STOR 390 Final Paper

Serena Zhang

2024-04-17

Introduction

Physical appearance plays a critical role in how people are perceived and treated by others. The body of empirical literature investigating physical appearance and its potential outcomes for individuals is vast – in particular, psychologists and other researchers have investigated the "beautiful is good" phenomenon: one's physical appearance may lead others to assume certain qualities about them, and beauty is assumed to be correlated with positive traits (Dion, Berscheid, et. al., 1972). Therefore, conventionally attractive people may experience certain benefits from their appearance, and less attractive people may experience certain consequences. In popular terminology, this phenomenon has been coined as "pretty privilege".

One area of influence beauty may play a role in is in labor market outcomes. Previous economic and psychological research on this topic has generally shown that people receive economic premiums for beauty. Roszell, Kennedy, and Grabb (1989) investigated the relationship between physical attractiveness and income attainment among the Canadian working population. In this paper, I analyze the statistical demerits of their methodology. I then replicate their study on panel data from an American sample and provide my own analysis of the data. Finally, I discuss the normative consideration of the paper from a utilitarian perspective.

Analysis of Methods

The researchers used Canadian panel data containing information about respondents' *attractiveness*, *income (1979)*, *income (1981)*, *education*, *occupational status*, *socioeconomic status*, *gender*, *gender job composition*, and *age*. The sample size for this study was 1,062 respondents and was generally representative of the Canadian population. Attractiveness was ranked by respondents' interviewers on a 5-point scale, ranging from homely (1) to strikingly handsome or beautiful (5).

The researchers first reported Pearson correlations for *income (1981)* and each of the predictors, as well as *attractiveness* and each of the predictors. They then performed a regression analysis of *income (1981)* on *attractiveness*, controlling for the other predictors. In order to test for interactions between *attractiveness* and some of the control variables, they then performed the same multiple regression on subsets of the sample.

In order to attempt to reproduce the authors' study, I used data from the Wisconsin Longitudinal Study (WLS). WLS data comprises a random sample of 10,317 people who

graduated high school in Wisconsin in 1957, and were repeatedly surveyed until 2011. The data is extremely extensive and covers almost every aspect of the respondents' lives over time (with over 24,000 variables), including physical appearance ratings and income data. Physical appearance was rated on a scale of 1 ("not at all attractive") to 11 ("extremely attractive") based on yearbook photos of the respondents. Benchmark photos were provided on the scale to provide judges with a standard by which to rate respondents. Photos were also in black and white to control for any influences that clothing color or background could have on ratings. Ratings were given by six men and six women who were from around the same age cohort as the respondents; they had an average age of 78.5 when ratings were conducted in 2004 and 2008. These ratings were then "normalized" to remove coder fixed effects by subtracting each judge's mean rating from their ratings, and all judges' centered ratings for an individual were averaged to provide the attractiveness rating for that individual. The normalized scale ranged from -4 to 4. After removing null and invalid responses for each of the variables I included in my replication, I was left with a sample size of n = 3,907.

A clear advantage of this data are the more rigid precautions that were put in place to ensure the integrity of ratings. However, it does have some shortcomings. The data inherently does not contain variation in age between respondents, since they were sampled from one graduating year. It also does not report gender compositions for the respondents' jobs. Therefore, those variables could not be included in my regressions. Otherwise, the predictors I used were *attractiveness*, *income (1974)*, *education*, *occupational status* (a measure of occupational prestige), *family income* (parental income, an indicator for socioeconomic status), and *gender*. The response variable was *income (1992)*.

I first conducted a simple regression of *income (1992)* on *attractiveness*, which found that for a single unit increase on the normalized attractiveness scale, respondents' income increased by $1,593.36 (Roszell et. al reported $1,988 in their simple regression). Here, *attractiveness* was a significant predictor at $\alpha = 0.01$. The researchers also found that *attractiveness* was significant in their multiple regression; however, in my multiple regression of *income (1992)* on all predictors, *attractiveness* was not a significant predictor (see Table 1). All other predictors except for *family income* were significant.

