北京大学

硕士研究生学位论文

题目： 经济差的时候毕业：维吉尼亚高等教育毕业年组的收入差别

Graduating in a bad economy: cohort differential wage effects on Virginia higher degree holders

姓　　名： Jeremy M.W. Schutte

学　　号： 1401213315

院　　系： 汇丰商学院

专　　业： 西方经济学(E)

研究方向： None

导师姓名： Christopher Balding

二〇一六年 六 月

版权声明

Contents

中文摘要

经济差的时候毕业：维吉尼亚高等教育 毕业年组的收入差别

作者：Jeremy Schutte

指导教授：Christopher Balding

摘要：

我利用美国 维吉尼亚州高等毕业生的收入收入数据来证明在经济情况差的时候毕业会对首先几年的工作收入有不良的影响。

我的研究的原始性在于我怎么用，还有处理它的数据。第一要解释的是原来的数据里面所有的观察代表了一个毕业组一年收入的分布。毕业组分毕业的年份还有毕业证的类型包括：两种 准 学士学位的毕业证，四年、还有五年的本科还有本科后的毕业证，所师还有所师后的毕业证，博士的毕业证，还有专业学位的毕业证。

每一组的每一个观察的分布只有三个部分：第 25,第 50 和 第 75百分位 间的数据。因为大部分收入的分布是有log-normal分布的分配，我用我发现的一个方法来利用第 25 和第 75 百分位间的数据来 模拟每一个分布代表的观察。我第一次做这个就模拟 4.6 百万多个模拟观察。我做了很多研究之后再发现额外的数据，加进去之后就有 5.3 百万多的模拟观察。为了证明模拟观察的可靠性还有不怎么影响我研究结果的有偏 性，我先用第 50 百分位间的数据来估算每一个分布的错误。我的结果证明我做的模拟的错误是极小的，也最大的让我低估不良经济指数在一个毕业组毕业的时候对它将来几年减少它的收入的结果。另外，从模拟分布的观点来看，我做得很好。

我利用本来的分布观察数据、还有这些模拟的观察数据，通过几种方法来证明毕业时的经济指标有了非常 显著 性的 统计，还有经济重要性的影响。这些指标 包括美国的失业率与GDP发展率，还有维吉尼亚的失业率与GSP发展率。

我基本上所有的研究结果是证明我的一个理论: 毕业时差的经济指标会对首先几年的工作经验有一定的影响。我另外的论点没有得到很多支持，是：毕业时差的经济指标会对长期的收入也会有一定的，规模小一些的（比首先几年的工作）影响。

第一个方法是用原来的分布观察，还有一种多变（其实每一套 预测指标都 通过十万遍。）的 随机选择 训练集和测试集做回归来看那些预测指标对测试集有最强的预测能力。结果是所有的经济指标（包括毕业的那一年，前一年，还有后一年的指标）都有预测能力，而且加了越多经济指标对测试集的MSE(平均值的标准误差)就越小。

第二个方法就是利用模拟观察数据来分析毕业时经济指标对所有的工作经验有多大的平均影响。附录 表 1 至 3 有了 24 个回归使用我所有的经济指标还有其他的重要预测指标来估计这个。

第三个方法就是分开看不一样的毕业证拥有者从毕业时经济指标收到多大的影响。附录表 4 至 7 有了 32 个回归 （其实他们代表了 3200 个回归，看上面的解释）也是按毕业证分开做最

有趣还有最能容易来理解的回归。除了专业毕业证还有技术性副学士学位的拥有者,其他的毕业证的拥有者都是受到预期的影响。前面的两种毕业证的拥有者在这些回归都收到预期影响的相反反应。

第四个方法是看 2000 年的毕业生还有 2008-2011 年的毕业生有没有更严重的收入损失。这些组都是在经济很差的时候毕业，所以应该对他们首先几年赚的钱有更大的影响。他们确实受到了比较大的影响，不过我做这一步发现大部分的损失是通过没有全职工作的途径来减小他们的收入。我后来用的数据能让我看一个分布里面有多少人是全职的、有多少是兼职工作人员。

我最后两步是分开看毕业后首先五年还有五年后的观察来计算毕业后经济指标对收入有多少短期还有中期的影响。就像其他的每一步，毕业时的经济指标对短期收入有一定的影响。它对中期的收入也有不良统计显著 性的结果，不过经济重要性没有多少。

最后，我想提到两件事。第一是上面都是我自己用中文写的，我只请我老婆修改一下某个字和标点符号。第二的是我做出来的研究不单单在于这文章里面；其实大部分的功劳用在R的统计程序里面好好安排跟利用原来的数据。我在R里面最后用的代码有差不多 1500 行，而且做的过程中应该最少写了 5000 行代码。我也为了这个论文学习怎么用Latex来做出一个更漂亮的、整理得更好的研究。

关键词：劳动力市场进入，失业率，毕业生，GDP,GSP,高等毕业证，维吉尼亚，收入分布，log-normal模拟

Graduating in a bad economy: cohort differential wage effects on Virginia higher degree holders

written by Jeremy Schutte, directed by Christopher Balding

English Abstract

I use Virginia wage higher degree holder wage data to show that graduating in poor economic times has negative effects on the first few years of wages and on later wages.

The originality of my research lies in how the data I used was utilized and manipulated. The first thing to explain is that in my source dataset, one observation represents a year's distribution of earnings from a given degree cohort. Degree cohorts are divided according to graduation year and the type of degree they hold, including less than one year post-high school certificates, one to two year post high school certificates, two types of associate's degrees, four, five, and post-four year bachelor's degrees, master's and post master's degrees, doctoral degrees, and professional degrees.

Every cohort yearly observation was a distribution with three descriptors: the 25th, 50th, and 75th percentile earnings for the yearly wages of this cohort. Since most wage distributions are close to log-normally distributed, I use a trick I learned through much forum crawling and internet searches for creating log-normal distributions when the 25th and 75th percentile of the distribution are known to create simulated distributions for each cohort-year original observation. The first time I do these types of simulations, I simulate roughly 4.6 million wage-year observations for Virginia higher education graduates. Later after performing much of my analysis, I discover an expanded dataset that allows me to simulate a total of roughly 5.3 million wage-year observations, which I use for many later regressions. In order to demonstrate the reliability of the simulated data and to show it doesn't bias my results to any significant degree, I use the only measurement of error I possess, the 50th percentile of a given distribution, to measure the lack of fit of the distribution. My results indicate that the bias from the simulations is quite small, and that it works in the opposite direction of my results, which may be underestimated as a result. I also believe that the simulation of the distributions was done well.

I use the original dataset of distributional observations and the simulated set of observations to carry out multiple methods to prove that poor economic indicators at graduation lead to statistically and economically significant reductions in income after graduation. These indicators include U.S. unemployment and GDP and Virginia state level unemployment and GSP.

Almost all of my results support my main hypothesis: graduating in poor economic times has a definite effect on Virginia higher degree holders' first few years of income. My other hypothesis, that graduating into a poor economy has longer term effects on Virginia higher degree holders' income, also receives some support, but the effects are of a smaller scale.

The first method I use is to perform multiple (actually 100,000 times for each set of predictors tested) regressions from randomly chosen training and test set of observations from the original (not simulated) dataset to see the predictive power of a

given set of predictors (here I measure predictive power as the mean standard error using the training set fitted regressions predict the withheld test set of observations). The results of this process is that each of the graduation economic indicators in question (including lags and leads of one year) carry predictive power, and that adding more of them adds to the predictive power of a given model.

For the second method, I use the initial simulated dataset to estimate the average (across all available work years) effect of a given economic indicator at or around graduation on future wages. Appendix tables 1 to 3 contain 24 regressions using many economic indicators and other predictors to estimate these effects.

The third method is to analyze the effects of various economic indicators at graduation separately by degree type. The rest of the appendix tables are the result of this step, but of special interest are appendix tables 4 to 7, which represent 3200 regressions aggregated into 32 regressions in a manner described in the individual degree regressions section. They are also the most interesting and easy to explain of the individual degree regressions. Except for professional degree holders and possessors of associate's degrees of a technical nature, all individual degree regressions have the expected average reactions to economic indicators at graduation. The previously mentioned two degree types economic indicator estimations move in the opposite direction as expected.

The fourth method is to see whether year 2000 and year 2008-2011 graduates suffered particularly large losses. Cohorts of these graduation years graduated into particularly bad economies, so their first few years of wages should have suffered appreciably. They really did suffer worse than other cohorts, but from these regressions I found the primary mechanism of lower earnings was lower full time employment. It was only the expanded dataset I found close to the time of finishing this research that let me know the full time employment percentage of a wage year of a given cohort.

The last two steps I took were to separately consider observations from only the first five years after graduation and from only more than five years after graduation to see the short term and longer term effects of poor economic indicators around graduation. Just like every other method I used, this showed that income is definitely affected in the short term by economic indicators around graduation. I also found statistically and economically significant effects on wages more than five years after graduation from bad economic indicators around graduation, but they were not as severe as the effects over the first five years.

Finally, I want to mention two things. The first is that I wrote the Chinese abstract myself with very minimal editing from my wife. The second is that the research shown in this thesis is not nearly the amount that I put into it, since most of the effort took place in the R programming environment. My final lines of code in R totaled about fifteen hundred, and I must have written at least five thousand lines of code to get those fifteen hundred. I also learned how to use LaTeX for this paper so I could have a more eye-pleasing and better organized thesis.

Keywords: Labor market entry, unemployment rate, graduates, GDP, GSP, higher education degrees, Virginia, wage distributions, log-normal distribution simulation

1 Introduction and motivation

A well researched topic in labor economics is the effect of a weak economy on recent graduates — both on the immediate impact of finding a job and on the long term impact on earnings. The economy one graduates into is almost completely outside of the control of the individual, but strongly impacts one's long term job prospects and wages. This is most true over longer periods of history or time[1], but even within generations, the timing of graduation can make a difference to one's current and future earnings.

I use Virginia wage data, with observations representing distributions of wages by year, graduation year, and degree of higher education, along with measures of economic activity [2] to estimate the dollar impact of a weak economy on short and medium term earnings. This data is of particular interest because it represents yearly data and over four million total wage observations. [3] [4]

It seems from a review of the literature that most related studies estimate lost earnings indirectly, for example from estimated lost earnings for those who are unemployed. Fewer studies address the impact on employed workers. I add to this smaller literature by using the aforementioned data to estimate negative differences in earnings per extra percentage of Virginia or U.S. unemployment. Using Virginia Gross State Product growth or U.S. GDP growth as the proxy for economic health yields an increase of earnings per year per extra point of GSP. This wage gap for those who graduate into a time of higher rates of unemployment or lower rates of growth varies across degree type and over years of work experience but usually moves in the expected direction. Actual wage rates are relatively stable over time when controlling for years of experience, degree, and inflation, which suggests that there has been no real wage growth for all Virginia holders of higher degrees over the preceeding fifteen years. However, this may be largely due to the recessions following the bursting of the tech bubble at the end of the 90's and the housing bubble in 2008.

