Ian Hu
Prof. Wadhwa
ECON 123CW
Final Paper
6/2/22

Effect of Education on Poverty Rates

ABSTRACT

This paper will focus on the relationship between education and poverty. More specifically, the education rate comprises adults 25 and older who have completed college whereas the poverty variable is the poverty rate of each state. The data is cross-sectional from the United States, including the 50 states and the District of Columbia. Other variables include average K-12 spending per pupil, average post-secondary spending per pupil, urban percentage, gini coefficient, cost of living index, and unemployment rate. The modeling begins with a best subsets model which prompted me to create the full model. From there, t-statistics and F-statistics were used to confirm the significance of the variables and model and compared it with a restricted model. The restricted model was proven to be required and the research ultimately found a significant negative relationship between education and poverty rate.

BACKGROUND

Poverty is everywhere—it affects essential parts of life, such as housing, food, and education, which bleeds into future generations. It's a roadblock for both individual and overall economic development and success. As defined by the Organization for Economic Cooperation and Development, the poverty rate is "the ratio of the number of people (in a given age group) whose income falls below the poverty line; taken as half the median household income of the total population" (OECD). The poverty rate can have different divisions, including those under 17 years of age, working adults, and seniors over 65. This paper focuses on the poverty of legal adults, i.e. those over 18 years of age.

Poverty rates have been on the rise as of recent years, and its effects have only been magnified throughout the recent pandemic, causing mass unemployment and financial hardship. Since 2020, the poverty rate has increased to 11.4%, a full percentage point above the previous year. Strikingly, this was the first rise in over 5 consecutive years of poverty rate decline. Additionally, the median household income level decreased from 2019 to 2020, from $69,560 to $67,521. Note that these figures do not include stimulus payments nor tax credits (*Income, Poverty, and Health Insurance Coverage*).

Although difficult, it is possible to escape the vicious cycle of poverty. Many believe that education is a sure highway towards opportunity and higher income. A better education equips one with better knowledge and more workable skills, all of which are desired in the workforce, regardless of field. The sheer fact of having a college diploma is a signal towards employers that the individual is productive, intelligent, and eager to work.

The relationship between education and poverty are visible in two main methods. Firstly, “investment in education increases the skills and productivity of poor households [by enhancing] the income level and as well as the overall standard of living” (Awan). Poverty affects education mainly through the impediment to opportunity due to financial constraints. These two effects compound on each other to be vicious in nature, reinforcing the barrier of poverty and bleeding into future generations.

Given the importance of poverty, this research paper aims to answer the question, “To what extent does the completion of higher education affect poverty rates in the United States?” Using cross-sectional data including the 50 states of the US and the District of Columbia from 2018-2021, the relationship between poverty and education is analyzed. Additional control variables are introduced, such as average K-12 spending per pupil, average post-secondary spending per pupil, urban percentage of the population, gini coefficient, cost of living indices, and unemployment rates. Utilizing these additional variables will help broaden the understanding of how education impacts poverty rates.

LITERATURE SURVEY

Bharti (2021) measured mainly the relationship between poverty rate and education rates of adults 25 years and older. Bharti similarly used cross-sectional data from the United States, focusing on cost of living, unemployment rates, labor force participation rates, urban percentage, and GDP per capita. Bharti initialized his research with simple linear regression models to study the ceteris paribus effect, i.e. he first modeled poverty rates against education. Following the one-variable model, he supplemented the estimation with cost of living, unemployment, labor force participation, urban percentage, and GDP per capita. Similarly, he created a model which excluded the insignificant variables—unemployment, urban percentage, and gdp per capita—and ultimately found that they were jointly insignificant. Overarchingly, he found a statistically significant relationship between education and poverty rate—it was negative in nature.

