12-2014

An Examination of the Correlation Between Educational Attainment and Upward Economic Mobility Within and Without ‘High Tech Clusters’ in the US in 2012

Hannah Cecilia Oakley
hoakley@vols.utk.edu

Hannah Oakley

December 5, 2014

Honors Thesis Project

An Examination of the Correlation Between Educational Attainment and Upward Economic Mobility Within and Without 'High Tech Clusters' in the US in 2012

A relatively new book by economist Enrico Moretti titled The New Geography of Jobs encompasses Moretti's research on the topic of income inequality across cities in America. With the recent boom in the innovation sector, the demand for high levels of education has increased. Higher educational attainment, in turn, should increase wages and mobility. However, highly educated people tend to conglomerate, creating education segregation, as Moretti coined it. Since these highly educated communities are unique in their values and expertise, the ideas generated and products produced in that specific place are often hard to replicate in other cities. Moretti showed that high school graduates living in these more highly educated localities garner higher wages and salaries due, in part, to knowledge spillovers from interactions between skilled and unskilled workers. This presence of skilled workers reduced the overall effect of education on unskilled workers' salaries, meaning they needed less education to attain proportionate increases in salary.

I want to examine further if less educated workers do in fact benefit the same as better-educated workers in these innovation clusters. Considering the higher standard of living in most high tech clusters, I would think that less skilled workers gain much less from higher wages than their more educated peers. More people may be employed—which is an invaluable necessity for middle-class aspirations—but the larger issue of real upward economic mobility did not seem to be adequately addressed in Moretti's analysis. I want to take his research further by examining the effects of educational attainment on upward mobility, which I defined as the probability of moving from the bottom quintile to the top quintile. I then want to examine if that correlation changes when located in a high tech cluster. Moretti identified seven high tech clusters: San Jose, San Francisco, Seattle, Austin, Raleigh, Washington D.C., and Minneapolis.

In their Equality of Opportunity project, Raj Chetty et al. measured mobility of current thirty year olds (a single set of birth cohorts) as a function of parental income. They found that children from families below the 85^{th} percentile have better outcomes when relative mobility is greater in their area, meaning that location matters most for children from low income families. This would seem to support Moretti's theory that high tech clusters act as a rising tide to lift all ships. However, upward income mobility was significantly lower in communities with large African American populations.

The relationship between parental income and educational attainment in measuring relative upward mobility is more connected than I originally thought, according to the paper "Recent Developments in Intergenerational Mobility." Wealthier parents are better able to invest in their children's human capital. Wealthier parents also tend to have more education, and their children are often rewarded with higher earnings as a result. These findings illustrate that intergenerational mobility depends on parental means; those from more privileged backgrounds have more economic mobility than those with low-income backgrounds. Thus, to some degree, economic mobility is a function of privilege.

More than parental income and potentially even educational attainment, location affects a person's economic mobility prospects. The Pew 2013 Mobility Report found that the descendants of poor families living in low-mobility metro areas will take four generations to reach their metro area's mean income, while descendants of poor families living in high-mobility areas will only need three. Thus, focusing investment on these lower-mobility areas with better public services and access to quality education could improve opportunities.

I believe economic upward mobility will increase as educational attainment increases. I think counties with high mobility statistics will have higher levels of educational attainment, due in part to knowledge spillovers. I do not think that living in a high tech cluster will increase upward mobility for the less-educated adults. I am using the Equality of Opportunity data, which measured economic mobility across commuting zones in America in 2012. I will then merge that with a crosswalk (also from the Equality of Opportunity project) that matches commuting zones with counties. Lastly, I will merge the mobility/crosswalk datasets in Stata with my census data from 2012. My county-level census data includes the following metrics from 2012: educational attainment, median income, employment status, and poverty rates.

