that (Xi, Yi), i = 1,...,n, are independently and identically distributed (i.i.d) across observations which is true as each respondent was chosen and analyzed independently.

- The slope coefficient on our explanatory variable 'race' is -4.98, significant at a 99% confidence level.

- The coefficient of our constant is 10.63.

We can also see our model below:

$$\hat{equityshare_i} = 10.63 - 4.98\hat{race_i} \tag{3}$$

This means that on average, equity share of non-white people was approximately 4.98% less than white people. Our constant of 10.63 can be interpreted as follows: among our survey respondents, those that identified as racially white had an equity share of 10.63% on average. In contrast, those who were non-white had an average equity share of 5.65%. This divergence in the percentage of equity share is surprising as it gives us an understanding of how differently white and non-white people allocate their assets.

Though our analysis is interesting and shows a clear difference in how different racial groups allocate their resources, we may violate *LSA 1*[3]: the conditional mean of the error term might not be zero due to omitted variable bias. There lie possibilities that "our model generates significant racial gaps in equity share because white and non-white have very different wages" *(Boerma 2022)* and/or differences in financial literacy. As a result, we need to consider other variables, such as education and income level, to see whether these findings are consistent after controlling for these variables.

3.2 Multiple Linear Regression

The results from the simple linear regression are likely to be affected by omitted variable bias. For this reason, we investigate whether race is still a significant explanatory variable for differences in equity share after controlling for other variables.

- As described in the introduction, we add *bachelor_degree* to control for education and *min_income* to control for income. Education can be significant because of its role in financial literacy. Income can have explanatory power because richer people will have more excess funds and the necessity to allocate these funds efficiently.

[3]LSA 1 states that the conditional distribution of u_i given X_i has a mean of zero

- Both education and income are dummy variables, *bachelor_degree* is equal to 1 if the respondent has a bachelor's degree or higher and 0 otherwise, and *min_income* is equal to 1 if the respondent has an income higher than $100,000 and 0 otherwise.

$$equityshare_i = \beta_0 + \beta_1 race_i + \beta_2 bachelor\_degree_i + \beta_3 min\_income_i + \mu_i \tag{4}$$

In this case, the null hypothesis (5) and the alternative hypothesis (6) are:

$$H_0 : \beta_1 = 0 \ and \ \beta_2 = 0 \ and \ \beta_3 = 0 \tag{5}$$

$$H_1 : \ \beta_1 \neq 0 \ and/or \ \beta_2 \neq 0 \ and/or \ \beta_3 \neq 0 \tag{6}$$

However, as discussed later, we only care about β_1 being different than 0. Therefore, our null hypothesis remains the same as in *equation 2*.

As seen in *Table 3*, education and income are highly correlated. This induces imperfect multicollinearity that can significantly affect the coefficient of these two regressors. This effect can also be seen in *Figure 1* where higher education indicates higher income; interestingly, we notice that there is a more significant portion of white people with higher education levels. Since both variables can potentially reduce omitted variable bias, we include them as control variables. In this way, we shift our focus to the effect that the regressor race, our variable of interest, has on equity share. Even after including these control variables, we still do not violate *LSA 4* [4] as they are not perfectly correlated. Education and income, being control variables, are not the object of interest in our study, and the coefficient of these regressors will not be interpretable because they are biased [5] and do not have a causal interpretation. This leads to a modification of *LSA 1*. By considering the race variable as X_{1i}, education as W_{1i} and income as W_{2i}, the first least square assumption for causal inference with control variables states that the error u_i has conditional mean that does not depend on the X's given the W's, that is:

$$E(u_i|X_{1i}, ..., X_{Ki}, W_{1i}, ..., W_{ri}) = E(u_i|W_{1i}, ..., W_{ri}) \tag{7}$$

[4] LSA 4 states that we do not have multiple explanatory variables that are the same or that are a perfect linear function of other variables,, in other words Var$(X) \neq 0$.