Azih (2017) seeks to uncover the factors behind poverty rates, and uses variables such as education, divorce rates, female-headed household, crime, welfare, unemployment, and minimum wage. Azih focuses mainly on social and economic welfare. The study is cross-sectional, focusing on the 50 states of the US in 2013 (Azih). More grounded in the familial social structure, Azih sheds light on underlying roots of poverty and finds that public spending on education has a moderate negative correlation with poverty rates. Interestingly, Azih focuses on the ceteris paribus effect of these variables on poverty and looks at multiple one-variable regressions. The model utilizes ANOVA, Pearson's correlation coefficient (r), and an Eta-squared range to determine the significance and subsequent strength of the variable on poverty rates (Azih). Relevant to this paper, Azih found the following:

1) A negative, moderate relationship between education spending and poverty.
2) A strong, positive relationship between unemployment and poverty

Azih's focus on both social and economic factors contributing to poverty allows for a wider and broad explanation of the causes of poverty rates (Azih).

Ladd (2012) focuses on the US education system and how certain education policies have a lesser effect than imagined or perhaps provide a negative effect. Ladd examines the effects of No Child Left Behind, test evaluations of teachers, and the promotion of educational competition between and among students. Ladd's overarching hypothesis is that these programs fail to address the challenges faced by students whose families suffer lower socioeconomic status (Ladd). Comparatively to what this paper is discussing, Ladd reverses the relationship between poverty and education rates and seeks to uncover how poverty affects education rates given variables such as family income, parental education level, and parental employment status. It is important to note that Ladd explains that the effect of family income on education can be explained as either causal or correlated. Regardless, this distinction is important to note for later analysis. Ladd's analysis focuses on K-8 education, both domestic and international. As such, the analysis differs comparatively to this paper, which includes K-12 in addition to higher education, albeit limited to public and domestic education. Ladd's use of a bivariate regression model uncovers "[strong] evidence that child poverty in itself may be causally linked to education outcomes" (Ladd, 6).

HYPOTHESIS AND MODEL

The paper mainly hypothesizes that the education rate will have a strong negative correlation with the poverty rate. That is, as the completion rate of college increases, the poverty rate should decrease. This thinking derives from the human capital theory, where "education and training are investments that make individuals generally more productive. Individuals who are more productive will, according to [the human capital theory] also have higher earnings and be more employable" (Carneiro, 255). Subsequently, K-12 and post-secondary spending is expected to have similar negative correlations, however the latter to have a larger overall effect,

due to the effect a college degree has on obtaining a higher-paying job in comparison to a high school diploma. Regarding the signs of urban percentage and cost of living, the expectation of their signs are unclear. College has been made much more accessible within the last 20 years, given increases in public expenditures on post-secondary education attainment, however the extent to which it has impacted the poverty rate is still yet to be analyzed (Tilsley). Regarding cost of living, a higher cost of living may imply higher wages, on average, and perhaps more individuals over the poverty threshold, however a lower cost of living implies a cheaper location and thus less individuals that are poverty-stricken. The gini coefficient is expected to have a positive relationship. The gini coefficient measures wealth distribution, so a higher gini index corresponds to a higher wealth distribution inequality. From Marxian theory, poverty and inequality are integral components of capitalism, and as such, the gini coefficient, which measures income inequality, is expected to have a positive correlation with poverty (Peet). The Keynesian school of thought believes that unemployment is largely involuntary and causes poverty, i.e. a positive correlation (Prasetyo). As the unemployment rate increases, so will the poverty rate, as individuals struggle to make income.

The estimated model is listed below in Figure 1:

$$Pov_i = \beta_0 + \beta_1 Educ._i + \beta_2 Spd_i + \beta_3 PS_i + \beta_4 Urb_i + \beta_5 Gini_i + \beta_6 CoL_i + \beta_7 Unemp_i$$

Figure 1: Description of Variables

Variable	Description	Units
Pov	Poverty Rate of state i	Percent
Educ	Education Rate of state i	Percent
Spd	Average K-12 Spending per Pupil of state i	Current US Dollars
PS	Average Post-Secondary Spending per Pupil of state i	Current US Dollars
Urb	Urban Percentage of state i	Percent
Gini	Gini Coefficient of state i	Index (0-1)
CoL	Cost of Living index in state i	Index (0-100)
Unemp	Unemployment rate in state i	Percent

