**Research Program Overview**

May 2018

John Vandivier

**I. Overview**

This 2-page document summarizes research completed this semester and next steps planned for completion of dissertation work and for publication of academic paper(s).

Research is organized into a collection of projects. Each project is composed of data files, code, and plain-language analysis. Plain-language analysis is conventionally presented in a 2-pager format. Each 2-pager is composed of 3 sections describing what was done, why it was done, and what was found. 3-pagers are sometimes created to include additional notes for inclusion in a final paper and to include a section on recommended future work. 4-pagers are working academic papers.

This paper is organized into 3 sections. The first section gives an overview of the paper. The conventional content of a research project is described. The conventional contents of 2, 3, and 4-pager artifacts are also described. Section 2 lists current research projects and their related artifacts. Dependencies between some projects are noted. The third section describes planned next steps.

**II. Outline of Existing Research[[1]](#footnote-1)**

The current research program is focused on understanding the relationship between alternative education and labor outcomes, public disposition towards online learning, adoption bottlenecks, and the ability or inability of alternative education to substitute for traditional learning. Analysis is emphasized regarding online learning and the software development industry.

1. SurveyMonkey Baseline Attitudinal Survey
   1. This research identified whether individuals tasked with hiring and firing decisions within firms were particularly optimistic or pessimistic about online learning. It secondarily examined the general population on the same question.
   2. 2-pager-survey-monkey-1-off.docx
2. Udacity Scrape
   1. This research scraped data from public profiles of an online learning platform. Profile data included employment status. The effect of online learning on employment was estimated, and the employment rate among users was compared to general employment.
   2. Projects 3-6 are supplementary to this study.
   3. 2-pager-udacity-scrape.docx
   4. 4-pager-udacity-scrape.docx
3. Udacity GitHub Augmentation
   1. Portfolios simultaneously operate as credentials and as outputs of productivity. GitHub portfolios were compared against Udacity learning data to see whether either theoretical relation was observed.
   2. 2-pager-udacity-github.docx
   3. 3-pager-udacity-github.docx
4. Udacity LinkedIn Augmentation
   1. This study replicated analysis of the effect of alternative education on employment outcomes, using measures of employment derived from LinkedIn data rather than Udacity profile data.
   2. 2-pager-udacity-linkedin.docx
5. Udacity Kairos Augmentation
   1. Kairos is an image-based machine classifier for ethnicity, gender, age, and some other factors. This research integrated Kairos data with data scraped from Udacity user to improve findings.
   2. 2-pager-udacity-kairos.docx
6. Classifier Variance Analysis by LinkedIn Data
   1. This research measures comparative ethnic classifier accuracy by comparing self-reported ethnicity to machine estimates generated from LinkedIn profile data.
   2. 2-pager-udacity-classifiers-linkedin.docx
7. Classifier Variance Analysis by Self-Report
   1. This research measures ethnic classifier accuracy by asking respondents to use a classifier and asking them about their level of agreement with the classifier estimate.
   2. 2-pager-udacity-classifiers-survey.docx
   3. 3-pager-udacity-classifiers.docx

**III. Next Steps**

1. Run the attitudinal study several times over the next year or longer to generate panel data.
2. Classifier variance analysis which discovers the size of variance created by submitting different versions of a name to a classifier. For example, first-name-only, with or without initials, and with or without capitalization. This can be included under 2, 6, or 7.
3. Present working research to working dissertation committee to acceptability of the topic and identify areas for revision.
4. With committee, identify which, if any, projects are suitable for an academic paper apart from dissertation work.
5. Once current work related to scraping learning platforms and surveying individuals about learning platforms is settled, engage Dr. Robert Axtell regarding an agent based analysis for a second component of the dissertation.
6. Determine whether the two components discussed in III.3 are sufficient content for the dissertation, or whether a third topic is needed. If a third topic is needed, decide that topic.
7. Identify a third committee member. Interesting members include:
   1. Tabarrok or Cowen, because they may provide access to MRUniversity, an online education platform.
   2. Storr, because he oversaw a directed reading on Political Economy of Education Policy, brings an Austrian point of view, and was recommended to me by David Schmidtz.
   3. Many other possibilities, including someone from another department such as an education, psychology, or business department.

1. Research is generally open and accessible at <https://github.com/Vandivier/data-science-practice/tree/master/stata/udacity-exploratory-analysis> [↑](#footnote-ref-1)