By assuming that each observation-distribution is log normal, I am able to simulate the individual observations, which creates quite a large dataset to work with. Most similar studies do not have access to such a large number of observations, so I am able to get very small standard errors on most important predictors of interest, and an analysis of the misprediction of the simulated data shows that the bias on the predictors of interest is such that my estimates for economic indicators is, if anything, a slight underestimation of the results that would occur if the simulated distributions were a perfect fit.[5]

Theoretical explanation for the wage gap

When the unemployment rate is high, marginal workers are deprived of the ability to improve human capital in two important ways: firstly, they are unable to work and accumulate experience in a job suitable to their studies, since there are less of such jobs available. This time spent unemployed means that when they are able to obtain employment, they have less accumulated human capital and are paid less

[1] Of course in the long run, the level of capital accumulation and technological development within an economy has an enormous impact on wages.

[2] These include the US and Virginia yearly unemployment rates, as well as the US GDP and Virginia GSP (Gross State Product) yearly growth rates

[3] My initial dataset has 1592 observations, but each observation has the median, 25th percentile, 75th percentile wage, and number of workers that represent a graduation year - wage year - degree type distribution.

[4] Most of the regressions in this paper were performed with this dataset. Shortly before finishing this paper, I was able to access additional wage data for an additional three degree types, which were used in some later regressions and other descriptive graphs.

[5] However, from what data I have, the distributions seem to be a very close fit.

than others in different time periods with similar educations who were able to quickly secure work. This depresses wages of their cohort in the medium and perhaps long term.

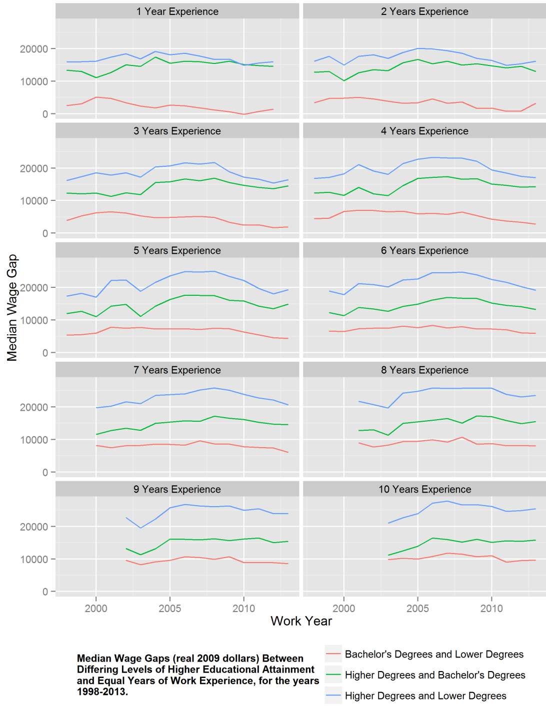
There is a secondary effect for marginal workers who do obtain employment: they are often underemployed relative to jobs they could have secured in a better economy. This both contributes to the lower initial wages of their graduation cohort (after all, these workers are underemployed), and it also contributes to lower medium and perhaps long term earnings gaps, again because of the lower accumulation of human capital.

Since this study's data only provides observations for employed workers, it is likely that any estimates of negative economic indicators on earnings represent a lower bound to the actual effect. The reasoning behind this goes as follows: the lost potential human capital from being unemployed, as discussed above, is not reflected in the data to the extent that some of those previously unemployed in each period continue to be unemployed. When they finally enter the workforce, their lower earnings will further increase the estimation of lost earnings. This means that for cohorts with many years of work experience, the estimated effect is likely to be close to the true effect (although it will still be somewhat understated since regressions include observations for all levels of experience, including the first few years). The estimated wage differentials are likely to be understated for cohorts with only a few years of work experience. [6]

The wage gap over time

A tangential question is whether the wage gap between those with different levels of education has been increasing over time, as previous studies indicate happened in the 70's, 80's, and 90's. My data shows, that at least for Virginia wage earners, the earnings gap between differing levels of *higher* education has remained relatively stable over the last decade and a half, as you can see from the following graph. Although there is no clear trend, it seems that wage gaps for most workers with higher degrees have declined in recent years.

Figure 1: The wage gap over time



Unfortunately, it is not possible from this data to estimate the change in the wage gap between those with higher education and those with high school education or less, since it only reports wages for those with degrees past high school. Most research and news around this issue shows that the wage gap between those with just a high school diploma and those with a college diploma has been widening for many years.

Another question of related interest is whether there are any standout cohorts — ones

[6]However, this estimated effect should shrink over time, if the effect of graduating in bad economic times is most felt close to graduation as I assume it does.

that earned a substantial income premium or had a substantial income loss because of the year they graduated in. While there is certainly the effect mentioned above (a couple hundred dollar difference per year per extra year of unemployment at graduation) for most degree types, there are no standout cohorts, with one lone exception. [7]

Unfortunately, the dataset only covers 16 years of wage data and only has information for graduates from 1993 on, so the results are not as robust as a longer-term dataset would reveal. The results are also sensitive to the "Great Recession" in America, and only a few years' worth of observations are available to estimate the effects of this downturn on future earnings. Revisiting this data in five years time would likely yield a much different picture of the results of the recession. Furthermore, the earnings data is conditional on employment. This suggests an attenuating effect on the effect of unemployment on earnings, especially for those cohorts who have recently entered the labor force, since it is likely that those who are employed in a recession are, on average, higher quality workers than those employed in better economic times. If true, this would mean that the estimates from this study would be a lower bound for the effects of economic indicators on short and medium term earnings. Fortunately, the dataset will supposedly be updated in the foreseeable future, and so this study could hopefully be revisited at a later time.

[7] The 1998 graduates of a less than one year certificate program earned substantially more in median and 75th percentile earnings than any surrounding cohort, and this effect persisted for more than a decade. Perhaps there was a special training program that year, and perhaps it was merely the effect of graduating at the top of the tech bubble with a certificate such as a Cisco systems' administrator certificate when demand for such jobs was the highest. Whatever the case, the 1998 graduates represent somewhat of an anomaly in the wage data. You can very clearly see this standout cohort in the "Summary data in graphs" section of the final appendix.

2 Literature review

There are many studies that touch upon variations of wages and employment, both long and short term, due to various economic shocks, including demographic shifts, recessions, and differences in immigration rates.

Various job market conditions can have an effect on the initial and lifetime earnings of a cohort of workers. One such condition is a substantially larger cohort than previous ones can lead to a reduction in relative wages between the two, for example between older workers and job market entrants across differing time periods. Welch (1979) finds that as the post-World War II baby boom cohort entered the labor force from 1967-1975, relative wages between job market entrants and older workers declined significantly. The primary reason for this is probably the relative decline in capita, both directly lowering income from lower capital per worker and indirectly by limiting new business opportunities because of the relative lack of startup capital.

It seems that a similar effect should be expected in the event of a recession. Following the work of Ludwig von Mises or virtually any modern Austrian economist, during a recession it is revealed that due to distortion of interest rates caused by the central bank[8], many investments that were made at previously low interest rates are revealed to be highly unprofitable at the readjusted interest rates. Or in other words, the amount of actual capital available in the system is seen to be much lower than it was previously thought. Businesses, realizing something closer to the

[8] Some Austrian economists maintain that absent a central bank or lender of last resort, a free banking system would result in a lending market in which entrepreneurs were able to make appropriate capital investment decisions. This author believes that so long as fractional reserve banking exists, there will be a distortion of interest rates, but that the mal-investment under a system with no lender of last resort would be substantially less, since banks would be much less willing to lend at low reserve ratios due to the potential for bank runs.

true value of their capital[9], are much less likely to hire workers or engage in new ventures.

Other papers have analyzed the effect of recessions on employment and long term earnings. Oreopolous, Till von Watcher, and Heisz (2012) find that Canadian college graduates who graduated during a recession in the 80's suffered significant earnings losses that faded after 8 to 10 years, and which were especially severe for college graduates from lower tier schools. They find that these lower wages can be mainly explained by an initial reduction in employer quality. Kahn (2010) finds large, negative, and persistent effects on wages from graduating during bad economic times using the National Longitudinal Survey of Youth. I also find a significant but mostly short term effect on wages from graduating in bad economic times.

Elsby, Hobijn, and Sahin (2010) clearly show that employment rates drop during recessions, but claim that the most recent American downturn in 2007 may result in long term unemployment for a subset of workers. They briefly touch upon the fact that younger workers face greater unemployment during downturns than other groups.

Farber (2011) found that one in six workers lost work during the '07 to '09 downturn, and that there were very low rates of reemployment and large earnings losses. Because of a lack of but a few years of wage data, it is too soon to see the long term effects of workers displaced by the "Great Recession", but a future comparison of those who entered the work force more recently to past cohorts may find substantial long term negative differences above and beyond those present in Farber's study and this one.

[9] It is true that during a recession, pessimism drives the price of various forms of capital below its long term average value. However, during the preceding boom, the distorted interest rates and excessive optimism drives it far above its long term average value.

Some papers have estimated the loss of lifetime earnings resulting from layoffs during recessions. Davis and Wachter (2011), using data from 1974 to 2008, claim that 2.8 years of pre-displacement earnings are lost when mass layoffs occur during recessions with unemployment levels of 8% or above vs. 1.4 years of lost pre-displacement earnings when such layoffs occur during times of unemployment of 6% or less. For reasons discussed in the introduction, the lost earnings conditional on employment are quite possibly underestimated by this paper, as it is not possible to account for the unemployed with this data. This paper makes no estimate of the amount of time spent underemployed or unemployed because of bad economic conditions at graduation, and so only wages conditional on some kind of employment can be estimated. This means that the total average losses from graduating in poor economic times are likely to be largely underestimated when only considering the coefficients on the predictors of interest in this paper.

Many papers have addressed the issue of immigration's effect on native workers' wages. Borjas (1999) finds that a 10 percent increase in labor supply from immigration lowers the wages of competing native workers by 3 to 4 percent. Using a more thorough methodology, Ottaviano and Peri (2011) find that immigration increased the wages of American natives, but lowered the wages of already in place immigrants. Both papers examine somewhat exogenous shocks to the labor supply, as does this paper, since it covers two recessions.