The dependent variable of interest is the poverty rate of each state. Education rate is the percentage of adults 25 and older who have completed college. Both K-12 and post-secondary spending are the average public education expenditures per pupil per state. Urban percentage is the rate of individuals that live in an urban area as opposed to rural. The gini coefficient is a numerical measure of the wealth dispersion in a given area—it essentially measures income inequality. A gini coefficient of 0 represents a perfect income distribution whereas a coefficient of 1 represents perfect income inequality; one individual has all of the income. It is important to note that different countries with vastly different wealth levels can have similar gini indexes—the statistic simply measures the distribution of wealth, not the wealth itself. The cost of living is a weighted average of the cost of groceries, housing, transportation, utilities, health, etc. and is indexed between 0 and 100. The unemployment rate is the number of economically defined unemployed individuals (i.e. those actively searching for a job but are unable to find one) divided by the number in the labor force. The data is collected mostly from 2018 to 2021 and comprises 51 observations: the 50 states of America and the DIstrict of Columbia.

RESULTS

The analysis begins with a best subset selection model to simply find the best model among the chosen variables. The output is listed below in Figure 2:

Figure 2: Best Subset Model

```
Subset selection object
Call: regsubsets.formula(Pov ~ . - State, data)
7 Variables  (and intercept)
      Forced in Forced out
Educ      FALSE      FALSE
Spd       FALSE      FALSE
PS        FALSE      FALSE
Urb       FALSE      FALSE
Gini      FALSE      FALSE
CoL       FALSE      FALSE
Unemp     FALSE      FALSE
1 subsets of each size up to 7
Selection Algorithm: exhaustive
         Educ Spd PS  Urb Gini CoL Unemp
1  ( 1 ) " "  " " " " " " " " " "  " " "*"
2  ( 1 ) "*"  " " " " " " " " "*"  " " " " "
3  ( 1 ) "*"  " " " " " " " " "*"  " " "*"
4  ( 1 ) " "  "*" " " "*" "*"  " " " " "*"
5  ( 1 ) "*"  "*" " " "*" "*"  " " " " "*"
6  ( 1 ) "*"  "*" " " "*" "*"  "*" "*"
7  ( 1 ) "*"  "*" "*" "*" "*"  "*" "*"
```

Interestingly, with a one-variable model, the best subsets model chose Unemp. (unemployment) as the best variable. However, with a two-variable model, the preferred model includes Educ. (education) in conjunction with Gini (gini coefficient). Unemployment was expected to be the best variable in a one-variable model, but the selection of education and the gini coefficient was surprising and interesting to note. Given the interconnectedness of unemployment and poverty rates, unemployment rate was expected to be an integral variable, regardless of model size. Moving forward, for a seven-variable model, it can be seen that the best subset model has chosen all the variables. Consequently, one can infer that all seven variables explain the data well enough to be selected—all seven variables are sufficient for a full model.

After creating the full model, the results of the multiple regression are listed below in figure 3.

Figure 3: Full Model

```
Call:
lm(formula = Pov ~ . - State, data = data)

Residuals:
       Min         1Q     Median         3Q        Max
-0.0182766 -0.0063461 -0.0002311  0.0071444  0.0210580

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.079e-01  3.797e-02  -5.476 2.10e-06 ***
Educ        -9.403e-02  3.687e-02  -2.550   0.0144 *
Spd         -2.645e-06  5.313e-07  -4.979 1.09e-05 ***
PS          -5.339e-08  2.186e-07  -0.244   0.8082
Urb         -6.293e-02  1.257e-02  -5.008 9.86e-06 ***
Gini         8.514e-01  8.614e-02   9.884 1.23e-12 ***
CoL          9.806e-05  9.548e-05   1.027   0.3102
Unemp        1.049e-02  2.041e-03   5.140 6.39e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.009932 on 43 degrees of freedom
Multiple R-squared:  0.8781,    Adjusted R-squared:  0.8583
F-statistic: 44.25 on 7 and 43 DF,  p-value: < 2.2e-16
```

