Actual and Perceived Effectiveness of Alternative Education

John Vandivier

**I. Background and Outline**

Alternative education consists of alternative credentials, paths, and pedagogies[[1]](#footnote-1). These alternative entities are defined by the negation of their traditional counterparts.

This study focuses on the case of the United States, although certain parts are applicable elsewhere. In this context, traditional education refers to public education from Kindergarten through the 12th grade. Traditional higher education refers to a 4-year degree from an accredited university in the United States.

A path is a sequence of steps or a series of tasks. An educational path results in the acquisition of learning or a learning credential. The traditional path for a student in the United States is to enroll in a university in the year following high school graduation. An alternative path would be to enter the work force following high school graduation, and to enroll in college later in life. Alternative paths may result in traditional or alternative credentials, but traditional paths always lead to traditional credentials.

Traditional pedagogy refers to learning-by-lecture. Learning-by-lecture is the most common form of teaching in traditional institutions of education, but it is also utilized in alternative institutions of education. Likewise, alternative pedagogies may be utilized within traditional or alternative institutions of education. Alternative pedagogies include, non-exhaustively, learning-by-doing, learning-by-teaching, flipping the classroom, self-directed learning, and the Socratic method of teaching.

The main hypothesis in this study is that alternative education is a comparatively preferred path to a career for many careers. This study specifies the industrial and other conditions under which alternative education is and is not practically preferred. The results of this study are equally important at the macroeconomic and microeconomic levels; for policy and for personal financial considerations.

Section 1 describes the content, importance, and organization of the paper. Section 2 – 4 describe original research and results. Section 5 synthesizes results and includes third party research. Section 6 concludes with recommendations for practical implementation.

**II. Attitudinal Study**

II.I Attitudinal Study Architecture

A study was sampled on February 26, 2018, to discover attitudinal data regarding alternative education[[2]](#footnote-2). The survey was executed online using SurveyMonkey software. A total sample of 141 was comprised of 103 purchased responses and 38 responses generated by word of mouth requests. The questions used are listed in Appendix A.

Exploratory data analysis was conducted for two related variables of interest[[3]](#footnote-3). A general index of favorability index was constructed for responses to three similar survey questions, and one survey question was treated as a standalone variable of interest.

Questions 2-6 were answered on a scale from 1-10. The general favorability index is the sum of the responses for questions 2, 3, and 5. These responses all positively and significantly cross-correlate. The content of these questions is highly related but distinct.

Question 2 (Q2) asks whether a respondent thinks alternative credentials can qualify a person for a job. Question 3 asks whether a respondent thinks alternative credentials will soon become conventional. Question 5 asks whether a respondent thinks online education in general is a good thing for society. Taken together, this index reflects an attitudinal disposition toward the whole ecosystem of online learning, alternative education, and alternative credentials. Q2 is a standalone variable of interest because it specifically asks whether alternative credentials can qualify an individual for an entry level job.

While age and income were reported by group, pseudo-continuous factors were generated by translating the groups into ordinal numbers. Respondents indicated whether they were familiar with certain online learning providers, and the yes responses are summed into a continuous count variable. Squared and cubic factors were generated from continuous variables. Some marginal effects are observed. Some samples were obtained through Survey Monkey’s paid respondent service, and this was captured in a dummy variable which ended up being omitted due to collinearity.

Initial data exploration indicated a significant effect of belonging to the income group with household income $175,000 - $199,000 per year, but this group only had one respondent. No other income group, nor was any other robust categorical variable, was identified by so few respondents. For robustness, exploration was reconducted with a modified sample set in which this observation was dropped.

Exploratory strategy involved a regular procedure for 4 sessions, and an ad-hoc session investigating structural intuitions and curiosities. The regular procedure was executed for both variables of interest using both data sets, leading to the count of 4 sessions. The procedure identified interesting models using the following guidelines:

1. Regress all variables to identify maximum explanatory power.
2. Obtain a weak factor model where all factors have p < .5
3. Obtain a maximum adjusted r-squared model.
4. Obtain a strong factor model where all factors have p < .1

During the regular procedure, the highest p-value factors are dropped 1-by-1 by until the next model is obtained. Factors appearing in any weak model were noted for further exploration. Following the regular process, eventually strong factors were identified with p < .1.