The researchers then checked for interactions by rerunning the multiple regression on subsets of the data. In order to test for interactions between *gender* and *attractiveness*, I subsetted the data into male and female subsets and ran the preivous regression on each. Unlike Roszell et. al's findings that attractiveness was only significant for male respondents, attractiveness was not significant in either regression (see Table 2).

The authors' main conclusion generally reflects those of other literature on this topic, in that attractiveness is positively associated with income. However, there are also some statistical issues with the authors' methodology. The interviewers tested for interactions between attractiveness and some of the covariates by performing the same multiple regression on subsets of the sample (e.g., regressing with female and male subsets). In this scenario, gender is essentially interacting with every predictor; therefore, the interaction effect between attractiveness and gender cannot be isolated. It is more appropriate to include an interaction term in the multiple regression using the full sample. In addition, the authors did not transform income and reported their results in the form of dollar amounts. It would have been more appropriate to log transform income variables to make their distributions more

Table 1: Simple and multiple regression following Roszell et. al's methodology

	Dependent variable:	
	income_1992	
	Simple regression	Multiple regression
	(1)	(2)
income_1974		1.920*** (0.057)
attractiveness	1,593.360*** (514.596)	292.961 (394.726)
education		1,945.596*** (262.029)
occ_status		2.527*** (0.310)
gender		−5,003.459*** (1,189.971)
famincome		0.114 (0.086)
Constant	42,653.880*** (632.185)	−18,829.560*** (3,043.259)
Observations	3,910	3,910
R^2	0.002	0.418
Adjusted R^2	0.002	0.417
Residual Std. Error	39,526.860 (df = 3908)	30,200.910 (df = 3903)
F Statistic	9.587*** (df = 1; 3908)	467.938*** (df = 6; 3903)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 2: Regressing with gender subsets

	Dependent variable:	
	income_1992	
	Female subset	Male subset
	(1)	(2)
income_1974	1.407*** (0.102)	1.952*** (0.072)
attractiveness	326.499 (336.666)	346.046 (617.316)
education	1,174.389*** (251.985)	2,370.696*** (383.778)
occ_status	2.382*** (0.288)	2.647*** (0.460)
famincome	0.033 (0.089)	0.143 (0.121)
Constant	−9,499.002*** (2,790.811)	−26,151.490*** (4,374.553)
Observations	1,516	2,394
R^2	0.273	0.346
Adjusted R^2	0.271	0.345
Residual Std. Error	16,506.540 (df = 1510)	36,181.090 (df = 2388)
F Statistic	113.443*** (df = 5; 1510)	252.580*** (df = 5; 2388)
Note:		*p<0.1; **p<0.05; ***p<0.01

normal given their extreme right skew. Then, the effect of attractiveness on income would be reported in terms of percent change, which is potentially a more useful interpretation than raw dollar amount since an increase of $1,988 in income per unit increase in attractiveness is valued differently for people with lower versus higher incomes.

Another limitation is the way in which attractiveness was measured. Attractiveness ratings were given by a few interviewers of the subjects, and interviewers varied among subjects (in total, there were hundreds of interviewers). Attractiveness ratings were therefore based on, at most, three people's perceptions of the subject and do not necessarily represent how the subjects would be perceived on average by the population. There was no system for controlling for differences between interviewers, as a variety of factors could affect how attractive an interviewer found a subject to be based on personal preferences (e.g., there were no benchmarks provided for the level of attractiveness each step on the scale would represent). A clear problem also exists in that the interviewers interacted with subjects in person, and external factors such as personality traits, clothing, and setting of the interview could have influenced their ratings.

One final statistical concern with the authors' methodology is potential omitted variable bias. The authors do not control for certain factors that are commonly controlled for in similar studies and potentially related to both income and attractiveness (Hamermesh and Abrevaya, (2013); Frieze et. al, (1991); Scholz and Sicinski, (2015)). The authors also do not include skill in their regression analysis, which is a predictor of income and could be related to attractiveness (e.g., if more skilled workers hypothetically valued physical appearance less and spent less effort on their appearances). A proxy variable for skill such as years of work experience and/or IQ would ensure that any perceived relationship between physical attractiveness and income is not explained in part by skill. They also do not include a control for race, which is an appearance-based factor that influences income attainment and could be related to subjective measures of attractiveness. As racial discrimination is one of the most prevalent forms of discrimination in the labor market, it is important to ensure that any effects of race are isolated from "appearance" as a whole. Of course, the authors were limited in that they were working within the confines of data previously collected by another organization; however, they also do not explain their model selection process and why they chose to include the covariates they used. The authors' finding that physical appearance is positively related to income may not necessarily be wrong, but without controlling for these factors, it is not guaranteed that other variables are not confounding the perceived link between attractiveness and income. Therefore, the authors' findings may not be as reliable or accurate as the findings of other studies that came to similar conclusions but included other potential confounders in their regression analyses.