Cutler, Huang, and Lleras-Muney (2014) show that graduating in poor economic times not only results in lower incomes over time, but also in lower health outcomes. They find that the higher the level of education, the lower the negative effects of graduating into high unemployment on health are. This provides another mechanism by which medium and even long term income could be influenced by graduating into a poor economy. [10]

Genda et al. show that there are different effects of entering the labor market during bad economic times for job entrants, and that the persistency of these effects varies by both education level and country. In Japan, they show that those with low education suffer persistently lower wages because of poor economic conditions when entering the labor market. In America, they show that the negative effects of entering the labor market for entrants with low education on earnings quickly disappears with improving economic conditions, but that the negative effect for more highly educated cohorts is more consistent. I did not have time to measure the differential effects from graduating in a poor economy on short term versus long term wages by degree level in this study, but that is a possible extension that could be done with a little more time.

3 The data

The primary source of data for this study was from SCHEV's Guide to the Mid-Career Wages Report, specifically EOM 18. EOM 18 provided yearly wage data for Viriginia degree holders. The wage data was from 1998-2013, and it was for graduates from 1993-2011. It further subdivides observations according to degree type. [11]

The degree type notation I use in regressions and tables in this paper are as follows:

LT1YRCert Those whose highest degree is a less than one year post-high school certificate.

1or2YRCert Holders of a one to two year higher education certificate.

AssTech Holders of an associate's degree with technical degree credits.

AssBach Holders of an associate's degree with Bachelor's degree credits.

4yrBach Holders of a four year bachelor's degree.

Mast Holders of a master's degree.

Doct Holders of a doctorate level degree.

Prof Holders of a professional degree.

Each observation (there are 1592 of them) represents a distribution of yearly earnings for a cohort — here I am defining a cohort as the wage earners who graduated in the same year and hold the same degree type. Hence, there is a maximum of 16 observations per cohort. 16 for those who graduated from 1993-1997, and one less per year for those who graduated after 1997.[12] The distribution only contains the 25th percentile, Median, 75th percentile earnings, and the number of wage earners it represents.

Because each observation represents an entire distribution of earnings in and of itself, the data actually represents 4,608,775 observations of yearly wage data. I simulate these observations in my regressions, as discussed in the section on data simulation.

In some regressions and tables, I used an expanded dataset that covered the following additional degree types:

5yrBach Holders of a bachelor's degree that took five years to complete.

PostBach Holders of a post bachelor's degree, but not clearly defined.

[10] If these health effects are sizable and persistent, they may even end up becoming the primary determinant of lower wages that come from graduating into a poor economy over the medium and long term. Unfortunately, my study has no measures of health outcomes, and even if it did, it might not be appropriate because of its complete lack of data for those with a high school diploma or lower.

[11] SCHEV's Guide to the Mid-Career Wages Report also provides discipline area level observations as well as job type, but I did not use these as part of most regressions.

[12] This is because there is only wage data through 2013. Hence, 2011 graduates have at most two years of wage data, 2010 graduates have three years of data, and so on.

PostMast Holders of a post master's degree, also not clearly defined.

Including these degree types adds 707526 simulated observations, as well as an additional approximately seventy five thousand simulated observations for the original eight degree types. [13] I also collected yearly US GDP data from the U.S. Bureau of Economic Analysis (BEA), yearly US unemployment data from the U.S. Bureau of Labor Statistics (BLS), and yearly Virginia unemployment and GSP (Gross State Product) data from the Virginia State website. I created one of the variables, Exp (for work experience), by just subtracting the years since graduation in a given work year for a given cohort. Of course, this is a proxy for work experience, and it surely contains measurement error. Specifically, it should overestimate the experience of many workers, which wouldn't be as much of a concern if it was an equal overestimation. Instead, it most likely overestimates years of experience for those who graduated into times of high unemployment. It would seem that such an overestimation would lower the residual wages (after accounting for experience) among graduates of poor economic times, and perhaps understate the effect of poor economic indicators at graduation on short and medium term wages. To account for this, the first sets of full regressions on the simulated dataset omit my proxy for experience entirely, to see whether the magnitudes and directions of the coefficients still support the economic argument put forth in this paper.[14]

Data simulation

Wages usually follow a more or less log-normal distribution. There is a simple rule of thumb to check this: Equal spaced percentiles from the median, for example the 75th and 25th percentiles, should share the same inverted ratio. For example, the ratio between the 75th and the median, and the median and the 25th should be approximately equal in a log-normal distribution.

An equal weighted (by the observations each distribution represented) ratio of the 75th to the median, as well as the median to the 25th wage for my entire dataset was 1.417 and 1.362, respectively, so I believe that fitting the lognormal distribution for each represented distribution is appropriate. The following section also addresses this issue in much more detail.

There is another neat trick with the log-normal distribution if you have equal spaced percentiles from the median. You can use them to directly calculate the log-standard deviation and log-mean, which gives you everything you need to simulate the log-normal distribution. [15] I used this information to simulate a full dataset of 4,608,775 observations by creating 1592 log-normal distributions[16] and drawing log-normal wages from them according to the number of workers within each distribution.

Wage distribution simulations potential problems

Whether the simulation of the distributions has a bias on the estimates from the main regressions of this paper is a critical issue. The only way to measure the error of the distributions is to look at the

[13] I would have used all of these observations for all regressions and tables in the paper, but received the updated dataset rather late, and did not have enough time to update all of the previous regressions and tables. I do use the updated dataset as a robustness check and for other additional regressions, and find no real evidence that my results are significantly changed by using the smaller vs. larger dataset.

[14] The directions of the economic indicators at graduation on regressions with no control for experience are correct, but the magnitudes are greater by several degrees, showing that omitting experience is a serious problem. Unfortunately, I could not think of good ways to correct for this overrestimation of true experience, but I don't think it is too much of a problem as it should bias the results in the opposite direction that I am trying to show.

[15] I only discovered this trick after digging through dozens upon dozens of forum posts discussing similar issues of fitting distributions to a paucity of data points.

[16] Later I followed the same procedure to simulate the larger represented dataset with additional degree types of 5,391,219 observations from 2178 log-normal distributions

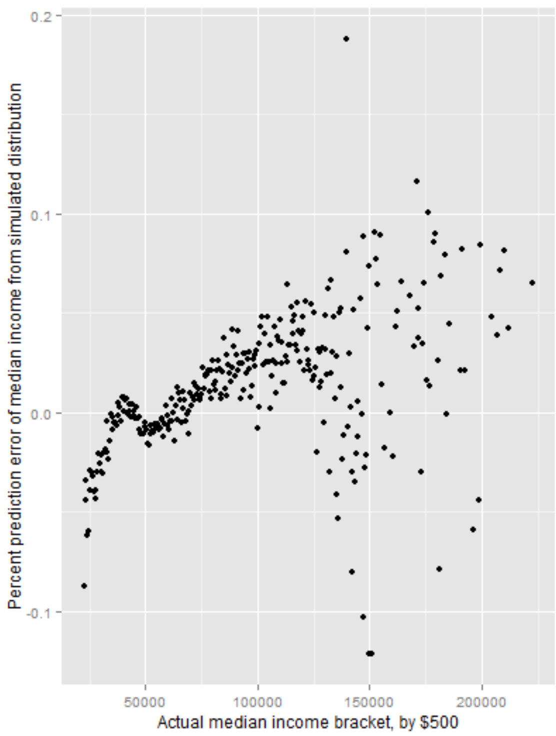
predicted median earnings from the log-normal distribution created with the 25th and 75th percentile earnings and compare it to the actual median earnings. This is because there are only three data points for each wage year and graduation year cohort: the 25th percentile, median, and 75th percentile. If the error from simulating the data has a large bias, it would call into question the central findings of this paper.

The first question is how good the simulated fit is to the median income. On average, the predicted mean incomes from the log-normal distributions are almost a perfect fit with the actual median incomes. However, averages usually cover up the full story, and so it is with income distributions.

There is a more or less linear relationship between reported median income and the predicted median incomes from the fitted log-normal distributions. For low actual median incomes, the predicted median incomes are too high, and for high actual median incomes, they are too high, as can be seen by the graph above (which bins actual median incomes and finds the weighted average among each income-bin). The actual errors are quite small (on the order of -0.1% to 0.1%), but show a clear early pattern and grow with income, as can be seen by the accompanying figure.

A simple regression on the percentage deviation from predicted median earnings on actual median earnings shows real median earnings to be higher than predicted by 0.6% for every 10,000 dollar increase in real earnings with a t-value in the high twenties. This shows that the earnings distributions of higher earners are more right skewed around the median than the distributions of lower earners, a rather intuitive result.[17]

Figure 2: Distributional error (measured at the median) across median incomes



If the relationship between the distribution of earnings around the median and simulated median earnings holds during bad economic times, it would mean that many of the simulated data points around the median would be biased upward in relation to good economic times, since lower wages in general would lead to less underprediction of high median wages.

This bias would tend to underestimate the actual effect of bad economic times at graduation on future wages, if anything, so this does not seem to be a big concern.

So the question is whether the fit of the simulated distributions are sensitive to economic indicators at time of graduation. Here, I run a regression on the percentage prediction error versus the economic indicators at graduation. As a control variable, I use the percent of full time (versus part time workers) reporting, since this will shift the distribution of earnings to

[17] There is a general perception that high wage earners have even more relative dispersion of wages than other workers; this suggests that the highest performing workers are compensated at an even higher percentage rate over their coworkers versus lower wage earners. This provides some support for this idea.

the right. With this control, the effect of a one percent increase in US unemployment at graduation is associated with a 0.1% underestimation of actual earnings by the simulated distributions. The effect of a one percent increase in GDP growth rate at graduation is negligible. Those two regressions can be seen in the nearby table.

Table 1: Graduation Economic Indicators Effect on Simulated Distribution Error

	Dependent variable:	
	perc.med	
	(1)	(2)
USGU	-0.001^{**}	
	(0.0004)	
USGG		-0.0001
		(0.0004)
P.FT	0.063^{***}	0.065^{***}
	(0.006)	(0.006)
Constant	-0.049^{***}	-0.055^{***}
	(0.006)	(0.005)
Observations	13,526	13,526
R^2	0.010	0.010
F Statistic (df = 2; 13523)	67.792^{***}	65.496^{***}

Note: *p<0.1; **p<0.05; ***p<0.01
The following OLS regression was performed:
$perc.med = \beta_0 + \beta_1 * USGU(orUSGG) + \beta_2 * P.FT + \varepsilon$
Observations are weighted by full time workers reporting
perc.med is the % diff. of the median and simulated median
USGU and USGG are US unemployment and growth at graduation
P.FT is the percent of full time workers reporting

It seems that although there is some bias from simulating log-normal wage distributions using the 25th and 75th percentile wages, that it is not systematically related to economic indicators, and if anything, it would just reduce the expected finding of long term wages being lower when graduating during poor economic times. That is, it would have an attenuating bias on the magnitude of the effect of graduating in bad economic times.