The literature I came across focused mainly on the correlation between income and mobility, while educational attainment was always considered as a secondary factor. I decided to focus my project on measuring the correlation between CZA-level mobility statistics (calculated by Chetty et al. from their Equality of Opportunity project) and county-level educational attainment data to further examine the direct impact of one on the other. I want to see if a correlation exists between level of education and real upward mobility in different counties throughout America. The mobility statistic I use is the probability of an individual moving from the bottom income quintile to the top income quintile as of 2012 (prob_p1_k5). The mean of this mobility variable is 9.09 percent, meaning there is, on average, a 9.09 percent chance of moving from the bottom quintile to the top across America. The bottom 10 percent of commuting zones offers only a 4.69 percent chance of moving from the bottom quintile to the top. Lastly, the top 90 percent of commuting zones has a 15.16 percent mobility statistic. The complete summary of this variable is below.

```
. sum prob_p1_k5, d

                          prob_p1_k5
-------------------------------------------------------------
      Percentiles      Smallest
 1%     .0281827       .0221016
 5%     .0379507       .0221016
10%     .0469857       .0221016       Obs                3129
25%     .0605109       .0221016       Sum of Wgt.        3129

50%     .0820753                      Mean           .0909584
                        Largest       Std. Dev.       .045957
75%     .1086637       .3571429
90%     .1515837       .3571429       Variance        .002112
95%     .1770658        .469697       Skewness       1.975757
99%     .2432432        .469697       Kurtosis       10.22434
```

The educational attainment statistics divide the total adult population in each county into the percent who have less than a high school diploma, the percent who have only a high school diploma, the percent who have attended some college or have an associate’s degree, and the percent who have a bachelor’s degree or higher. In all of

America, the average percent of adults with a high school diploma is only 34.7 percent. The nationwide average for percent of adults with at least a bachelor's degree is 19.6 percent. These nationwide educational attainment statistics are summarized in the chart below.

```
. sum PerAdultsLessHSDip PerAdultHSdipOnly PerAdultSomeAssocDeg PerAdultBachelorHigher

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
PerAdultsL~p |      3274    16.23155    7.489053   2.451472   55.12266
PerAdultHS~y |      3274    34.68881    7.031156   9.068703   70.73171
PerAdultSo~g |      3274    29.47712    5.494207   11.23358   48.81603
PerAdultBa~r |      3274    19.60253    8.732502   3.658537    72.7883
```

I merged the mobility measures dataset with a CZA-to-county crosswalk in order to later be able to match my county-level educational attainment data (from the 2008-2012 American Community Survey for adults 25 and older). The variables I focused on from that educational attainment data were the percent of the adult population with less than a high school diploma, the percent with only a high school diploma, the percent with some college or an associate's degree, and the percent with a bachelor's degree or higher.

However, my educational attainment data suffered from perfect multicollinearity. My right hand side education variables were the percent of the total adult population with a certain level of education. Thus, including all of them in one equation was impossible because together they summed to be 100 percent. A decrease in one variable meant an automatic increase to another. My only options at this point were to reinterpret my coefficients as a change relative to the omitted variable or choose a different functional form. Changing my functional form was necessary. Log-log specification—although not as readily intuitive—gets around the collinearity issue by taking the derivative of these variables and interpreting them essentially as percent changes.

Even still, after converting to a log-log specification, I had crippling multicollinearity. To salvage my regressions, I decided to only include the log of the percent of the adult population with a bachelors' degree or higher. This specification implicitly assumes that in the baseline (0% having a bachelor's degree or higher), everyone went to some college or less. Obviously, this is a shortcoming of my model, which will impact the intuitiveness of my coefficient interpretations and accessibility of my final results.

In addition to my educational attainment data, I also included some 2012 census data as controls. Although I am focused on the correlation between education and economic mobility, I need to include other factors to prevent unwanted omitted variable bias. I decided to include the county-level poverty rate and unemployment rate for 2012 as well as 2012 median household income.