In this regression model, the poverty rate (Pov) is the dependent variable, with education rate (Educ), K-12 spending (Spd), and post-secondary spending (PS), urban percentage (Urb), gini coefficient (Gini), cost of living index (CoL), and unemployment rate (Unemp). The model suggests the following (note that variables that increase by 1% have been adjusted):

As the education rate increases by 1%, the poverty rate decreases by 0.0009403, keeping all variables constant.

As average K-12 spending increases by $1, the poverty rate decreases by 0.000002645, keeping all variables constant.

As average post-secondary spending increases by $1, the poverty rate decreases by 0.00000006293, keeping all variables constant.

As the gini coefficient increases by 1%, the poverty rate increases by 0.008514, keeping all variables constant.

As the cost of living index increases by 1, the poverty rate increases by 0.00009806, keeping all variables constant.

As the unemployment rate increases by 1%, the poverty rate increases by 0.01049, keeping all variables constant.

With all of the variables equal to 0, the poverty rate would be expected to lie at -0.2079, however this is not meaningful as a negative poverty rate is impossible.

At the 0.05 significance level, the factors Educ, Spd, Urb, Gini, and Unemp are all significant. However, PS and CoL are statistically insignificant. The adjusted R-Squared of the model is 0.8583, which means that the variables explain roughly 85.8% of the variance within the poverty rate. The model supports my main hypothesis that poverty rate and education rate would have a negative relationship, as expected.

Following the results of the full model, the variables PS and CoL (post-secondary spending and cost of living index) are dropped. The model is listed below in Figure 4.

Figure 4: Full Model

```
Call:
lm(formula = Pov ~ . - State - PS - CoL, data = data)

Residuals:
       Min         1Q     Median         3Q        Max
-0.0185550 -0.0066968  0.0001371  0.0060523  0.0225238

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.033e-01  3.430e-02  -5.929 3.98e-07 ***
Educ        -8.633e-02  3.464e-02  -2.492   0.0164 *
Spd         -2.429e-06  4.778e-07  -5.083 6.97e-06 ***
Urb         -5.867e-02  1.170e-02  -5.014 8.78e-06 ***
Gini         8.438e-01  8.262e-02  10.214 2.67e-13 ***
Unemp        1.040e-02  1.997e-03   5.210 4.56e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.009829 on 45 degrees of freedom
Multiple R-squared:  0.8751,    Adjusted R-squared:  0.8612
F-statistic: 63.04 on 5 and 45 DF,  p-value: < 2.2e-16
```

The model predicts:

As education rate increases by 1%, the poverty rate decreases by 8.633e-02%

As K-12 spending increases by $1, the poverty rate decreases by 2.429e-06.

As urban rate increases by 1%, the poverty rate decreases 5.867-02%.

As the gini coefficient increases by 1%, the poverty rate increases by 8.438e-01%.

As unemployment increases by 1%, the poverty rate decreases by 1.040e-02%.

With all of the variables equal to 0, the poverty rate would be expected to lie at -2.033e-01, however this is not meaningful as a negative poverty rate is impossible.

Additionally, all explanatory variables in this model are significant at the 0.05 level. This model also supports the hypothesis that poverty rate and education rate have a negative relationship. The adjusted R-squared of the model is 0.8612, meaning that the 5 variables Educ, Spd, Urb, Gini, and Unemp explain roughly 86.12% of the variation in Pov.

To summarize the two models in a more visual manner, Figure 5 below shows the comparison between the coefficients, intercepts, and R-squared values.