To test this, and hopefully to put to rest some of the concern of working with the simulated dataset instead of individual observations, I performed an additional simple comparison regression to see if the prediction error of the distributions (which is systematically related to higher median incomes) would have a significant effect on the results when adjusted for.

The adjustment I made was rather adhoc, since figuring out the math for the actual adjustment for the distributional error is well beyond my or most graduate students' knowledge. Nonetheless, I think the adjustment was appropriate in that it helped shift the distributions in the appropriate direction and here is how it was performed:

- For each simulated wage, find it's associated percentile from its simulated distribution
- Adjust the median income by a percentage of the median error based on its distance from this median error
- This distance percentage was calculated as 1 minus four times the absolute value of 0.5 minus the percentile found from earlier.
- Hence, a full adjustment would be performed for simulated wages at the median, and a lower adjustment for values in either direction, although not before or after the 25th percentile or 75th percentile wage, since data on the 1st and last percentile earnings are not available.

This shows that the bias from this part of the simulation error is small and points in the same direction as the measured effect. Hopefully this analysis can put to rest concerns with the simulated shape of the wage distributions, as using the simulated distributions seems to underestimate the effect of poor economic indicators at the time of graduation on the true distributions.

Summary statistics and economic indicator names

In the accompanying table is the basic data for this simulated dataset, along with the economic indicators used in the regressions.

Table 2: Comparison regressions with and without adjustments for distributional error

	Dependent variable:	
	log(income) ∗100	log(adj.income) ∗100
	(1)	(2)
USGU	−0.602*** (0.018)	−0.604*** (0.018)
Wage1999	4.724*** (0.188)	4.733*** (0.188)
Wage2000	6.783*** (0.184)	6.796*** (0.184)
Wage2001	8.170*** (0.180)	8.186*** (0.180)
Wage2002	6.142*** (0.178)	6.154*** (0.178)
Wage2003	5.422*** (0.176)	5.432*** (0.176)
Wage2004	8.289*** (0.174)	8.304*** (0.174)
Wage2005	8.324*** (0.172)	8.338*** (0.172)
Wage2006	8.603*** (0.170)	8.618*** (0.170)
Wage2007	8.383*** (0.169)	8.398*** (0.169)
Wage2008	7.274*** (0.168)	7.286*** (0.168)
Wage2009	7.459*** (0.167)	7.472*** (0.167)
Wage2010	5.352*** (0.166)	5.361*** (0.166)
Wage2011	2.358*** (0.165)	2.361*** (0.165)
Wage2012	1.225*** (0.165)	1.227*** (0.165)
Wage2013	4.078*** (0.166)	4.085*** (0.166)
Exp	2.511*** (0.076)	2.516*** (0.076)
Degree1-2YRCert	−7.434*** (0.445)	−7.457*** (0.445)
DegreeAssTech	8.156*** (0.336)	8.172*** (0.336)
DegreeAssBach	−4.367*** (0.396)	−4.378*** (0.396)
Degree4yrBach	5.237*** (0.294)	5.251*** (0.294)
Degree5yrBach	15.887*** (1.345)	15.932*** (1.345)
DegreePostBach	41.070*** (1.030)	41.160*** (1.031)
DegreeMast	45.785*** (0.313)	45.878*** (0.314)
DegreePostMast	46.535*** (1.021)	46.635*** (1.021)
DegreeDoct	64.335*** (0.682)	64.458*** (0.683)
DegreeProf	54.758*** (0.479)	54.871*** (0.480)
I(Exp^2)	−0.067*** (0.004)	−0.067*** (0.004)
Exp:Degree1-2YRCert	0.181 (0.114)	0.182 (0.114)
Exp:DegreeAssTech	−0.109 (0.089)	−0.109 (0.089)
Exp:DegreeAssBach	2.118*** (0.099)	2.123*** (0.099)
Exp:Degree4yrBach	3.796*** (0.079)	3.803*** (0.079)
Exp:Degree5yrBach	6.641*** (0.326)	6.649*** (0.326)
Exp:DegreePostBach	2.153*** (0.379)	2.153*** (0.380)
Exp:DegreeMast	0.651*** (0.085)	0.650*** (0.085)
Exp:DegreePostMast	1.798*** (0.282)	1.798*** (0.282)
Exp:DegreeDoct	2.392*** (0.211)	2.392*** (0.211)
Exp:DegreeProf	6.943*** (0.133)	6.947*** (0.133)
Constant	1,043.585*** (0.340)	1,043.475*** (0.340)
Observations	5,391,219	5,391,219
R^2	0.184	0.184
Adjusted R^2	0.184	0.184
F Statistic (df = 48)	25 million	25 million

Note: *p<0.1; **p<0.05; ***p<0.01

Full regressions with (1) and without (2) an adjustment for simulation error

Table 3: Simulated dataset summary stats

Statistic	Mean	St. Dev.	Min	Max
ln(SW)	10.78	0.54	7.85	15.23
Exp	7.46	4.63	1	20
VAGU	3.96	1.05	2.31	7.05
USGU	5.46	1.19	4.00	9.60
VAGG	2.13	1.48	−1.22	4.88
USGG	3.07	1.49	−2.80	4.70
VAGUL	4.11	1.13	2.31	7.05
USGUL	5.56	1.17	4.00	9.60
VAGUF	3.90	1.08	2.31	7.05
USGUF	5.41	1.25	4.00	9.60

Data is for a total of 4,608,775 observations.
ln(SW) means the natural log of the simulated wage.

The economic indicators are as follows:

VAGU the Virginia unemployment rate at graduation

VAGG the Virginia GSP growth rate at graduation

USGU the US unemployment rate at graduation

USGG the US GDP growth rate

VAGUL one year lag of VAGU

VAGUF one year lead of VAGU

USGUL one year lag of USGU

USGUF - one year lead of USGU

Please see the final appendix for a much better visual description of the data. There, you will find correlation graphs and tables for the economic indicators, as well as many visual representations of the dataset.

4 Test Error Regression Fits

The table at the end of this section was a means of seeing which models had the most predictive power for wages, and for learning more about the patterns within the different parts of wage distributions. It uses a simple validation strategy repeated a very high number of times to estimate the mean standard error of a particular model. The method used was as follows:

- First randomly select a training and test set, consisting of 80% and 20% of observations respectively.
- Use the training set to run a linear regression.
- Use the coefficients from the regression to predict the held out test set observations
- Calculate the mean standard error of the test set
- Repeat this process 10,000 times, and find the average mean standard error
- Do the entire process with the Median of the cohort wage, 25th percentile of the cohort wage, and 75th percentile of the cohort wage as response variables.

- Repeat the entire process 10 more times to find an estimate of the standard error of this process and less biased estimates of the means

The resulting average mean standard errors are an excellent basis of comparison for the predictive value of various model fits; in fact they are much better for comparison than measures like adjusted RSS or AIC because they are not merely a training fit of the data, but are a measure of the predictive power of the various predictors on the withheld test sets, at least over the time period and data studied. I used mean standard errors instead of mean squared areas because the former has an easily interpreted value as the standard deviation of the predictive accuracy of the model.[18]

Each linear model used years of work experience, a set of dummy variables for degree held, and a set of dummy variables for the work year in question. It also includes the extra regressors listed in the extra regressors' column (so the first row represents only the base predictors described above), and any extra regressors in other models referenced by number. Any extra variable also includes interaction dummies for each degree. [19] Additional terms are as follows:

IExp2 a term for squared experience

MMSE the mean standard error of the prediction of the median wage on the test set

TMSE the mean standard error of the prediction of the 25th percentile wage on the test set

SMSE the mean standard error of the prediction of the 25th percentile wage on the test set

Errors The standard errors of the process itself are in parenthesis

Table 4: Repeated Validation Predictors Table

	TMSE	MMSE	SMSE	Extra.Reg
1	2176 (1.2)	3066 (2)	4781 (4.1)	
2	1780 (0.9)	2481 (1)	4731 (5.5)	I(Exp^2)
3	2162 (2.1)	3014 (7.2)	4713 (9.1)	VAGU
4	2167 (1.7)	3044 (2.4)	4722 (4.2)	VAGG
5	2159 (2.6)	3016 (5.6)	4717 (3.8)	USGU
6	2157 (1.5)	3035 (1.7)	4735 (3)	USGG
7	2156 (1.3)	2976 (9.2)	4684 (14)	3+4
8	2140 (2.2)	2942 (9.4)	4707 (6.9)	5+6
9	1772 (4.7)	2470 (6.7)	4672 (14.5)	2+3
10	1752 (1)	2418 (1.2)	4677 (5.6)	2+4
11	1752 (1.3)	2397 (4.2)	4640 (15.8)	2,3+4
12	1768 (7.3)	2466 (10.6)	4677 (4.4)	2+5
13	1736 (1)	2395 (1.7)	4685 (4)	2+6
14	1738 (1.7)	2367 (4.5)	4670 (5.5)	2,5+6
15	1730 (1.2)	2350 (4.6)	4660 (11.5)	2,3+6
16	1752 (3.2)	2406 (2)	4646 (7.6)	2,4+5
17	1736 (1.4)	2343 (7.7)	4622 (16.8)	2,3,4 + VAGUL
18	1738 (1.8)	2369 (4.7)	4569 (9.8)	2,3,4 + VAGUF
19	1721 (2.4)	2313 (7.8)	4541 (9.1)	17 + VAGUF
20	1745 (1.6)	2365 (2.1)	4617 (18.3)	3,4 + USGUL
21	1741 (2.2)	2370 (2.4)	4675 (7.4)	3,4 + USGUF
22	1747 (2.6)	2364 (8.6)	4626 (23.6)	20 + USGUF
23	1773 (0.8)	2427 (21.4)	4680 (3.6)	2 + VAGUL
24	1778 (4.2)	2455 (8.5)	4673 (2.8)	2 + USGUL
25	1758 (3)	2448 (10.2)	4637 (5)	2 + VAGUF
26	1746 (1.9)	2434 (7.4)	4668 (2.9)	2 + USGUF
27	1738 (5)	2350 (20.7)	4618 (7.7)	23 + VAGUF
28	1738 (1.6)	2379 (9.3)	4642 (7.5)	24 + USGUF

Now let's discuss what these tests seemingly reveal. From (2) we see that adding a squared term for experience greatly adds to the predictive power of a model for the 25th percentile and median wage earners. It holds very little predictive power for 75th percentile workers. This seems to suggest that 75th percentile workers or higher have returns to experience which are much closer to a constant linear return than the other percentiles, which is an intriguing idea.