[5] If a certain variable Z is a determinant of Y and correlated with X, $\lim_{n\to\infty} \hat{\beta}_1 = \beta_1 + (\frac{\sigma_u}{\sigma_X})\rho_{Xu}$

Our multiple regression model with control variables is:

$$\widehat{equityshare}_i = 6.84 - 3.81\widehat{race}_i + 6.97\widehat{bachelor\_degree}_i + 5.55\widehat{min\_income}_i \quad (8)$$

The results from the multiple linear regression in *Table 2 Specification (3)* show that the constant β_0 goes from 10.63 to 6.84 and the coefficient of the race regressor from -4.98 to -3.81. The coefficient of the *bachelor_degree* regressor is 6.97, and *min_income* is 5.55. However, as stated before, these two coefficients do not have a causal interpretation. The coefficient of the race regressor has a standard error of 0.25 and a t-statistic of -15.02, meaning that it is significant at a 99% confidence level. Moreover, since the difference between the 1% and 50% percentile for equityshare is 4.22% (*Table 1*), the race regressor is also economically significant as a 3.81% change is close to 4.22%. The reduction in the magnitude of the coefficients of the constant and of β_1 are evidence of a decrease in omitted variable bias. The interpretation of these results is the following: controlling for education and income level, belonging to a non-white racial group still leads to an average difference of 3.81% in equity share. Therefore, despite controlling for these variables, non-white people still allocate and invest their resources differently than white people. Past research indicates that even "if successful policies were designed to reduce racial differences in income and asset allocation, three-quarters of the wealth gap would still remain" *(Blau and John 1990)* which seconds our finding that white and non-white people allocate funds differently. These findings are truly defining as they give us an understanding of how different investment portfolios and asset allocations can be when considering race as a factor.

4 Limitations of results

After conducting our analysis, we notice that our results come with caveats. First, even after controlling for the mentioned variables, we still need to consider geographic differences and how they may affect variations. This is mainly because our survey data aims to portray the asset allocation of people all over the US without considering if they live in a metropolitan or suburban/rural town which may affect our results. Second, our large outliers, though portray a realistic depiction of real-world wealth stratification, also substantially affect our results when extracting descriptive statistics and in our regressions which is why we have tried to use medians and other statistical methods to improve the readability for our research. The coefficients of *bachelor_degree* and *min_income* are statistically significant to a 1% significance level. However, they can not be classified as economically significant as they are highly correlated, making them uninterpretable.

5 Conclusion

Our analysis assessed whether racial differences substantially affect how different racial groups invest their resources in equity. Using our regression models and conducting empirical research in our basic and multiple regression models, we see how race, after controlling for other factors, still plays a vital role in how different people allocate and invest their resources. These results are surprising as a four percentage point difference in equity share is statistically and economically significant. This makes us ponder that apart from income levels and financial literacy, racial groups may have natural, cultural, and historical differences in why they allocate their resources differently. These results are interesting especially for policy makers and economists who could use such finding to devise policies to migitate and reduce wealth gap in the United States.

References

Blau, Francine D., and John W. Graham. “Black-White Differences in Wealth and Asset Composition.” The Quarterly Journal of Economics 105, no. 2 (1990): 321–39. https://doi.org/10.2307/2937789.

Boerma, Job, Karabarbounis, Loukas, and Federal Reserve Bank of Minneapolis. Research Department. Reparations and Persistent Racial Wealth Gaps. Reports. Federal Reserve Bank of Minneapolis, 2022. https://jstor.org/stable/community.34096082.

Boulware, Karl David, and Kenneth N. Kuttner. “Wealth Stratification and Portfolio Choice.” AEA Papers and Proceedings 110 (2020): 411–15. https://doi.org/10.1257/pandp.20201111.