Due to these issues with the authors' methodology, I also conducted some analyses of my own on the WLS data. Rather than using the raw value of income as the authors did, I log transformed the income variables to account for their severe right skew and make their distributions more normal. I then regressed *income (1992)* on all of the previous predictors. *Attractiveness* was again not significant (see Table 3).

I also ran another regression with all of the previous covariates that additionally controlled for *tenure* (number of years of employment at current and previous job), *IQ*, and *race* to account for potential omitted variable bias (Table 3). After dropping null and invalid

observations, the sample size was n = 3,072. I found that *attractiveness* was a significant predictor of income at $\alpha = 0.1$: a one unit increase on the normalized attractiveness scale leads to a 1.8% increase in income, holding all other variables constant. All of the controls were significant except for *race* and *family income*.

Table 3: Regressions with logged income variables

	Dependent variable:	
	income_1992	
	With prev. used covariates	With added covariates
	(1)	(2)
income_1974	0.316*** (0.016)	0.290*** (0.016)
attractiveness	0.016 (0.011)	0.018* (0.010)
race		−0.074 (0.116)
tenure		0.013*** (0.001)
gender	−0.299*** (0.036)	−0.283*** (0.035)
iq		0.003*** (0.001)
occ_status	0.0001*** (0.00001)	0.0001*** (0.00001)
famincome	0.021 (0.021)	0.026 (0.020)
education	0.032*** (0.007)	0.036*** (0.007)
Constant	6.522*** (0.226)	6.186*** (0.229)
Observations	3,071	3,071
R^2	0.388	0.411
Adjusted R^2	0.387	0.409
Residual Std. Error	0.712 (df = 3064)	0.699 (df = 3061)
F Statistic	324.385*** (df = 6; 3064)	237.542*** (df = 9; 3061)
Note:		*p<0.1; **p<0.05; ***p<0.01

I then tested for interactions between *attractiveness* and *gender* by running another regression on all predictors with an included interaction term *gender x attractiveness* (see Table 4). Here, the interaction term is significant at $\alpha = 0.1$, but the significance of *attractiveness* on its own disappears. The coefficient on the interaction term indicates that for women (*gender* = 1 if female, 0 if male), a one unit increase on the normalized attractiveness scale leads to a 3.8% increase in income. This is the inverse of the result that Roszell et. al (1989) found (that attractiveness was only significant for men). Ultimately, I was only able to replicate the authors' methodology and results from their simple regression. However, I do find a positive relationship between attractiveness and income in my regression with logged income variables and additional controls.

Normative Concern

There is concern for potential arbitrary labor market discrimination when attractiveness, which is irrelevant to someone's skill level or ability to successfully perform their job, plays a role in the income that employees receive. In the hypothetical scenario that two employees

Table 4: Full regression with interaction term

	Dependent variable:
	income_1992
income_1974	0.291*** (0.016)
attractiveness	0.002 (0.014)
tenure	0.013*** (0.001)
iq	0.003*** (0.001)
race	−0.069 (0.116)
education	0.036*** (0.007)
occ_status	0.0001*** (0.00001)
famincome	0.027 (0.020)
gender	−0.281*** (0.035)
attractiveness:gender	0.038* (0.021)
Constant	6.183*** (0.228)
Observations	3,071
R^2	0.412
Adjusted R^2	0.410
Residual Std. Error	0.699 (df = 3060)
F Statistic	214.259*** (df = 10; 3060)
Note:	*p<0.1; **p<0.05; ***p<0.01

exist with identical traits in every other regard (schooling, skill, position, etc.) at the same firm, but the more attractive employee makes more money, a clear normative concern in labor market discrimination exists.