Next, look at (3) vs. (5), (12) vs. (9), and (8) vs. (7). It seems that Virginia and US indicators are both mostly equal in their ability to fit wages.

Now, look at (8) vs. (7). The US predictors together are more accurate than the VA ones together, and their explanatory power is greater with more predictors. This is a pattern that persists for all regressions. With more predictors, the models become better at predicting. Remember that these are predictions on a held-out test set repeated 10,000 times randomly. There is not over-fitting since

[18] I was inspired to use this process after taking the course Statistical Learning from Stanford University Online.

[19] The referencing of other formulas was done to save space, and make the table more manageable.

a useless term would not increase the predictive accuracy of the model on the test set, unlike a measure such as RSS for a full fit of the data.[20]

Look at (15) vs. (16). Virginia GSP growth seems to add less explanatory power than US GDP growth. This may be due to the discontinuous nature of the data that I adjusted for[21] or it may be due to lower accuracy of the Virginia GSP measure. This is true across many of the regressions. For example, compare (6) and (4) or (8) and (7).

Another strange pattern is how the lags for VA and US unemployment, as well as the leads, are better predictors than the numbers during year of graduation. This is somewhat troubling, but is ameliorated by the fact that including more predictors increases the predictive accuracy of the model in general. When included with the year of graduation predictors, the overall performance of the model improves.

While including additional predictors does a better job at predicting actual salaries, it makes the coefficients on them very hard or impossible to predict due to the high degree of correlation between all of economic indicator predictors. Hence the focus will be on performing regressions with a single economic predictor or single combination of unemployment predictors and interpreting the results.

Finally, it's important to admit that adding in economic indicators at graduation does little to increase the predictive power of the model. What is clear is that each indicator does in fact add to the predictive power of the model and so is significant, but that any indicator's importance pales in comparison to experience or the type of degree received. This is common sense[22] and will be made explicit through the regression results in the next section.

5 Regressions

All of the regressions for this section and the next can be found in the appendix[23] The most important of them are tables 1 - 3, which show the effects of different economic indicators with different specifications using pooled OLS and the entire simulated dataset. Table 1 uses this formula —

$$ln(SW_i) = \beta_0+\beta_1 EI_g+\mathcal{B}_1 Yr_t+\mathcal{B}_2 Deg_d+\mathcal{B}_3(Deg_d * Yr_t)$$

— where $ln(SW_i)$ refers to the log simulated wage of individual i, EI_g refers to the economic indicator at graduation year g (or lag or lead) being used [24], Deg_d refers to a set of dummies for degree levels, only one of which is active for a given individual, and where Yr_t refers to a set of dummies for the year of work.

Table 2 adds terms for experience and experience squared along with their degree interactions, and Table 3 further interacts the economic indicators at graduation (or

[20] A measure such as RSS will always increase when adding more predictors, but the average of many, many test sets should not.

[21] There were two series of gross state product available on the Virginia State website, one of which ended in 1997, and one of which started in 1997. I adjusted the former based on the percentage difference of the overlapping year.

[22] This should go without saying, but experience and the type of education one has should be a much more important determinant of wages than temporary economic indicators at graduation, ceteris paribus. Over the long run, economic growth becomes one of the most important predictors of wages, but its predictive power in shorter time periods is not as great as these other more important determinants.

[23] To be clear, the regressions in the appendix are for sections 5 and 5.1. They were the initial set of regressions I performed for this paper, and the degree level ones took a long time to perform. Perhaps here is the best place to confess that the amount of time necessary to reproduce these regressions with the expanded dataset was prohibitive, so I used them instead. Nonetheless, the only differences in the dataset are a few less common degree types which were not in the original dataset and about 2% additional workers reporting, as described in the data section.

[24] See the data simulation subsection for a description of these indicators, and see the data section for a description of degree types.

leading or lagging graduation) with the experience and experience squared terms.

As can be seen from the tables, all of the coefficients are statistically significant (usually extremely so), and almost all of the economic indicators point in the expected direction, with a positive effect on wages with higher GDP or GSP growth, and a negative effect for higher Virginia or US unemployment.[25]

The inclusion of my proxy for experience greatly reduces most of the economic indicator coefficients, but they still remain. All terms, except perhaps the interactions with squared experience terms, are significant and point in the right direction for the last two tables.

The results pretty clearly show that there is a strong relationship between poor economic indicators at graduation and lower wages averaged over the years of this study (remember from the summary data table that the average "experience" for a simulated worker in this dataset is about 7 years.)

One interesting aspect of these regressions is the relationship between including experience as a predictor and the estimated effect of the unemployment predictors versus growth predictors: without experience as a predictor, growth seems much more of an important predictor of the difference in wages; with experience, it seems like unemployment is a better predictor, for both state and national level data.

Individual Degree Ressions

The regressions for the fully simulated data set should be rather reliable because of the large number of observations. However, due to the size of the dataset (more than four million observations), repeated simulations of the dataset were computationally prohibitive. This was not so much the case for degree level regressions, since the degree with the largest amount of simulated observations (a bachelor's degree) was about half of the size of the entire simulated data set.

This was the process that was followed for degree level regressions, which should make their estimates more precise than the regressions using the entire dataset in the previous section or the sections to come.

- Simulate a dataset for each degree-level.
- Run a given regression using those data-sets, extracting the coefficient estimates and standard error estimates.
- Repeat the process 100 times
- Use the average coefficient estimate from the 100 regressions on a given degree-level regression as the point estimate.
- Use the average standard error from the regressions, plus the standard error of the 100 standard errors, as the standard error for a given degree-level regression.

That was the process used to produce the degree-level regressions in the appendix. Because of the large number of simulations per regression (100), the confidence in the point estimates should be a lot higher than those for the full dataset regressions, although how to implement standard errors was not very clear to me, so I used the very conservative method as described above.[26]

Let's see what we can learn from the degree-level regressions.

Just as with the full data set regressions, there is an opposite effect for the lagged

[25]Except the the lag for Virginia unemployment and US unemployment at graduation in the base table. This finding is quite puzzling, and I am not quite sure what to make of it or how it really relates to experience.

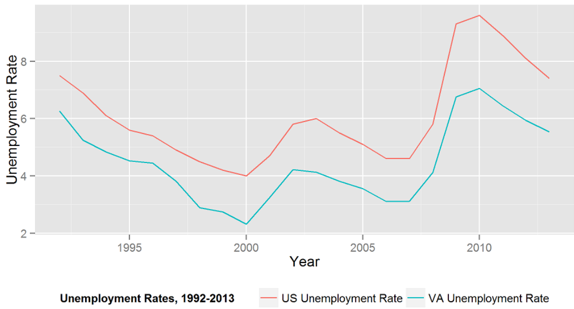
[26]It seems, if anything, that the standard errors are overestimating the real standard error of the point estimates, but it is better to error on the side of caution rather than overstate my results.

unemployment indicators and the lead indicators. Looking at regression tables 4, 10, and 11, we can see that there is a significantly higher estimated negative effect for the unemployment rate one year after graduation than the year of graduation. This makes some sense; in bad economic times, it would be hard to find a job after graduation. If the economy worsened in the coming year, the long term effects of not being able to find a job or being underemployed could be quite severe.

The estimates on the lagging indicator are much harder to figure out, and they stand out just as sharply, if not more so, than in the full-degree regressions. The only explanation that seems to make sense is that unemployment during the year of graduation is too much of a lagging indicator, and the unemployment a year forward is a much better measure of the economic environment that one graduates into. There isn't much to explain an increase of wages for those who graduate a year after high unemployment, unless the unemployment rate reverts to the mean extremely rapidly, which it does not seem to. There are a few fairly sharp rebounds in the data (see the accompanying figure), and there were two serious bubbles that burst during this time frame (with seemingly minimal long-term repercussions), so perhaps that would go a long way toward explaining the positive effect on wages when graduating a year after a high unemployment rate, especially if that unemployment rate was a rather lagging indicator (and so the measure was more of a lag of two years, rather than one).

There are a few interesting effects going on that can be seen most clearly in appendix tables 4 to 7. There, you can see that the average effect across all wage years of graduating in poor economic times moves in the expected direction for all but two degree types – technical associate's degrees and professional degrees. This is an interesting phenomenon that deserves more study.

Figure 3: Unemployment over time



The 2000 Effect

Just considering the effect on wages from graduating into a bad economy across the whole dataset should be sufficient to establish the link between poor economic times at graduation and lower wages. However, there are a few especially interesting cohorts that can be analyzed to get a better idea of this effect and to hopefully serve as robustness checks.

The graduates of the year 2000 had the lowest first year full time employment, by far, of any cohort in this study. However, by the second year their full time employment percentages had jumped back to normal levels.[27] This seems like a good opportunity to see the effect of being underemployed during the year after graduation on one's future years of employment.

Unfortunately, the 2001 wage data for 2000 graduates exhibits extremely unusual behavior. If you choose to discount the strangeness of this data and look at future years data, there seems to be a significant effect on earnings of 2000 graduates on the several years after 2001, even after controlling for full vs. part time work and job type.

[27] For all degree types, the percent of full time workers one year after graduation is 68% for 2000 graduates. It averages 77% for all other graduates. Interestingly, this jumps up to 83% for 2000 graduates vs 82% for all others two years after graduation.