Figure 5: Model Comparison

Independent Variable	Model 1	Model 2
Educ	-0.09403	-0.08633
Spd	-0.000002645	-0.000002429
PS	-0.00000005339	
Urb	-0.06293	-0.05867
Gini	0.8514	0.8438
CoL	0.00009806	
Unemp	0.01049	0.01040
Intercept	-0.2079	-0.2033
Adjusted R-Squared	0.8583	0.8612

To find the best model, an ANOVA test was run comparing the two. The full model is listed below.

$$Pov_i = \beta_0 + \beta_1 Educ._i + \beta_2 Spd._i + \beta_3 PS._i + \beta_4 Urb_i + \beta_5 Gini_i + \beta_6 CoL_i + \beta_7 Unemp_i$$

We will test the joint significance of the PS and CoL variables with the following hypothesis

$$H_0: \beta_3 = \beta_6 = 0$$

$$H_A: At\ least\ one\ \beta_i \neq 0,\ for\ i = \{3, 4\}$$

Running an ANOVA test with the function *anova(full, restricted)*, the output is shown below in Figure 6:

Figure 6: Analysis of Variance Table

```
Analysis of Variance Table

Model 1: Pov ~ (State + Educ + Spd + PS + Urb + Gini + CoL + Unemp) -
    State
Model 2: Pov ~ (State + Educ + Spd + PS + Urb + Gini + CoL + Unemp) -
    State - PS - CoL
  Res.Df       RSS Df   Sum of Sq      F Pr(>F)
1     43 0.0042419
2     45 0.0043477 -2 -0.00010579 0.5362 0.5888
```

As $p > 0.05$, one fails to reject the null hypothesis. One concludes that there is insufficient evidence to support the alternative hypothesis and the variables are statistically not different from 0. Thus the restricted model is required. Additionally, this conclusion aligns with the adjusted R-squared values of the two models, as the restricted model had a slightly higher value, at 0.8612 compared to 0.8583. However, the difference in R-squared values is negligibly small, so arguments can be made regarding the efficacy and appropriateness of each model. Regardless of which model is more effective, both support the notion of a negative relationship between education rate and poverty rate, validating the original hypothesis.

CONCLUSION

This study aimed to better understand poverty's relationship with education, controlling for variables such as educational spending, urban percentage of the population, costs of living, wealth inequality, and unemployment. Using the 50 states of America and its District of Columbia, the study was able to quantify the effects of said variables on poverty.

The study concluded the following:

1) College completion rates, average K-12 spending per pupil, and urban percentage of the population all have significant negative relationships with poverty rates, all else equal.
2) The gini coefficient and unemployment rate both have significant positive relationships with poverty rates, all else equal.

3) Average post-secondary education and average cost of living indices have an insignificant relationship with poverty rates, all else equal

Note that the study tested these variables in conjunction with each other, so the aforementioned conclusions are conditional on the inclusion of the listed variables in this study.

To acknowledge, a limitation of this study includes that poverty is a slow-moving force. Policies aimed at reducing poverty may take years to see significant change, if any at all. To better understand the effects, it may behoove another to analyze the effects of these variables as a time-series, rather than cross-sectional. Many of the variables used in this study are also slow-to-change. Furthermore, the study is limited to the United States, so other countries may experience different effects due to cultural or social norms, relative income levels, geography, or differences in other economic factors, such as wealth distribution or government policies. Another factor of interest would be to look at the varying impacts of education based on levels. As the model in the study suggested, average post-secondary spending per pupil on education was insignificant, however this may be overshadowed by the inclusion of the college completion rate. Additionally, the study only looks at poverty rates of adults, and different conclusions may occur looking at the poverty rates of children (0-17) or seniors (65+) with different variables.

It is crucial to encourage education for the population. To increase knowledge and human capital is to increase productivity and raise standards of living. States can invest more in educational spendings or training programs targeted for public schools and universities. States can also improve programs towards financial support. Financial or basic needs assistance can help alleviate the pressures of poverty, such as subsidies for education, health, families, or transportation.