A fifth exploratory session was conducted to investigate whether dropping survey questions effected the significance of remaining variables. In principle, the factors used in this session are observables which could be obtained without participant response.

II.II Attitudinal Study Results

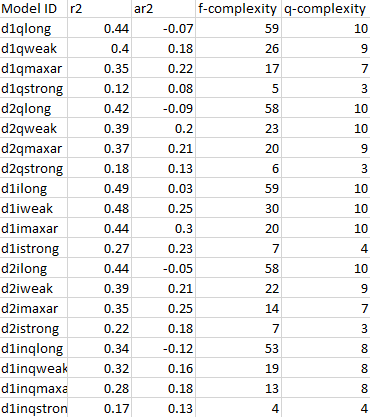
The five sessions lead to 20 models of interest, described in Table 1. Table 1 includes measures of explanation and complexity. R-squared is raw explanatory power, and adjusted r-squared attempts to value simpler models by applying a penalty to every additional factor. Adjusted r-squared is useful, but it misrepresents the concept of simplification in certain ways.

F-complexity counts the number of factors in a model. This is easier to interpret than adjusted r-squared, and it is more straightforward for calculating model price use in a business context. It includes the constant, and it includes extraneous variables omitted from degrees of freedom calculation.

Q-complexity describes the number of questions asked, explicitly or implicitly, to solve a model. It is possible to speak of the f-complexity of a question based on the number of allowed responses. Suppose you have two models with a given adjusted r-squared, but they require a different number of questions to be asked to obtain all factor observations and solve a left-hand estimate. Assuming similar cost per question, the q-simpler model is the cheaper model for practical use.

Factors which were significant in any model are reported in Table 2. Robustness checks were carried out to identify points of strength and importance within the weak factor set, and to identify super-strong factors within the strong factor set. A concept of vertical and horizontal robustness is introduced, where vertical robustness refers to significant factors identified through regular p-value analysis.

**Table 1 – Models of Interest with Explanation and Complexity**



**Table 2 – Significant Factors**

|  |  |  |  |
| --- | --- | --- | --- |
| **Factor Name** | **Significance (p)** | **Strong Model** | **Vertical Effect** |
| eq4\* | 0.026 | d1istrong | 0.43 |
| eq4squared | 0.013 | d2istrong | 0.22 |
| eq4cubed | 0.05 | d2istrong | -0.02 |
| eq6 | 0.001 | d2istrong | 1.31 |
| eq6squared\* | 0.006 | d2qstrong | 0.14 |
| eq6cubed\* | 0.001 | d2istrong | -0.018 |
| male\* | 0.05 | d1inqstrong | 2.29 |
| \_income2 | 0.021 | d2qstrong | 1.84 |
| \_income6 | 0.003 | d1inqstrong | -17.39 |
| \_industry6\*\* | 0.015 | d2qstrong | 2.18 |
| \_industry9\*\* | 0.007 | d2istrong | 8.96 |
| cprovider1\*\* | 0.075 | d1inqstrong | 1.54 |

\*Vertically and Horizontally Robust  
\*\*Vertically and Horizontally Robust and Important Effect Size

Horizontal robustness checks the stability of values across models, across samples, and, in principle, across studies. Multiple study meta-analysis was not conducted in this case, but it follows the conceptualization of horizontal robustness.

Horizontal robustness was measured as the ratio of the frequency of a coefficient of the same sign as the average effect to the frequency of any coefficient, across the twenty models of interest. If a coefficient had the same sign in every case, it received a horizontal robustness of 1. Such variables are reported in Table 3. Values roughly as low as .29 were observed, and arguably factors with a horizontal robustness of less than 1 could be considered interesting. The horizontal effect size is the average effect across models of interest.