Table 5: The effect of graduating in the year 2000 for Virginia graduates

	Dependent variable:			
	log(income)			
	(1)	(2)	(3)	(4)
FirstYear	−0.065*** (0.001)	0.084*** (0.001)	−0.065*** (0.001)	0.018*** (0.001)
Exp	0.028*** (0.0001)	0.020*** (0.0001)	0.028*** (0.0001)	0.023*** (0.0001)
P.FT		2.227*** (0.003)		1.262*** (0.005)
TwoK	0.023*** (0.003)	0.009*** (0.003)	0.018*** (0.003)	0.012*** (0.003)
FirstYearTRUE:TwoK	−0.009** (0.004)	0.241*** (0.004)	−0.005 (0.004)	0.134*** (0.004)
TwoKTRUE:factor(Wage)2002	−0.063*** (0.004)	−0.015*** (0.004)	−0.063*** (0.004)	−0.036*** (0.004)
TwoKTRUE:factor(Wage)2003	−0.043*** (0.004)	−0.010** (0.004)	−0.045*** (0.004)	−0.026*** (0.004)
TwoKTRUE:factor(Wage)2004	−0.021*** (0.004)	−0.009** (0.004)	−0.023*** (0.004)	−0.015*** (0.004)
TwoKTRUE:factor(Wage)2005	−0.001 (0.004)	0.001 (0.004)	−0.004 (0.004)	−0.002 (0.004)
TwoKTRUE:factor(Wage)2006	−0.006 (0.004)	−0.011*** (0.004)	−0.002 (0.004)	−0.006* (0.004)
TwoKTRUE:factor(Wage)2007	0.005 (0.004)	−0.007 (0.004)	0.009** (0.004)	0.001 (0.004)
TwoKTRUE:factor(Wage)2008	0.016*** (0.004)	0.010** (0.004)	0.019*** (0.004)	0.014*** (0.004)
TwoKTRUE:factor(Wage)2009	0.013*** (0.004)	−0.007 (0.004)	0.015*** (0.004)	0.003 (0.004)
TwoKTRUE:factor(Wage)2010	0.008** (0.004)	−0.014*** (0.004)	0.009** (0.004)	−0.004 (0.004)
TwoKTRUE:factor(Wage)2011	0.004 (0.004)	−0.011*** (0.004)	0.005 (0.004)	−0.004 (0.004)
TwoKTRUE:factor(Wage)2012	0.014*** (0.004)	−0.008* (0.004)	0.014*** (0.004)	0.002 (0.004)
TwoKTRUE:factor(Wage)2013				
Constant	10.331*** (0.002)	8.719*** (0.003)	10.161*** (0.004)	9.308*** (0.005)
Observations	5,391,219	5,391,219	5,391,219	5,391,219
R^2	0.176	0.233	0.269	0.278
Adjusted R^2	0.176	0.233	0.269	0.278
Residual Std. Error	0.481 (df = 5391178)	0.464 (df = 5391177)	0.453 (df = 5391144)	0.450 (df = 5391143)
F Statistic	28 mil.***	40 mil.***	26 mil.***	27 mil.***
Note:				*p<0.1; **p<0.05; ***p<0.01

The following is the base regression, which compromises all of column (1): $log(income) = \beta_0 + \alpha_1 * FirstYear * TwoK + B_2 * TwoK * Wage_t + B_3 * Degree_d + \beta_4 * Exp$. TwoK is a dummy which means the simulated wage comes from a 2000 graduate cohort. FirstYear is a dummy for the first year of work. Columns (2) and (4) include a control for the percent of full time workers of a given graduation year - wage year cohort. Columns (3) and (4) include a set of dummy variables that control for job industry. Not printed are values for wage year dummy variables, degree type dummy variables, and job types. The base year for the wage year interactions is 2013.

Looking at the nearby regression table, "The effect of graduating in the year 2000 for Virginia graduates", we can see the unusual behavior just mentioned.

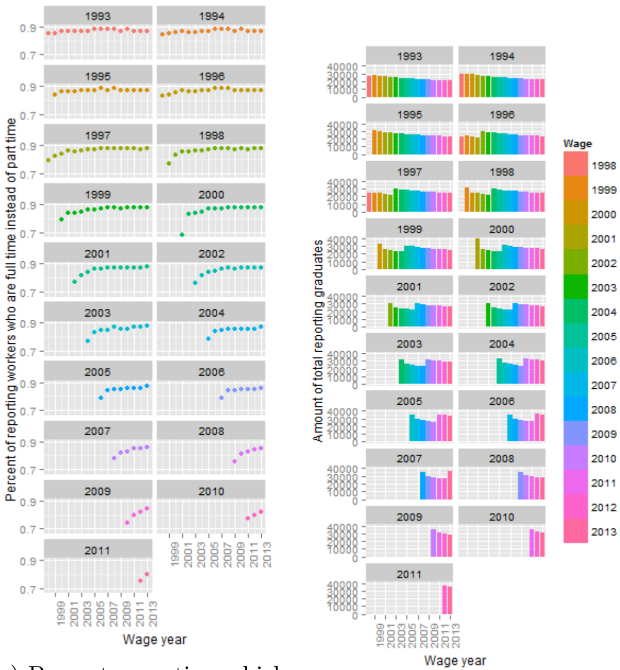
The regression controls for year fixed effects, experience, degree type, and dummy variables representing year 2000 graduates and first year jobs. Strangely, wages are slightly higher for this group on average during their first year after graduation[28], the same year when the percentage of full time workers is so low for this cohort.

This is not the only unusual thing going on with the data. It is not a priori true that wages will be lower for recent graduates who manage to secure jobs during bad economic times; in normal times, the graduates who would not secure work in bad times could be quite marginal workers who weigh down the average earnings of all full time employed recent graduates. So it is possible that during bad economic times, average full time wages for recent graduates who are employed could be higher than average full time wages for recent graduates who are employed in better times.[29] There is also the likely case of wages of recent graduates being higher conditional on full time employment during the first year after graduation, since those who find full time work quickly are likely to be better workers than those who do not. However, even if ei-

[28] Estimated at about 1.4%, as can be seen from regression (1) by adding together the dummy for 2000 graduates with its first year interaction dummy.

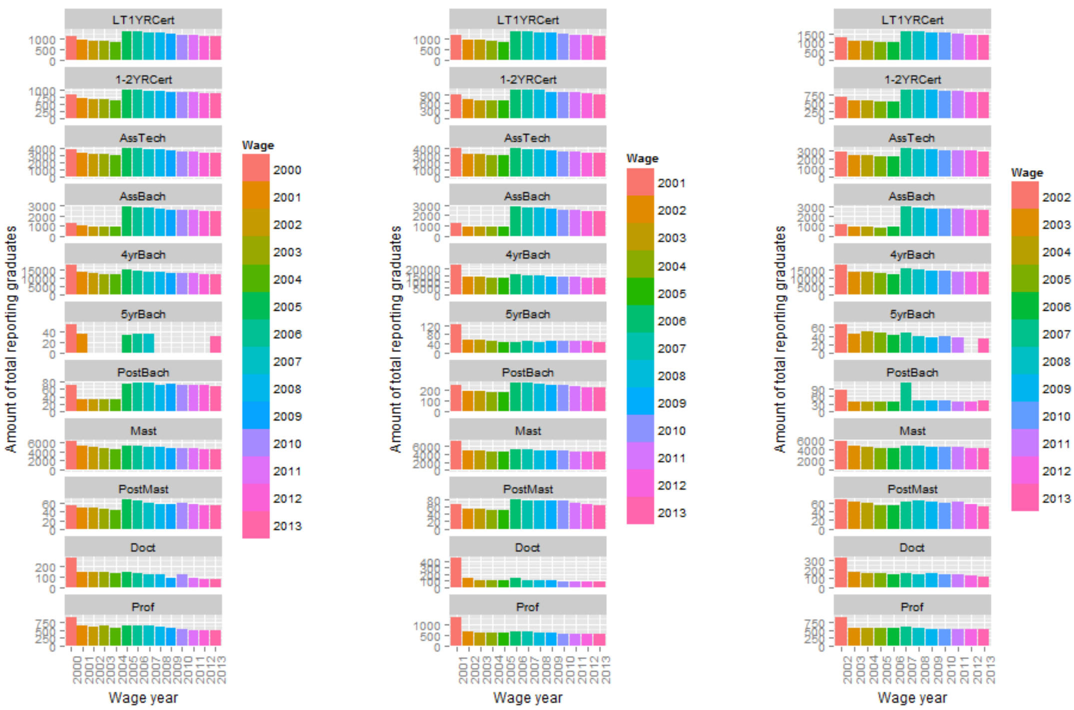
[29] This author does not believe that this is the most likely scenario. Bad times should cause companies to cut entry level wages for all but the most promising of candidates.

Figure 4: Percentage full time and total reporting workers by year and graduation cohort



(a) Percent reporting which are full time

(b) Total reporting



(c) 1999 graduate cohort

(d) 2000 graduate cohort

(e) 2001 graduate cohort

Figure 5

ther or both of these related phenomena are real ones, they are unlikely to lead to sharply higher wages conditional on being full-time employed. Disconcertingly, this is exactly what the data says for year 2000 graduates. Controlling for the percent of workers who have full time employment, first year wages for 2000 graduates are a full 25% higher than first year wages for other cohorts.

This strange finding led me to examine the data for 2000 graduates much more carefully. Looking at the total number of reporting graduates in graphs b-e of the figure titled "Percentage full time and total reporting workers by year and graduation cohort", one sees a pattern. Observations from one year out of school are more numerous than from later work years for almost all graduation cohorts, but observations for the first year of potential wages for the 2000 cohort is far higher than subsequent years.

Looking into the details of the first year of potential wages for 2000 graduates, it seems like the number of reporting graduates is sharply higher across most degree types, most importantly for those with four year bachelor's degrees (although the pattern is more striking for doctorate degrees and professional degrees). Compare the 1999, 2000, and 2001 cohorts by degree types to better understand the issue. There is an enormous number of graduates reporting for the 2000 cohort in the first year after graduation, during a period of difficult economic times. This seems to suggest that there is something wrong with the data for this year, especially because this excess of observations is isolated to this specific wage year for this specific graduation cohort.[30]

Ignoring, for a moment, the elephant in my data in the form of higher first year wages for 2000 graduates when a large percentage could not obtain full-time work, the pattern of wages over the following few years suggest that not being able to secure a full-time job has a significant effect on future earnings, at least for several years. Looking at regression (1), we see that the earnings effect of graduating in the year 2000 is 4% lower earnings and 2% lower earnings two and three years after graduation, respectively. This is partly because of lower full time employment for these workers, but if you control for job type and full time employment percentage, more than half of the lower earnings effect remains for the 2000 cohort. Most of the effect seems to have disappeared after five years.

Clearly, the results for the 2000 graduates are quite mixed. On the one hand, there is a quite low rate of full time employment for these graduates during their first year after school, and their wages are lower for several after graduation, even when controlling for the number who are full time employed, or even by job type! On the other hand, the data is quite suspicious for this cohort's first year's wage earnings. Wages are far higher than other cohorts when controlling for full time employment, and the total number of employed for this one wage year of this cohort represents a rather stark outlier from the rest of the dataset.[31]

The 2008 and beyond effect

The first year full time employment dips were not as pronounced for 2008-2011 graduates as they were for the year 2000 graduates, but first year employment was still significantly below trend for all four graduation years.

[30] Carefully examining the total reporting graph, you can clearly see that no other wage year for any graduation cohort sticks out like the 2001 wage year for the 2000 graduates.

[31] Due to the size of the dataset and limited time, I did not have time to perform repeated simulations for the 2000 effect or for the next two subsections. Nonetheless, the sheer size of the simulated dataset should put to rest most concerns regarding the bias of regression estimates for this regression and the ones to follow.

