

How Education Makes Differences on Individuals?

MACS 40800 - Project Proposal

Steven Cao
Jesús Pacheco
Sharon Shen
Yaoxi Shi

Introduction

Current opinion holds strongly that education is key to realising the potential and growth of individuals, regardless of their setting and background. By extension, it is also commonly held that investing more time and energy into education results in a positive net impact for an individual's economic status. The association between increasing education and growth of economic capacity at both national and individual levels has been repeatedly noted (Psacharopoulos 1994), leading policy to be in favour of allocating more and more resources towards education. Interpretations of why education leads to better outcomes (assuming that it does), however, are far from unanimous (Pritchett 2006). The main reason is that it is not clear what exactly about education leads to better outcomes: clear causal connections are lacking (deMeulemeester & Rochat 1995). Various factors more specific than just the generic grouping of "education" have to be taken into consideration, with examples being the type of major or the type of degree pursued, because it has also been shown that not all kinds of education are weighed equally when it comes to economic benefits (Lin 2004).

More granularity could be provided here as to the exact relationships between education and economic outcomes. We seek to explore such relationships through exploratory data analysis of the General Social Survey (GSS). We will be exploring variables for which we think there could

be an intuitive explanation, e.g. comparing years of education and degree earned against income, in order to observe clusters which may or may not validate our intuitions. The goal, ultimately, will be to parse which aspects of education are most linked to economic outcome, with one potential and direct benefit being a better-informed policy towards education.

We anticipate the results of our exploratory data analysis to be the identification of clusterings which could potentially correspond to familiar classifications. For example, we may find high density of observations across certain ranges within a scatterplot of specific income brackets and years of education, that may in turn correspond to the individual's socioeconomic status, or the type and nature of occupation.

Literature Review

Education is traditionally understood as one of the major determinants of income level (Carnevale, Rose and Cheah 2010). Empirical evidence from various researches show that over a lifetime, individuals who have attended college make more money than those without. Specifically, within the concept of education attainment, degree types play an important role in creating tiers of earning boost effect. The largest jump in monthly household income is when one moves from Associate's degree to a Bachelor's degree, with observed increase of \$2391 in 2009 (Fry 2014), while the wage difference is less significant between Bachelor's and Master's. It is possible that the marginal utility of degree, or effective economic payoff relative to initial investment in education, diminishes as one becomes more advanced in academic, despite the general emphasis of postsecondary education.

Moreover, existing literature also explores the variations in earnings by occupation. Given the same educational attainment, be it bachelor or master, people who work in healthcare and

STEM (science, technology, engineering and mathematics) fields earn much higher wages as compared to those who are in office support or blue-collar positions (Carnevale, Rose and Cheah 2010). When taking a closer examination at the distinction, STEM major students are more likely to have better income profile and greater job satisfaction than non-STEM graduates (Xu 2012). Based on a longitudinal national survey of students with respect to their academic choices and career path, Webber (2014) addresses the heterogeneity in economic returns among majors. Not surprisingly, STEM graduates realize the largest premium in wage differential, a lifetime earning of \$1.5MM more than high schoolers with no college experience, all else equal.

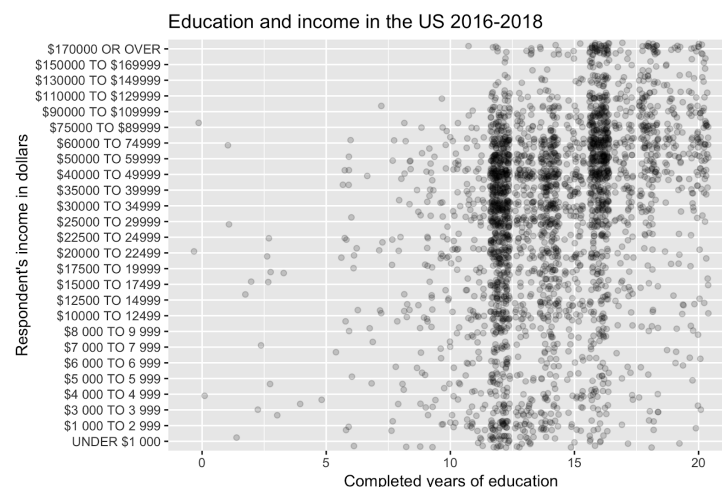
What the discussion of STEM brings to light is the impact of areas of specialization one chooses to study may have on one's earning potential. A study in Finland, for example, might be worth drawing some implications from, for our understanding of the job market dynamics in the U.S. In her research, Prix (2013) shows that although higher education is associated with betterment of personal annual gross earnings, whether the level of education lifts individuals' socio-economic status is dependent on the fields of study. Using the Finnish census data over three decades, it appears that the economic return and subsequent wage differential of an ordinary bachelor's degree in engineering is at least equal, if not higher, than that of an advanced degree in humanities (Prix 2013). Of course, degree level and major field are just some example factors among the broader facets of education that deem attention. Through this research project, our hope is thus to unveil the data distribution of income levels based on insights about certain educational attributes as such in the modern American context.

