

1. Us employment rate ~60%. So, udacity’s is less without correction
2. Simple regression on nnano1 shows insignificant negative correlation
3. Reg \_employed nnano1 nnano2 neg linear, positive marginal! “noob effect” but not expected to persist as n -> infinite (permanent increasing returns is theoretically problematic)
4. Reg \_employed nnano1 nnano2 nnano3 has more important, positive nnano3 compared to nnano2
5. Age matters; and I’m interested in alt creds as a step to the first job; so what about including age?
   1. Age highly significant and nnano1, 2, 3 pattern is robust, even better p!
   2. Including linear interaction flips the negative marginal effect, but reduces factor significance; nnano1, 2, 3 is now strictly positive but interaction is negative. What does this mean for young folks? It means getting a nanodegree is a worse idea as you age. This confirms intuition. But wait, couldn’t this effect have some marginal caveat? So let’s introduce quadratic and cubic effects to the interaction.
   3. Yes, now all interactions are highly significant. Now we see a clear but complicated picture. Getting nanodegrees as you age is linearly beneficial, marginally problematic over an intermediate region, and somewhat beneficial at the extremes. The cubic benefit might seem negligible, but note that it is roughly twice the cubic negative effect of age. If being super old per se is cubically bad, being super old and super education is cubically good to the exact same degree.
   4. The important part here is that the noob effect remains; although it’s expressed in quadratic and cubic variables. A single nanodegree doesn’t seem to help much, but several do help. This noob effect seems like ‘knowing just enough to be dangerous’ and it may signal justification for the imposter syndrome experienced by many professionals who switch to a career in programming utilizing self-teaching means or alternative education. But it seems to also support the notion that some who push through can achieve high skill.
   5. The pattern is robust to language effects; interestingly, speaking English has a negative effect…?

jesse7 is a case where two people were in the image, so I didn’t process it via Kairos. With large samples this could introduce a bias because people with a profile picture of the opposite gender may tend to be married or affectionate relative to others, and we know that married couples have different employment outcomes as do those of high agreeableness. Shane is a guy, but his profile pic was him and his son. I can tell because his son is too young to be working age, but his profile info indicates current employment. In theory I could discriminate by selecting the Kairos response with the largest age, but this is not a generic procedure which I can run across the code. I should still do it anyway.

Try dropping experience = 0 and see what happens to r2; it mixes effects of those never previously employed with those just too lazy to fill out profile

I always skipped with 2 or more people in pic for consistency

This noob effect seems to be an anti-sheepskin effect and thereby indicates human capital, not signaling. Instead of passing a marker society expects and repeating reward, you get what society expects and see a negative effect (eg one nanodegree or unit of education) but then you proceed to some arbitrary, peculiar, no-way-employers-see-this-point-as-a-sheepskin-threshold, and you get gains.

Private sector lit review:

<http://businessblog.udacity.com/2016/02/26/your-next-great-tech-hire-may-not-have-a-computer-science-degree-and-thats-a-good-thing/>

<https://news.ycombinator.com/item?id=9313088>

<https://www.inc.com/jessica-stillman/why-elon-musk-doesnt-care-about-college-degrees.html>

-review blog (google study on degree decipation)

-review reddit

-talk about ‘degrees of alternativeness’ vocational, private k-12, and charter are some of the ‘least alt alts’ and they work well; rank them in altiness; more alt is like homeschooling and what else?

Some github scrapes were taken by tracking the github profile down by hand; for example, Audrey Klammer had her linkedIn url twice within Udacity data, so I took this to mean she probably had a Github account and misentered it. So, I searched her name in GitHub users and found the correct GitHub url, then scraped it.