Table 6: The effect of graduating in the years 2008-2011 for Virginia graduates

	Dependent variable: log(income)			
	(1)	(2)	(3)	(4)
FirstYear	0.101*** (0.001)	0.100*** (0.001)	0.101*** (0.001)	0.102*** (0.001)
Exp	0.020*** (0.0001)	0.020*** (0.0001)	0.020*** (0.0001)	0.020*** (0.0001)
P.FT	2.185*** (0.003)	2.186*** (0.003)	2.185*** (0.003)	2.185*** (0.003)
Two8	−0.0001 (0.003)			
FirstYearTRUE:Two8	−0.006 (0.004)			
Two8TRUE:factor(Wage)2010	0.027*** (0.004)			
Two8TRUE:factor(Wage)2011	0.0001 (0.004)			
Two8TRUE:factor(Wage)2012	0.016*** (0.004)			
FirstYearTRUE:Two9		0.023*** (0.004)		
Two9TRUE:factor(Wage)2011		0.031*** (0.004)		
Two9TRUE:factor(Wage)2012		0.022*** (0.004)		
Two9		−0.006** (0.003)		
Two10			0.004 (0.003)	
FirstYearTRUE:Two10			−0.003 (0.004)	
Two10TRUE:factor(Wage)2012			0.025*** (0.004)	
Two11				0.026*** (0.003)
FirstYearTRUE:Two11				−0.048*** (0.004)
Constant	8.744*** (0.003)	8.744*** (0.003)	8.744*** (0.003)	8.744*** (0.003)
Observations	5,391,219	5,391,219	5,391,219	5,391,219
R^2	0.232	0.232	0.232	0.232
F Statistic	49 mil.***	50 mil.***	52 mil.***	54 mil.***

	(1)	(2)	(3)	(4)
FirstYear	−0.065*** (0.001)	−0.064*** (0.001)	−0.068*** (0.001)	−0.064*** (0.001)
Exp	0.028*** (0.0001)	0.028*** (0.0001)	0.028*** (0.0001)	0.028*** (0.0001)
Two8	0.013*** (0.003)			
FirstYearTRUE:Two8	−0.024*** (0.004)			
Two8TRUE:factor(Wage)2010	−0.055*** (0.004)			
Two8TRUE:factor(Wage)2011	−0.042*** (0.004)			
Two8TRUE:factor(Wage)2012	−0.005 (0.004)			
FirstYearTRUE:Two9		−0.028*** (0.004)		
Two9TRUE:factor(Wage)2011		−0.048*** (0.004)		
Two9TRUE:factor(Wage)2012		−0.008** (0.004)		
Two9		−0.013*** (0.003)		
Two10			−0.032*** (0.003)	
FirstYearTRUE:Two10			0.053*** (0.004)	
Two10TRUE:factor(Wage)2012			−0.022*** (0.004)	
Two11				−0.057*** (0.003)
FirstYearTRUE:Two11				0.031*** (0.004)
Constant	10.331*** (0.002)	10.331*** (0.002)	10.332*** (0.002)	10.331*** (0.002)
Observations	5,391,219	5,391,219	5,391,219	5,391,219
R^2	0.175	0.175	0.175	0.175
F Statistic	35 mil.***	37 mil.***	38 mil.***	mil.***

Note: *p<0.1; **p<0.05; ***p<0.01

The two sets of regressions in the table above are identical, except the first set includes a control for the percent of full time workers of a given graduation year - wage year cohort. The following regression is for the top set of tables: $log(income) = \beta_0 + \alpha_1 * FirstYear * TwoX + B_2 * TwoX * Wage_t + B_3 * Degree_d + \beta_4 * Exp$. TwoX represents Two8-Two11, which are dummies for having graduated in 2008-2011, only one of which was used for each column's regression. Running the regressions with all dummies produces extremely similar results, but depending on the order in which they are entered in the regression, show no value because of singularities. Not printed are values for wage year dummy variables and degree type dummy variables. The base year for the wage year interactions is 2013.

The question is what is the observable effect of having graduated during or after the 2008 financial crises. Two sets of four regressions, almost identical to those used for the 2000 graduates, were used to try and see these effects. The first set does not control for the percentage of full time workers. The second does.

Looking through the bottom regressions of the table "The effect of graduating in the years 2008-2011 for Virginia graduates", it is apparent that the first three years of wages were lower for 2008 and 2009 graduates, by a similar amount as the 2000 graduates[32]. However, by comparing it to the top set of regressions, you can see that all of the effect disappears when controlling for full time employment, so again it is the dip in full time employment that is the primary driver for lower wages during these years for these graduates.

Separate regressions were included for each year because of singularity issues, but including all of the dummies in one giant regression produces remarkably similar total estimates for the first year differential wage effects for the different graduation years.

Together with the 2000 graduates, there seems to be sufficient evidence that there is a very significant effect on the first few years of wages after graduating into a bad economy, but that almost all of the effect is driven by the lack of full time employment. This seems to invalidate the earlier theoretical discussion that suggested the lack of full time employment in one's early years could lead to persistent and significant lower earnings because of the lack of human capital accumulation during some of the most important years of one's working life. However, the results in a later section provide support for this very idea.

The first five years

Since it seems that the effects of graduating in bad economic times seems to be concentrated mostly around the first few years after graduation, I decided to drop all observations that represented wages six or more years after graduation, leaving the first five years of potential work experience for each cohort, with the exception of those who graduated too late to have had five years of wage data.[33]

The results can be seen in the table titled "Effect of economic indicators around graduation on five years of wage data". The indicators are all pointing in the right direction, although it seems a quadratic experience term interacted with the appropriate economic indicator has problems fitting the edge years. The initial effect on graduation for higher gross state product growth is positive, and declines over time at a decreasing rate. The initial effect of high unemployment at graduation is negative, and this effect lessens over time at a decreasing rate. These are both expected effects.

The average effect is on the order of 1% higher first year wages for every additional percent of Virginia gross state product at graduation, and 0.7% lower wages for each additional point of Virginia unemployment at graduation. Both effects disappear over time as described in the previous paragraph.

Finally, despite its greater usefullness as a predictor of wages in the test error regression fits, the lag of Virginia unemployment at graduation does not have nearly the level of influence on actual wages as the actual Virginia unemployment at graduation, and so my speculation as to the reasons for its greater predictive power

[32] Remember to add not only the interaction value between the wage year and graduation year dummy, but also the coefficient on the graduation year dummy to get the full estimated effect.

[33] I did not drop these cohorts. Instead, I used whatever data was available, as described in the data section.

Table 7: Effect of economic indicators around graduation on five years of wage data

	Dependent variable:		
	log(income) times 100		
	(1)	(2)	(3)
Exp	4.778*** (0.531)	2.228*** (0.652)	3.302*** (0.677)
I(Exp^2)	−0.333*** (0.088)	0.032 (0.111)	−0.073 (0.113)
VAGG	1.556*** (0.086)		
Exp:VAGG	−0.676*** (0.068)		
VAGG:I(Exp^2)	0.064*** (0.011)		
VAGU		−1.009*** (0.112)	
Exp:VAGU		0.388*** (0.097)	
VAGU:I(Exp^2)		−0.073*** (0.017)	
VAGUL			−0.351*** (0.131)
Exp:VAGUL			0.139 (0.111)
VAGUL:I(Exp^2)			−0.048** (0.019)
Constant	1,035.313*** (0.689)	1,042.890*** (0.821)	1,040.321*** (0.854)
Observations	2,147,321	2,147,321	2,147,321
R^2	0.163	0.162	0.162
Adjusted R^2	0.163	0.162	0.162
Residual Std. Error (df = 2147270)	0.442	0.442	0.442
F Statistic (df = 50; 2147270)	8,334.109***	8,328.906***	8,323.889***

Note: *p<0.1; **p<0.05; ***p<0.01

The following is the base regression for all of the above regressions: $log(income) * 100 = \beta_0 + \alpha_1 * Year_t + \beta_2 * Exp + \beta_3 * Exp^2 + B_4 * Degree_d + B_5 * Degree_d * Exp + B_6 * Degree_d * Exp^2$. Subscript t represents the year dummy of the wage in question and d represents each of the dummies for degrees. Columns one, two, and three include Virginia GSP at graduation, Virginia unemployment at graduation, and a one year lag of Virginia unemployment at graduation respectively, as well as each's interactions with experience and experience squared. The actual coefficients on any year dummies or degree dummies and their interactions are not shown for ease of interpretability.

in the test error regression fits seem to be wrong, but this is good; conditions at graduation and the year after graduation should have the greatest influence on wages.

After five years

The final robustness check that I performed was to see the effect of graduating in poor economic times on wages more than five years out from graduation. To do so, I dropped all simulated observations that represented workers five years or less out from graduation. Then, for ease of interpretability, I adjusted the Exp control variable by subtracting off six for each remaining observation.[34]

The desire, of course, was to find out medium and long term effects on wages from graduating in bad economic times. The results can be seen in the table titled "The effect of various economic indicators at graduation more than five years after graduation". To be able to interpret the table, you need to know the distribution of my so called Exp variable which I described above. A histogram is shown that shows this distribution, and the mean value of this "Exp" for all remaining simulated observations is 4.309.

With this in mind, it is easy to calculate the estimated effect of graduating in poor economic times on wages for the mean ob-

[34]So its value becomes 0 for observations 6 years after graduation, 1 for 7, and so on.