Bibliography

Awan, Masood Sarwar and Malik, Nouman and Sarwar, Haroon and Waqas, Muhammad (2011): *Impact of education on poverty reduction.* Published in: International Journal of Academic Research , Vol. 3, No. 1 (2011): pp. 659-664.

Azih, Chioma Austin-Mona. Analysis of Factors Contributing to Poverty in the United States: An Empirical Study. 2017 Methodist University. https://www.methodist.edu/wp-content/uploads/2018/09/mr2017_chioma.pdf

Bharti, Ekansh. The Effect of Education on Poverty. 2021. Georgia Tech Library. https://smartech.gatech.edu/bitstream/handle/1853/64480/Econ%203161%20Paper%2-%20Ekansh%20Bharti.pdf?sequence=1&isAllowed=y

"Cost of Living Data Series." Missouri Economic Research and Information Center, 2021, https://meric.mo.gov/data/cost-living-data-series.

"Income, Poverty and Health Insurance Coverage in the United States: 2020." Census.gov, US Census Bureau, 2 Nov. 2021, https://www.census.gov/newsroom/press-releases/2021/income-poverty-health-insurance-coverage.html#:~:text=Policy%20Directive%2014).-,The%20official%20poverty%20rate%20in%202020%20was%2011.4%25%2C%20up%201.0,million%20more%20than%20in%202019.

"Median Household Income, by State." National Center for Education Statistics, https://nces.ed.gov/programs/digest/d20/tables/dt20_102.30.asp.

"Completion Rates, 2015-2019 (Completing college, adults 25 and older)" USDA Economic Research Service, 24 Feb. 2021, https://data.ers.usda.gov/reports.aspx?ID=17829.

"Gini Coefficient as a Measure for Household Income Distribution Inequality for U.S. States in 2019." Statista, Sept. 2020, https://www.statista.com/statistics/227249/greatest-gap-between-rich-and-poor-by-us-state/.

Hanson, Melanie. “U.S. Public Education Spending Statistics” EducationData.org, March 15, 2022, https://educationdata.org/public-education-spending-statistics

Hegedus, Andy. Evaluating the Relationships Between Poverty and School Performance. NWEA Research, 2018. Institute of Education Sciences. https://files.eric.ed.gov/fulltext/ED593828.pdf.

Ladd, H.F. (2012), Education and Poverty: Confronting the Evidence. J. Pol. Anal. Manage., 31: 203-227. https://doi.org/10.1002/pam.21615

OECD (2022), Poverty rate (indicator). doi: 10.1787/0fe1315d-en (Accessed on 27 May 2022)

Peet, Richard (1975), Inequality and Poverty: A Marxist-Geographic Theory, Annals of the Association of American Geographers, 65:4, 564-571, Doi: 10.1111/j.1467-8306.1975.tb01063.x

“Poverty.” USDA ERS - Data Products, 5 Jan. 2021, https://data.ers.usda.gov/reports.aspx?ID=17826.

Prasetyo, P. and Thomas, P. (2021) A Simple Mitigation Model of Poverty Reduction in Indonesia. Open Journal of Business and Management, 9, 2742-2758. Doi: 10.4236/ojbm.2021.96152.

P. Carneiro, et. al., “The Economics of Vocational Education and Training” International Encyclopedia of Education (Third Edition), Elsevier, 2010, Pages 255-261, https://doi.org/10.1016/B978-0-08-044894-7.01737-1.

Tilsley, Alexandra, and Elizabeth Forney. “How Has Education Funding Changed over Time?” Urban Institute, https://apps.urban.org/features/education-funding-trends/.

“Unemployment Rates for States.” U.S. Bureau of Labor Statistics, U.S. Bureau of Labor Statistics, 2 Mar. 2022, https://www.bls.gov/lau/lastrk19.htm.

“Urban Percentage of the Population for States, Historical.” Iowa State University, Iowa Community Indicators Program, https://www.icip.iastate.edu/tables/population/urban-pct-states.