Has linkedin url and has github url alone ought to be a couple data points\*

Effect has known attenuation which is people that deleted their accounts, so it’s in the udacity json but not real

Make several papers:

1. SurveyMonkey 1-off Survey
2. Udacity scrape
3. Robustness testing with Kairos
4. Variance testing with NamePrism and others (footnote to Markus)
5. Github scrape
6. linkedIn scrape
7. Survey from above-scraped urls (udacity email survey)
8. Scrape github and survey (github survey)
9. Surveymonkey panel survey
10. Scrape stackoverflow
11. Company internal data extraction
12. ABM stuff
    1. Causal stripping: like take out gender. We know women have different personalities and preferences, but is this nature or nurture? As long as we don’t know, we do not create any function that says “women have this modifier, men don’t”
    2. We can then add the feature in question (women-specific modifier) and if some relationship of interest improves to reflect the real world this is confirming evidence. Generally, ABMs provide evidence for hypotheses this way: when you make some model variation and the model becomes more realistic we say the model variation is more plausible than it otherwise would have been, if never examined under an ABM
    3. Factor and relationship robustness in the face of feature-variation is the ABM equivalent of factor robustness testing under multiple regression models
13. Stuff with Markus
    1. Spokane\*…\*Need help from Caplan and perhaps Schmidtz about what to do here.
    2. Repec
    3. Multi-sample variance replication
14. Stuff with Ryan Turpyn
15. Stuff with Boettke, Cowen, Tabarrok, Storr, or whoever the last dude is (stratman???? But I think he didn’t get along w bryan)

Scrum Teaching – Agile Scrum Applied to Education

1. ABM study revealed that price was a key bottleneck; almost everyone consumed education if they could afford it
2. How can everyone afford it? If it’s free. Open source learning materials can solve this.
3. But, such materials are sometimes of a lower quality
4. But, this can present an opportunity for students; Who knows better how to improve such materials than the teachers and students that use them?
5. Student contributions to open materials can create a portfolio, giving them a leg up in career and college admissions
6. This respect for student opinions is scrum-like; what if teachers went whole-hog and engaged Agile Learning? The teacher taking the Scrum Master role and treating students as developers,
   1. Perhaps more like the teacher taking the Team Lead/Tech Lead role; a hybrid scrum master and senior developer who also teaches, does peer review, sets standards, etc
7. 2-week sprints would exist; students would have deliverables (contributions); retros would happen where the teacher would genuinely ask “what can I do to teach you better; what can we do together to build a higher quality product and execute with higher velocity?”
   1. The notion of a client here is interesting….perhaps every so often multiple classes could get together and vote on products
   2. Or, the real users of the open source materials can be called clients. The materials may be rated under some system of actual ratings, or metrics such as monthly downloads, views, references, etc
8. This would pair nicely with my Scientific Scrum stuff; I would set the stage for being a thinker in the Agile process space

Caplan discussion strategy:

1. Prepare 2-pagers for several microstudies, the weekend of 4/14
   1. SurveyMonkey 1-off Survey
   2. Udacity scrape
   3. Robustness testing with Kairos
   4. Variance testing with NamePrism and others (footnote to Markus)
   5. Github scrape
   6. linkedIn scrape (time permitting)
2. Target meeting on 4/19, after listening to his Conversation with Tyler
3. discuss questions for surveys
   1. Udacity Survey
   2. Github Survey
   3. SurveyMonkey Survey
4. Prepare a statement on Spokane + Ancestry Yearbooks study idea
   1. Ask what it is Caplan had recommended about looking into factors of college enrollment and completion, if I have that right? (Is it not already well studied?)
   2. Ask what else can be done with this data source that might be interesting
5. Ask about whether we can renew for the Fall
   1. And about whether ABM is needed for dissertation
   2. And if so, when should I turn my attention that way
   3. And about who should the third person be

Schmidtz discussion strategy:

1. Ask no more than 10 questions; 5+ are given below and should be enough, but a couple more can come.
2. Is he interested in working in econ of education more? His background doesn’t seem here
3. 2 questions are in Evernote. One is basically “Why don’t you act like your case is stronger? It seems strong to me.”
4. I can ask him these 3 also, similar to Caplan
   1. What questions should I include in my surveys
   2. What is Spokane data good for
   3. To what degree can I milk my microstudies, or must I merge them to be meaningful
5. I can also ask what econ of education topics, grants, and data sources he recommends
6. I can also ask how he would construct the ‘spectrum of alternativism in education,’ where charters don’t seem too alternative, and Nanodegrees and homeschooling are more alternative. What other practices and institutions does he see existing on this spectrum and where?