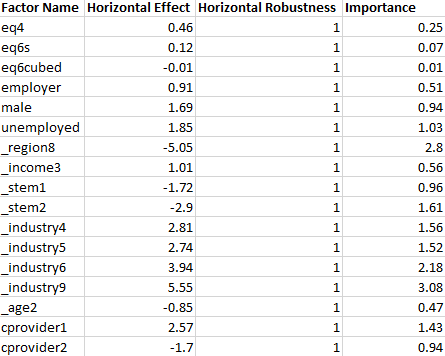
Factors which were significant in one strong model were sometimes significant in multiple strong models. As a result, a significant value for a factor is not uniquely identified. The vertical effect has the unique status of being the most significant effect of a factor across all models. The model in which such effect appears is identified as the strong model in the row for that factor.

Average effect size was calculated per factor across all models. The average of these average effects is considered the exploratory average effect. Factors are considered to have an important effect size if they have an average effect which exceeds the exploratory average effect.

From a space of 58 factors, 24 factors had an effect which was vertically or horizontally robust. 6 factors had an effect which was both, and 3 factors were bidirectionally robust and important. Two of these three factors are industry effects. Industry 6 is information technology and Industry 9 is the military.

The other super-strong factor was a continuous variable generated from a count of the e-learning providers a respondent identified as familiar to them. This establishes a strong familiarity effect, where respondents exposed to online learning platforms tend to favor such platforms. This factor had horizontally robust negative marginal effect.

**Table 3 – Horizontally Robust Factors and Importance**



Income didn’t have much of an effect, whether assessed continuously or categorically. Income group 6 had the largest coefficient of all factors analyzed, but this group has a subsample of 1 and it is the outlier discussed in section II.I.

Q4 and Q6 had interesting effects. Q4 asks whether cryptocurrency is overall a good thing for society. This question is intended to capture a general attitude toward technology. Research has established both the effect of a pro-innovation bias[[4]](#footnote-4) and an anti-innovation bias, or status quo bias[[5]](#footnote-5). A positive coefficient on this factor is consistent with the literature.

Q6 captures anti-foreign bias with a positive linear effect. The positive quadratic and negative cubic effects are interpreted jointly as an eventually negative marginal effect, where some intermediate region has a non-negative marginal effect.

A key result of the survey is to show that individuals who make hiring and firing decisions tend to be favorable to alternative education, as do the unemployed. This opposes the notion that employers may constitute a bottleneck in the social normalization of alternative education. The fact that employers and the unemployed both favor alternative education indicates that the less favorable labor group is the currently employed.

STEM workers are more favorable of alternative education. Those who do not work in STEM or who do not know whether they work in a STEM field are less favorable. This effect exists after including industry effects.

-does it matter that you actually do work in a STEM field or that you think you work in a STEM field even when you don’t? reasons:

1. Psychology mechanisms including identity mechanisms: is a technology-positive mindset and view associated with stem?
2. Actual scholarship and regulations for STEM (by state, nationally?)

Additional surveys are planned to occur quarterly through at least fall 2018, likely through spring 2019, perhaps longer.

Some observations were p

Omitting the outlier observation reduced maximum explanatory power of the index variable from .49 to .44. Maximum explanation for Q2 was reduced from .59 to .42. Models identified by exploring either data set were forward-tested on the other data set. Table 2 summarizes model performance in either data set.

-Statistics from Provider Marketing Materials

-providers have detailed data sets on their own users which can be used to generate interesting metrics that occasionally generate financial insight. These data sets are generally closely held as proprietary, making many of these statistics questionable. They are nonreproducible and may be the result of cherry picking or selection bias. On the other hand, reputation mechanisms and other forces drive significant accuracy in these measures. There are also no known data which contradict the claims mentioned shortly. In any case, these statistics belong in a discussion which seeks to declare the state of knowledge on alternative credentials.

-linkedIn gets you more recruiter views or something

-degreed claims (data or no) to increase retention

-Degreed says points don’t matter and PluralSight says don’t use Skill IQ for employment decisions…but are they being too modest??

Can a cost-benefit analysis be generated based on these marketing data? What about other providers with somewhat substitutable systems? Can we say “Firms should be using these platforms” and proceed to “firms should have X data [because it comes with the platforms]”

An effect which is low-significance but robust is more useful than an effect which is low-significance and non-robust.