Data and Methodology

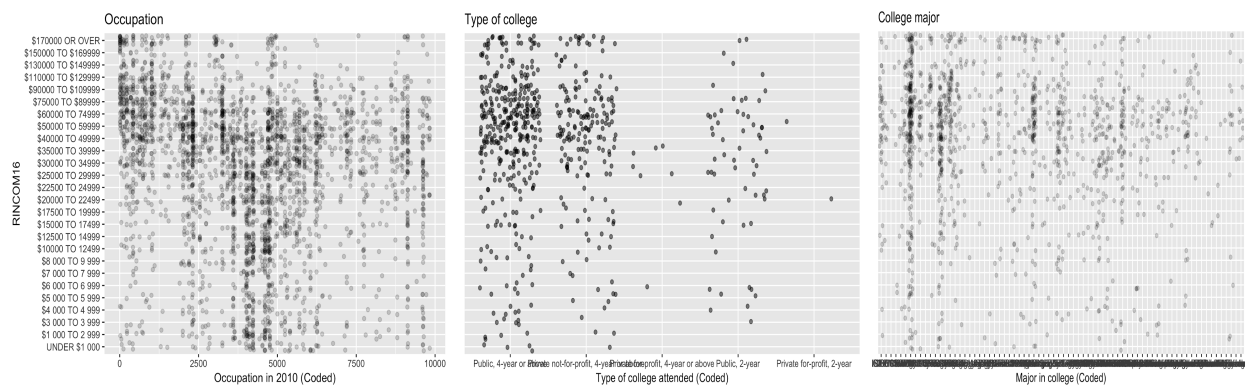
The data we will use to explore these questions is the General Social Survey (GSS) available online from NORC at the University of Chicago. We will use categories of variables such as background, education, income, occupation, stated opinions on different subjects and any other type of information that we might be inferring from the analysis. Luckily, the GSS has more than 5,000 variables ranging from sociodemographic to occupations and work-related topics. The survey has a national probability sample and is conducted biyearly since 1972.

Particularly, we know that the income variable is a difficult variable to disclose, but we rely on the approach taken by this survey using income brackets. Unfortunately, the brackets are not the same among surveys (in 1972 the annual income upper bracket was *25,000 USD and more*, while the latest version 2016 has an upper bracket of *170,000 USD and more*). Hence, we are using, at least for now, the information for 2016 and 2018. We present an example of the use of these data just to kick-off the project.

The relation between the completed years of education of respondents and their income in dollars is displayed in the next graph:



We can also see try to look for patterns in the next 3 graphs where we plot income with three variables: occupation, type of college and college major. At this early stage of the project, we can still do not see clear clusterability. But we are trying to get to know the data and identifying trends.



As mentioned above, we will first look for clusterability in the returns to education (income divided by the number of years invested on education) among different cohorts (occupations and majors). Using different techniques covered in this course, such as soft-partitioning and association rule mining we will try to discover some useful patterns and reveal more useful questions for further research.

Implications

In this project, we aim to explore how factors related to education (e.g. years of education, types of degree etc) would influence individual's socioeconomic status or the type of occupation. By using unsupervised machine learning methods to explore the patterns from these variables, this

project could contribute to the theoretical research about education and also would potentially influence the policy-making in this field.

In the literature, previous studies on education and individual's development have indicated that education significantly predicts increased income and social status of individuals, while the role of more specific factors that relate to education underlying this association still remains unclear (Pritchett 2006). By analyzing large scale data from GSS database, our project aims to provide more evidence to unfold the distribution of socioeconomic status and occupation with respect to specific variables of education, thus contribute to closing the gap in existing social education theories.

Practically, the finding of this project could provide insights for current and future policy makers in education related field. For example, by revealing what are the likely factors in education that contribute to the grouping of individuals' future income and socioeconomic status, policy makers could potentially allocate more resources to enhance these factors, to support citizens' personal growth and future development, therefore improve life satisfaction in a large scale. Also, in a micro level, the findings of this project would also very likely to benefit each individual. Recognizing what are the factors that play the most important roles in improving their future income or social status, individuals may make better decisions along their personal development, such as which major to choose, whether to go for another degree etc.

Reference

Carnevale, A., Rose, S. and Cheah, B. (2011). *The College Payoff: Education, Occupations, Lifetime Earnings*. [online] Ww2.ed.gov. Available at:
<https://www2.ed.gov/policy/highered/reg/hearulemaking/2011/collegepayoff.pdf>
[Accessed 16 Oct. 2019].

deMeulemeester, J., & Rochat, D. (1995). A causality analysis of the link between higher education and economic development. *Economics of Education Review*, 14, pp.351–361.

Fry, R. (2014). *For Millennials, a bachelor's degree continues to pay off, but a master's earns even more*. [online] Pew Research Center. Available at:
<https://www.pewresearch.org/fact-tank/2014/02/28/for-millennials-a-bachelors-degree-continues-to-pay-off-but-a-masters-earns-even-more/> [Accessed 16 Oct. 2019].

Lin, T. (2004). The role of higher education in economic development: an empirical study of Taiwan case. *Journal of Asian Economics*, 15, pp.355-371.

Pritchett, L. (2006). Does learning to add up add up? The returns to schooling in aggregate data, In: Hanushek, Erik A. and Finis Welch e., (Eds.). *Handbook of the Economics of Education*, 1, pp.635–695.

Prix, I. (2013). More or different education? The joint impact of educational level and fields of study on earnings stratification in Finland, 1985–2005. *Acta Sociologica*, 56(3), pp.265–284. doi: 10.1177/0001699312474381

Psacharopoulos, G. (1994). Returns to education: A global update. *World Development*, 22(9), pp.1325–1343.

Webber, D. (2014). The lifetime earnings premia of different majors: Correcting for selection based on cognitive, noncognitive, and unobserved factors. *Labour Economics*, 28, pp.14-23.

Xu, Y. (2012). Career Outcomes of STEM and Non-STEM College Graduates: Persistence in Majored-Field and Influential Factors in Career Choices. *Research in Higher Education*, 54(3), pp.349-382.