From a utilitarian perspective, the harm that results from this form of discrimination clearly outweighs any benefit it may produce. There are obvious damaging consequences for "less attractive" employees who are treated unfairly in terms of unequal compensation or other unfair treatment in the workplace. Negative implications for efficiency also exist if high-performing workers are receiving unfair consequences based on perceived lack of attractiveness, and low-performing workers are receiving unfair benefits based on perceived attractiveness. These high-performing workers are more likely to leave or cease to perform as well if they are not being compensated or treated fairly, and low-performing workers who are unfairly rewarded for their attractiveness are more likely to rise into positions that they are not the optimal candidates for. Ultimately, this issue results in a net negative as the consequences for employees and firms outweigh the benefits that attractive people would receive in the form of economic premiums. Therefore, utilitarianism would dictate that this form of arbitrary discrimination is wrong.

There is also the more likely possibility that the disparity in income by attractiveness is not through explicit employer-employee discrimination, but because people may see attractiveness as a proxy for other traits (whether they are true or not): for example, perhaps conventionally attractive people are viewed as more outgoing or approachable and thus are able to form better relationships with their coworkers or clients, making them ultimately

more productive workers. This scenario is a representation of the concepts of "beauty is good" or "pretty privilege", where explicit labor market discrimination may not exist, but an ethical concern is still present in that "beauty" is unfairly assumed to imply positive or morally "good" personality traits while "ugliness" is not. From a utilitarian perspective, the harmful outcomes that this kind of subconscious arbitrary discrimination elicits makes this scenario ethically wrong, regardless of a lack of harmful intent.

Conclusion

Roszell et. al (1989) find that physical attractiveness leads to increased income for middle-aged and older males in male-dominated occupations. Their finding that attractiveness is associated with an economic premium is generally supported by other research in this area. However, multiple factors lead me to be reproachful of their data and methodology, including how attractiveness ratings were made, how the authors tested for interactions, and potential omitted variable bias. Therefore, their results seem to be less robust or reliable. Other literature on this topic may provide more accurate results. In my own analysis, I find that attractiveness leads to increased income for women.

Their study still raises an important normative concern regarding arbitrary discrimination in the workplace. An ethical problem exists if workers with the same skills and backgrounds receive different incomes solely due to their levels of physical attractiveness. Future research to determine whether this form of arbitrary discrimination is due to explicit discrimination or rather because people subconsciously see beauty as a proxy for certain personality traits is needed (i.e., if some measure of personality traits could be included in the regression analysis). However, utilitarianism would dictate that, regardless of conscious or subconscious intention to discriminate, the income disparity between more attractive and less attractive workers is wrong because it results in a net negative for employees and firms.

Works Cited

1. Dion, K., Berscheid, E., & Walster, E. (1972). What is Beautiful is Good. *Journal of Personality and Social Psychology*, 24(3), 285. doi:10.1037/h0033731

2. Frieze, I. H., Olson, J. E., & Russell, J. (1991). Attractiveness and Income for Men and Women in Management. *Journal of Applied Social Psychology*, 21(13), 1039–1057. doi:10.1111/j.1559-1816.1991.tb00458.x

3. Hamermesh, D. S., & Abrevaya, J. (2013). Beauty is the Promise of Happiness? *European Economic Review*, 64, 351–368. doi:10.1016/j.euroecorev.2013.09.005

4. Roszell, P., Kennedy, D., & Grabb, E. (1989). Physical Attractiveness and Income Attainment Among Canadians. *The Journal of Psychology*, 123(6), 547-559. doi:10.1080/00223980.1989.10543009

5. Scholz, J., Sicinski, K. (2015). Facial Attractiveness and Lifetime Earnings: Evidence from a Cohort Study. *The Review of Economics and Statistics*, 97(1): 14–28. doi:10.1162/REST_a_00435

6. Wisconsin Longitudinal Study (WLS) [graduates, siblings, and spouses]: 1957-2020 Version 14.03. [machine-readable data file] / Hauser, Robert M., William H. Sewell, and Pamela Herd. [principal investigator(s)]. Madison, WI: University of Wisconsin-Madison, WLS. [distributor]; https://researchers.wls.wisc.edu/documentation/