Having all my data clean and ready, I thought the best way to measure and clearly illustrate the effects of a high tech cluster (denoted in my Stata output as 'HTC') would be to add it as an intercept dummy variable. An intercept dummy identifies the 7 high tech clusters with a 1, while all other localities are labeled with a 0. I wanted to determine if the presence of a high tech cluster implicitly raised mobility so that an individual was born into a higher level of mobility. Since my research question focuses on the correlation between educational attainment and economic mobility in and out of high tech clusters, I made my educational variables slope dummies to assess their impact on my mobility statistic, relative to my controls. Transforming my educational variables into slope dummies simply changes their slope when in a high tech cluster because, presumably, the effect of an individual's educational attainment on her upward mobility will differ within and without a high tech cluster.

In my final regression (which can be found in this paper's Appendix), I used the intercept dummy for high tech cluster (HTC) coupled with the slope dummy for the log of the percent of the population with a bachelors' degree or higher. I had to do this to avoid a truly devastating multicollinearity problem in my model. When HTC is 0 (when the locality is not a high tech cluster), if the percent of the population with a bachelors' degree or higher increased by 1 percent, the percent of those moving from the bottom quintile to the top quintile would decrease by 0.089%. This is statistically significant (according to its corresponding t-statistic and p-value). Thus, an increase education attainment actually decreases economic mobility. However, when HTC is 1 (located in one of Moretti's high tech clusters), the results are weakly statistically significant or merely suggestive, meaning that presence of a high tech cluster does not strongly boost educational attainment's effects on economic mobility, as I originally believed. Although barely statistically significant at the 10 percent level, the relationship between educational attainment and upward economic mobility within a high tech cluster is negative, meaning the probability of one moving from the bottom income quintile to the top quintile decreases as educational attainment increases.

I also included my control variables to avoid omitted variable bias. My control variables are also strongly statistically significant, with high t-statistics and p-values of zero, meaning unemployment and poverty rate are negatively correlated with mobility. So, if the unemployment and poverty rate increased by 1 percent, mobility would decrease by 0.067 percent and 0.03 percent, respectively.

My hypothesis was that education would have an outsized impact on mobility, but those with less education would not fare better in a high tech cluster simply because of potential knowledge spillovers and higher overall wealth. In his book, The New Geography of Jobs Moretti focused on their income level, which obviously grew in the presence of a high tech cluster. I wanted to see if that increase in income overcame the higher standard of living and translated to a higher level of mobility. My regressions have proven that living in a high tech cluster does not, in fact, significantly increase mobility for the lesser-educated baseline (with 0 percent having a bachelors degree or higher). In fact, within and without a high tech cluster, higher levels of education were negatively correlated with moving from the bottom income quintile to the top quintile.

In exploring why higher levels of educational attainment actually decrease mobility from the bottom quintile to the top quintile, I thought of many possible explanations. First, I wonder if my characterization of mobility as the movement from the bottom income quintile to the top quintile was too strict and narrow. Higher levels of educational attainment may have a different effect on different measures of economic mobility such as relative or absolute measurements. There is space for further examinations using different measures.

Another potential reason why higher educational attainment decreased mobility from the bottom to the top quintile could be that the cost of higher education is prohibitively expensive, especially for those in the lower income quintiles. So, those in the lower income quintiles are effectively barred from pursuing higher education, meaning only the already privileged are able to afford it. With the cost of higher education out of reach, there can be no upward mobility from the bottom to the top quintile due to educational attainment. In this sense, mobility is a function of privilege, and mobility exclusively for higher income quintiles is not necessarily something I value.

Another potential reason for my results could be that educational attainment cannot overcome other social and systemic factors that impact those in lower income quintiles. For example, our public school system does not work equally across income brackets. Funded primarily by property taxes, schools in wealthier sections of the country have greater access to quality resources, such as up-to-date textbooks, interactive learning technology, and their first choice of teachers. For children in lower income neighborhoods and schools, the public education experience can be unsatisfactory at best and discouragingly incompetent at its worst. A higher percentage of low-income youth receive a worse education than their more affluent peers, leaving them unprepared for higher education or so dejected they drop out all together. By addressing these systemic inequalities in our public school system, educational attainment could increase among our society's least privileged, in turn, setting them up for greater economic mobility.