Appendix

Table 1: Summary statistics

	Mean	St dev	Min	Max	Median
income	1312.457	7016.582	0	216342.7	83.35988
educ	9.903509	2.970996	-1	14	9
equityshare	13.12361	18.69435	0	100	4.220507
wageinc	192.5105	1101.279	0	35725.66	44.06165
debt	331.5532	1741.879	0	69186.77	46.3206
N	22085				

Table 2: Regression table

	(1)	(2)	(3)
	equityshare	equityshare	equityshare
race	-4.983***	-4.090***	-3.810***
	(0.258)	(0.257)	(0.254)
min_income		8.188***	5.555***
		(0.280)	(0.301)
bachelors_degree			6.971***
			(0.308)
_cons	10.63***	8.315***	6.839***
	(0.149)	(0.170)	(0.169)
N	21699	21699	21699
adj. R^2	0.018	0.066	0.100

Note: The unit of observation is a respondent to the Survey of Consumer Finances. Standard errors in parenthesess

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: Regression table

	(1)
	income
educ	9.660^{***}
	(0.156)
_cons	-13.92^{***}
	(1.286)
N	17940
adj. R^2	0.166

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

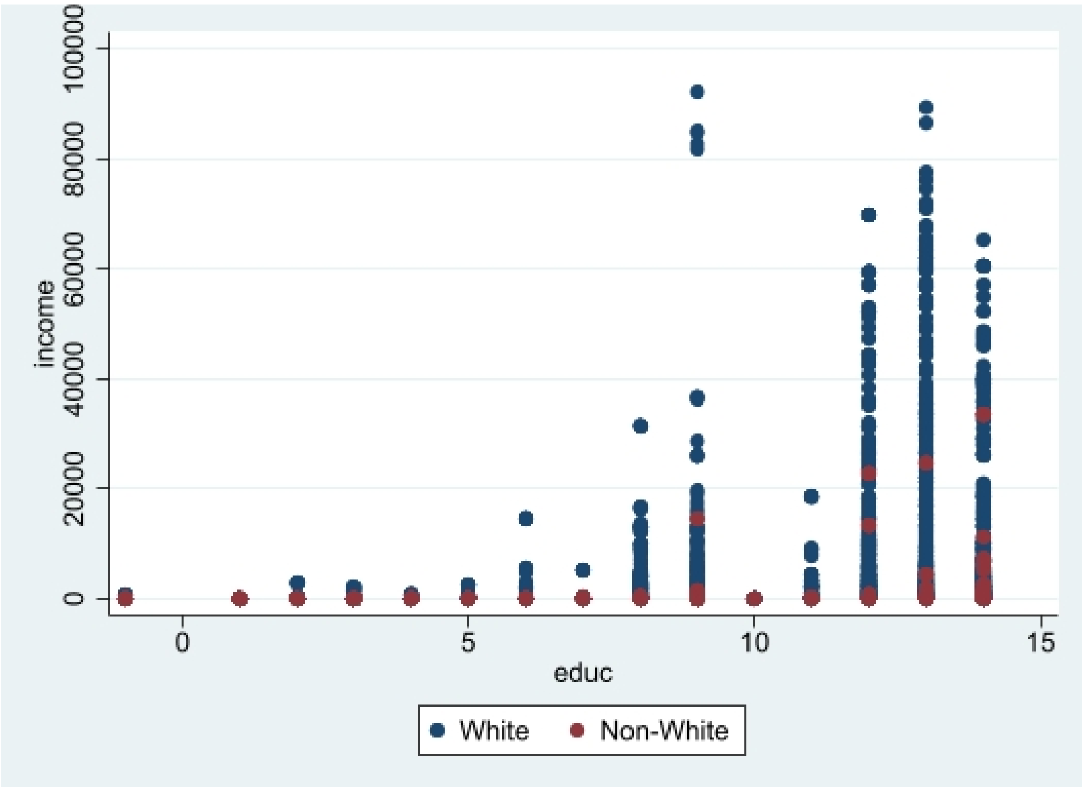

Figure 1: Shows the relationship between income and education among different racial groups. It is important to note that 12 on the x-axis (education) refers to bachelor degree which is the basis for one of our dummy variable *bachelor_degree*.