In Defense of STEM

I wanted to say the distinction didn’t matter, but it appears to be robust, important, and significant. The factor appears in maxar regressions, and removing it causes a notable drop in both raw and adjusted r-squared. I wanted to argue that STEM has such a broad reach (digital artists through scientists) as to become incoherent or low-meaning. I also wanted to point out that policies by state around STEM conflict in what they consider STEM. Nurses, for example, are considered STEM only in some states. Despite all of this, the data shows there are positives and they have important effect sizes.

Horizontal robustness:

1. Cross-model
2. Cross-sample
3. Cross-operationalization
4. Cross-study
   1. Cross-author
   2. Cross-method

Vertical robustness:

1. p-value
2. stable estimate as p-value threshold is moved up and down within one model

**III. Provider Data Study**

**IV. Some Case Studies**

**IV. Additional Notes from the Literature**

**IV. Some Practical Applications**

**Appendix A: Survey Questions[[6]](#footnote-6)**

1. Do you contribute to hiring and firing decisions at your company?
2. For many professions, alternative credentials can qualify a person for an entry-level position.
3. It will soon become fairly conventional for high school graduates to obtain alternative credentials instead of going to college.
4. When you add up the pros and cons for cryptocurrency, it's probably a good thing for society overall.
5. When you add up the pros and cons for online education, it's probably a good thing for society overall.
6. When evaluating an applicant's education, it is important is important to check whether the degree was awarded from a US institution.
7. Have you heard of any of the following online course providers?
8. Do you work in a STEM profession?
9. Which of these industries most closely matches your profession?
10. Age Group\*
11. Gender\*
12. Household Income\*
13. Region\*
14. Device Type\*\*

\*Responses included in purchased SurveyMonkey responses only

\*\*This factor was provided but structurally exempted from analysis

Unformated WIP content and notes below

* 1. theoretical and practical importance
     1. theory: signaling or human capital? Caplan basically answered this
     2. practical: financial consulting barely addresses education but it should; how should it do so? Also career and life planning. Gains to be had by starting earlier
        1. question: average age of first financial consulting or bank account? Connection to financial success?
  2. relation to the literature
     1. Previous papers calling for additional research and a dearth of data
     2. temporal changes as a need to update even the so-called established literature
     3. caplan’s work and other notable scholars in this field
     4. problematic conclusions in the field this far: govt spending bad and also failure to disaggregate and look at edge cases: even caplan looked at the mean and not so much the distribution nor a multi-specific approach by skill or industry (though he agrees it comes down to skills!)

1. outline
   1. Attitudinal Study
      1. Technical Description
      2. Results
      3. Related Material
         1. Stack overflow 2016-2018
         2. Gitlab and others
      4. Key Conclusions
      5. Future Work
         1. Repeat and make panel
   2. Udacity Study
      1. Technical Description
         1. Identification of Udacity in particular, and other key providers
      2. Results
      3. Related Material
      4. Key Conclusions
      5. Future Work
         1. Increase sample size and detail (what about nanodegrees?), look at other providers
   3. a series of concrete case studies to illustrate opportunity cost
      1. Fire Academy + EMT -> Paramedic -> other
   4. bottleneck diagnosis: employers, parents, students, teachers, or someone else?
      1. Students think it’s too hard or boring
      2. Parents and teachers think kids need to go to school
      3. Employers don’t realize the value

1. <http://www.afterecon.com/education/alternative-paths-traditional-education/> [↑](#footnote-ref-1)
2. For raw data and some additional comments, see: <http://www.afterecon.com/economics-and-finance/results-survey-education/> [↑](#footnote-ref-2)
3. Exploratory do files, raw data, and various spreadsheets are openly available at <https://github.com/Vandivier/data-science-practice/tree/master/stata/udacity-exploratory-analysis> [↑](#footnote-ref-3)
4. <https://en.wikipedia.org/w/index.php?title=Pro-innovation_bias&oldid=806646937> [↑](#footnote-ref-4)
5. <https://en.wikipedia.org/w/index.php?title=Status_quo_bias&oldid=829487998> [↑](#footnote-ref-5)
6. Raw results at: <https://www.surveymonkey.com/results/SM-FBQL8F5H8/> [↑](#footnote-ref-6)