Works Cited

Chetty, Raj, Nathaniel Hendron, Patrick Kline, and Emmanuel Saez. "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States." (2014). Web. 6 Nov. 2014. <http://www.equality-of-opportunity.org/>.

Moretti, Enrico. *The New Geography of Jobs*. N.p.: Houghton Mifflin Harcourt, 2012. Print.

Urahn, Susan K., Travis Plunkett, Erin Currier, Diana Elliott, and Sarah Sattelmeyer. "How Communities Factor into Economic Mobility." *Mobility and the Metropolis*. Ed. Alexandre Mas and Paul Jargowsky. Pew Charitable Trust, Dec. 2013. Web. 6 Nov. 2014. <http://eml.berkeley.edu/~bgraham/Published/PewMobilityReport_2013/PewMobilityReport_2013.pdf>.

Black, Sandra E. and Paul J. Devereux. 2011. "Recent Developments in Intergenerational Mobility." in O. Ashenfelter and D. Card, eds., Handbook of Labor Economics, Vol. 4, Elsevier,chapter 16, pp. 1487–1541.

Appendix

Percent of adults with less than a high school diploma, 2008-2012

	Percentiles	Smallest		
1%	4.609724	2.451472		
5%	6.819587	2.485339		
10%	8.108459	2.554028	Obs	3274
25%	10.64282	2.702703	Sum of Wgt.	3274
50%	14.59957		Mean	16.23155
		Largest	Std. Dev.	7.489053
75%	20.8406	46.55172		
90%	26.4107	46.6405	Variance	56.08592
95%	30.18592	47.34254	Skewness	.9509584
99%	38.79618	55.12266	Kurtosis	3.872123

Percent of adults with a high school diploma only, 2008-2012

	Percentiles	Smallest		
1%	17.13337	9.068703		
5%	22.48762	10.26033		
10%	25.45525	10.60897	Obs	3274
25%	30.0326	11.43418	Sum of Wgt.	3274
50%	35.02405		Mean	34.68881
		Largest	Std. Dev.	7.031156
75%	39.62557	54.87437		
90%	43.32553	54.90368	Variance	49.43715
95%	45.45157	55.10337	Skewness	-.2258408
99%	49.36897	70.73171	Kurtosis	3.166116

8% of the adult population in the bottom 10% of counties has less than a high school diploma. The top 10 percent of counties have 26.4% of their adult population with less than a high school diploma. The percent of adults with just a high school diploma in the bottom 10 percent of counties is 25.5%, which jumps to 43% in the top 10% of counties.

Percent of adults completing some college or associate's degree, 2008-2012

	Percentiles	Smallest		
1%	16.60959	11.23358		
5%	20.40434	11.24343		
10%	22.38893	12.03518	Obs	3274
25%	25.77705	12.46995	Sum of Wgt.	3274
50%	29.53912		Mean	29.47712
		Largest	Std. Dev.	5.494207
75%	33.04348	46.88716		
90%	36.61711	47.02093	Variance	30.18632
95%	38.63899	47.33542	Skewness	-.0025546
99%	42.10352	48.81603	Kurtosis	2.997656

Percent of adults with a bachelor's degree or higher, 2008-2012

	Percentiles	Smallest		
1%	7.670292	3.658537		
5%	9.705731	5.431983		
10%	11.02092	5.467091	Obs	3274
25%	13.69189	5.616606	Sum of Wgt.	3274
50%	17.44363		Mean	19.60253
		Largest	Std. Dev.	8.732502
75%	23.25129	60.49284		
90%	31.28251	63.19843	Variance	76.2566
95%	37.02741	71.24834	Skewness	1.509157
99%	49.14213	72.7883	Kurtosis	6.028416

In the bottom 10% of counties, 22.4% of adults have completed some college or have an associate's degree. That number increases to 36.6% in the top 10% of counties. The percentage of adults with a bachelor's degree or higher in the bottom 10% of counties is 11% and 31% in the top 10% of counties.

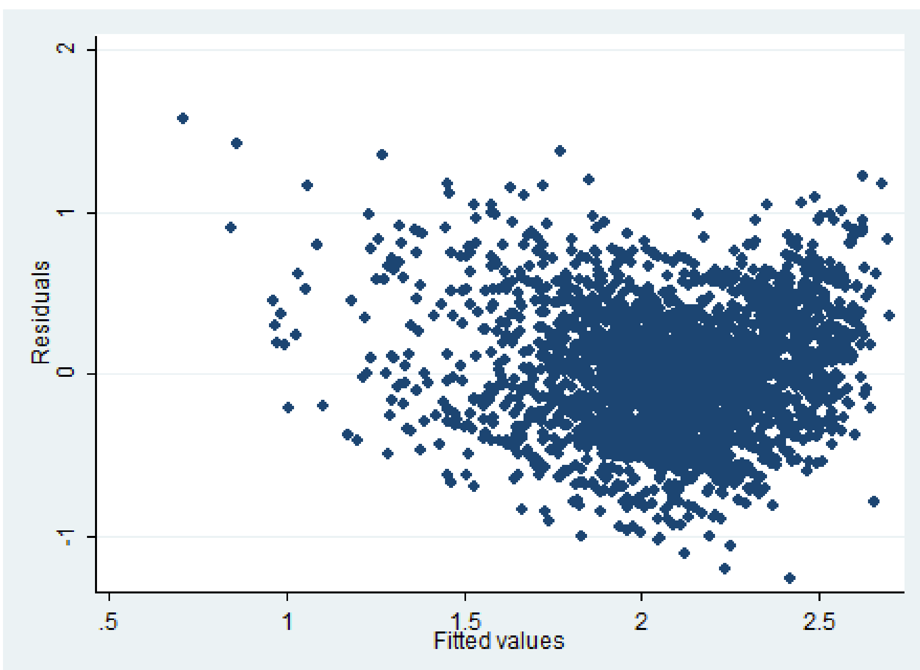
```
. sum ln_perc_mob ln_less_hs ln_hs ln_aa ln_ba Median_Household_Income_2012 Unemployment_rate_2012 PCTPOVALL_2012
```

Variable	Obs	Mean	Std. Dev.	Min	Max
ln_perc_mob	3129	2.10029	.4588358	.7930645	3.849503
ln_less_hs	3274	2.681985	.4660311	.8966888	4.009561
ln_hs	3274	3.523521	.2217968	2.204829	4.258894
ln_aa	3274	3.365217	.1957301	2.418908	3.888059
ln_ba	3274	2.890067	.406394	1.297063	4.287555
Median_~2012	3193	44948.44	11392.02	22126	121250
Unemployme~2	3272	7.90599	3.080865	.8	27.2
PCTPOVA~2012	3194	17.16941	6.539042	3.1	51.2

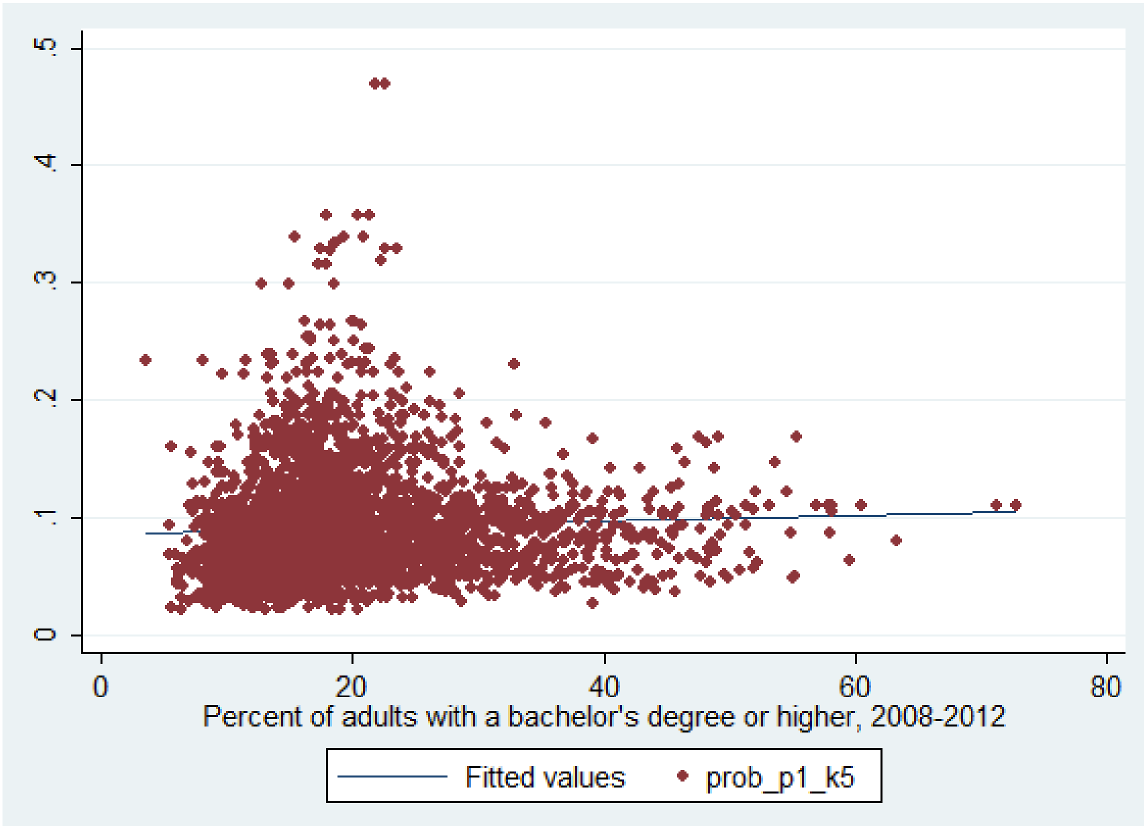
These are summary statistics for the variables I used in my final regression. Ln_perc_mob is the log of the percent mobility starting in the bottom quintile and moving to the top quintile. Ln_less_hs is the log of the percent of the population with less than a high school diploma. Ln_hs is the log of the percent of the population with only a high school diploma. Ln_aa is the log of the percent of the population with some college or an associate's degree. Ln_ba is the log of the percent of the population with a bachelor's degree or higher. Median~2012 is the median income in each county as of 2012. Unemployme~2 is the unemployment rate in each county as of 2012. PCTPOVA~2012 is the poverty rate in each county as of 2012.

```
. reg ln_perc_mob i.HTC#c.ln_ba  PCTPOVALL_2012 Unemployment_rate_2012 Median_H
> ousehold_Income_2012, robust

Linear regression                                      Number of obs =     3124
                                                       F(  5,  3118) =   198.07
                                                       Prob > F      =   0.0000
                                                       R-squared     =   0.3445
                                                       Root MSE      =   .37197

------------------------------------------------------------------------------
             |               Robust
 ln_perc_mob |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 HTC#c.ln_ba |
          0  |  -.0898595   .0234767    -3.83   0.000    -.1358909   -.0438282
          1  |  -.0457985   .0280881    -1.63   0.103    -.1008715    .0092746
             |
PCTPOVA~2012 |  -.0302694   .0021076   -14.36   0.000    -.0344018   -.0261371
Unemployme~2 |  -.0670134   .0040891   -16.39   0.000     -.075031   -.0589958
Median_~2012 |  -8.74e-06   1.24e-06    -7.05   0.000    -.0000112   -6.31e-06
       _cons |   3.784973   .0837706    45.18   0.000     3.620722    3.949225
------------------------------------------------------------------------------
```

This is my final regression.



The plot of residuals against fitted variables shows no discernable trend, meaning that my model does not suffer from serial correlation. This is obvious, however, because my data is not time series data, which makes a Durbin-Watson test irrelevant.



This is a scatter plot of the percent of adults with a bachelor's degree or higher against my mobility statistic, the probability of one moving from the bottom income quintile to the top quintile. As you can see, the correlation between the two variables is not very strong.