Table 8: The effect of various economic indicators at graduation more than five years after graduation

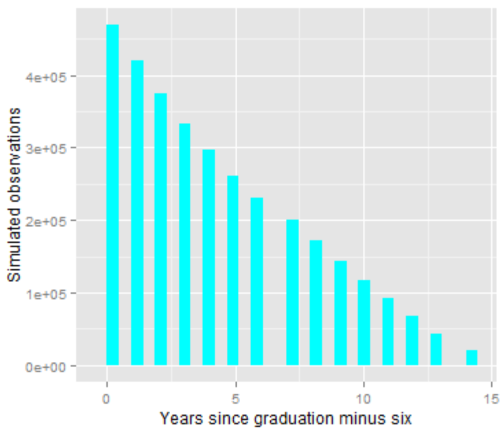
	Dependent variable:				
	log(income) ∗100				
	(1)	(2)	(3)	(4)	(5)
Exp	0.733*** (0.043)	2.481*** (0.060)	2.324*** (0.055)	2.222*** (0.060)	0.714*** (0.051)
VAGG	0.009 (0.038)				
Exp:VAGG	0.092*** (0.008)				
VAGU		0.837*** (0.070)			
Exp:VAGU		−0.347*** (0.011)			
VAGUL			0.858*** (0.063)		
Exp:VAGUL			−0.287*** (0.009)		
VAGUF				0.796*** (0.073)	
Exp:VAGUF				−0.311*** (0.012)	
USGG					0.108** (0.047)
Exp:USGG					0.064*** (0.010)
Constant	1,055.668*** (0.366)	1,051.367*** (0.517)	1,050.384*** (0.536)	1,051.881*** (0.506)	1,055.382*** (0.384)
Observations	3,243,898	3,243,898	3,243,898	3,243,898	3,243,898
R^2	0.146	0.146	0.146	0.146	0.146
Adjusted R^2	0.146	0.146	0.146	0.146	0.146
Residual Std. Error (df = 3243860)	50.127	50.118	50.119	50.122	50.128
F Statistic (df = 37; 3243860)	14,982.850***	15,018.710***	15,017.110***	15,005.060***	14,977.590***

Note: *p<0.1; **p<0.05; ***p<0.01

The following is the base regression for all of the above regressions: $log(income) * 100 = \beta_0 + \alpha_1 * Year_t + \beta_2 * Exp + B_4 * Degree_d$. Subscript t represents the year dummy of the wage in question and d represents each of the dummies for degrees. The columns include the following economic indicators: Virginia GSP at graduation (1), Virginia unemployment at graduation (2), a one year lag of Virginia unemployment at graduation (3), a one year lead of Virginia unemployment at graduation (4), and U.S. GDP growth at graduation (5). The regressions also include each indicator's interactions with experience. The actual coefficients on any year dummies or degree dummies and their interactions are not shown for ease of interpretability. *Important note:* Exp is defined differently for this regression for ease of interpretability. It is now years since graduation minus six, so that the intercept is zero (all observations within five years of graduation were dropped for these regressions).

servation. Ignoring Virginia GSP growth at graduation because of its lack of statistical significance[35], the observations with the mean amount of the newly formed Exp variable experience the following effects[36]: about 0.66% lower wages per point of extra Virginia unemployment at graduation, about 0.38% lower wages per point of extra lagged Virginia unemployment at graduation, about 0.54% lower wages per point of extra Virginia unemployment one year after graduation, and about 0.38% per extra point of U.S. GDP growth at graduation.

Figure 6: >5 Years since graduation



In contrast to previous sections, this analysis seems to show that there are persistent effects for graduating in bad economic times, but that they are less serious than in the years immediately following graduation. This is as expected, and is in line with my economic reasoning in the introduction.[37]

6 Conclusion

The results of the regressions, as well as using economic indicators to estimate the

[35] Note, however that the interaction with exp is significant at the 1% level and points in the right direction, representing about 0.4% higher wages per extra point of Virginia GSP growth for observations with the mean amount of this modified Exp variable.

[36] Remember that these estimated effects are in no way additive. They are just the estimations when considering a single economic indicator at graduation, and so must be considered separately.

[37] This section shows the leads and lags of economic indicators at graduation to be less significant than the indicators of the year of graduation, which is also comforting and goes against the results of the initial prediction regressions I performed.

test error of a given regression formula, seem quite robust in showing a statistically and economically significant effect on short term earnings of workers who graduate in poor economic times.

Further subdivision of the data based on extreme economic events[38] or recent graduates versus those who have already graduated for many years yields more insight, and shows the effect is largely isolated to the first few years after graduation and operates mainly through lower full time employment. If the level of full time employment is controlled for, the estimated effect of graduating in poor economic times largely disappears.

Comparing separate regressions which divide the observations to five or less years since graduation for a given cohort verus six or more years show that the short term effects of graduating in poor economic times are more serious, but that longer term effects are also present and still have strong economic and statisticial significance.

The challenges presented by working with data that was in the form of an income distribution for each observation were quite hard to overcome. They also present challenges of how to estimate and adjust for the true effects of the economic indicators. There seem to be two sources of bias that act to lower my estimated results. The first is the very slight lack of fit of the simulated log-normal distributions to the 50th percentile datapoint that I have, which very slightly reduces the estimation of the negative effects of graduating in poor economic times. The second is the inflation of the experience proxy I use for workers who graduate in poor economic times; this, too, seems to bias my results in the opposite direction that I find. Since my results still show large and significant effects of graduating in poor times, it seems that these biases are not a large cause for concern; however, they do create the possibility that my results represent a lower bound on the actual negative effects of graduating in poor economic times.

Although having individual level data would undoubtedly allow more precise estimates of the effect of being out of work on earnings when re-entering the workforce and in later years[39], the simulated dataset of workers who are in the workforce still allows a good understanding of the effect of poor economic times at graduation on wages during the years after graduation, both in the short and medium term.

I will conclude by saying that the economic effects of graduating in poor times were as expected and the results seem to be quite robust.

[38]Specifically the bursting of the tech bubble at the end of the 1990's, as well as the housing bubble later in the 2000's.

[39]My results seem to indicate that there is little long term effect of being unemployed or underemployed on earnings.

Appendices

References

Borjas, George 1999. 'The Economic Analysis of Immigration.' *Handbook of Labor Economics*, Volume 3, Part A, 1697-1760.

Cutler, Huang, and Muney 2015. 'When does education matter? The protective effect of education for cohorts graduating in bad times.' *Social Science and Medicine*, 127, 63-73.

Elsby, Michael, Bart Hobijn and Aysegul Sahin 2010. 'The Labor Market in the Great Recession.' *Brookings Papers on Economic Activity*, Spring, no. 2010, 1-69

Farber, Henry 2011. 'Job Loss in the Great Recession: Historical Perspective from the Displaced Workers Survey' *NBER Working Paper*, No. 17040, May 2011, 1984-2010.

Genda, Yuji, Kondo, Ayako, Ohta, and Souichi, 2010. 'Long-term effects of a recession at labor market entry in Japan and the United States.' *Journal of Human Resources*, 45, 157-196.

Kahn, Lisa 2010. 'The long-term labor market consequences of graduating from college in a bad economy.' *Labor Economics*, Volume 17, Issue 2, April 2010, 303-316.

Oreopoulous, Phil, Till von Wachter and Andrew Heisz 2012. 'The Short- and Long-Term Career Effects of Graduating in a Recession: Hysteresis and Heterogeneity in the Market for College Graduates.' *American Economic Journal: Applied Economics*, Vol. 4, Number 1.

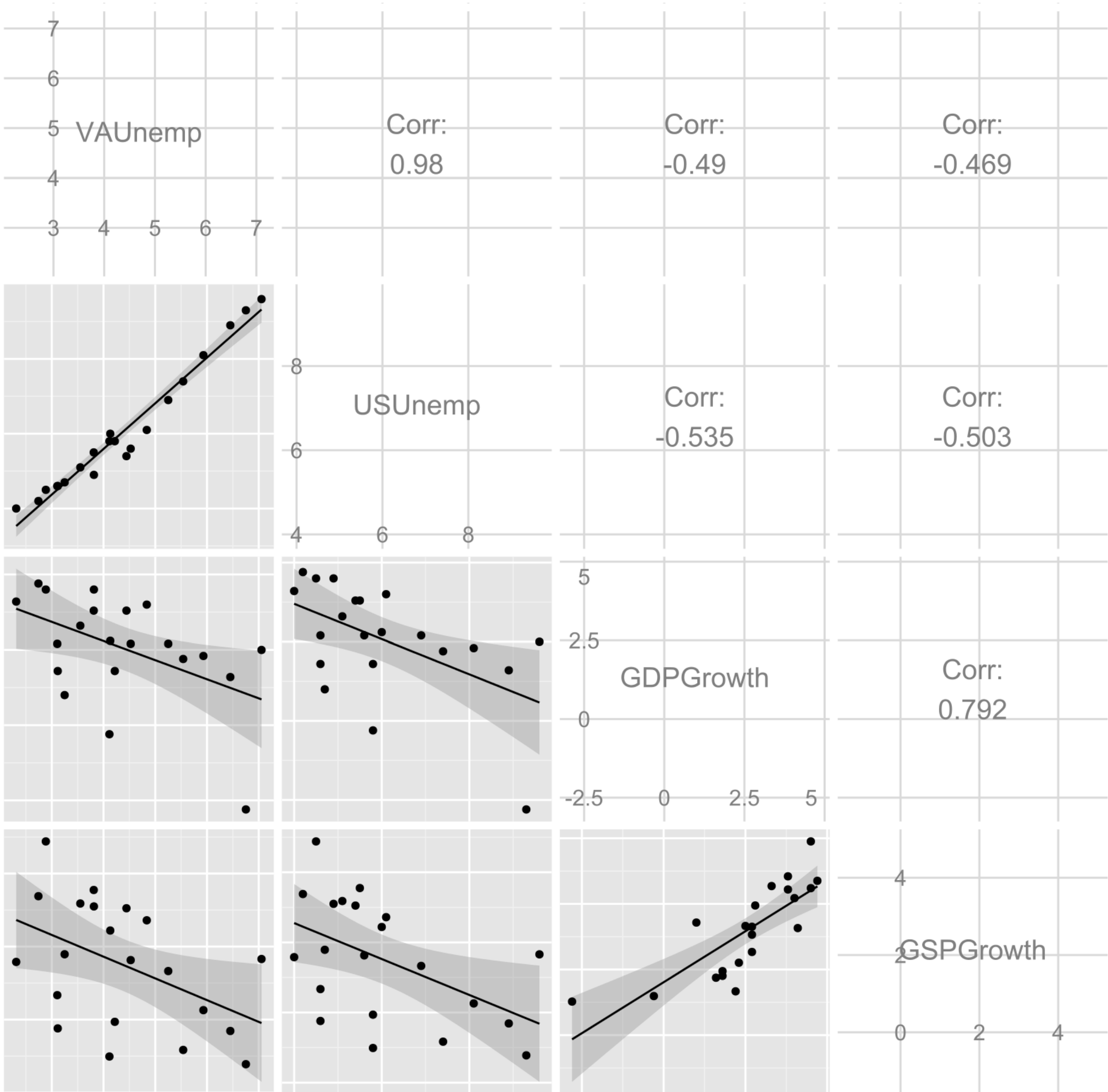
Ottaviano, Gianmarco and Giovanni Peri 2011. 'Rethinking the Effect of Immigration on Wages.' *Journal of the European Economic Association*, Online publication, October 25th 2011.

Steven J. Davis & Till Von Wachter 2011. "Recessions and the Costs of Job Loss," *Brookings Papers on Economic Activity, Economic Studies Program, The Brookings Institution*, vol. 43(2 (Fall)), 1-72.

Welch, Finis 1979 'Effects of Cohort Size on Earnings: The Baby Boom Babies' Financial Bust.' *Journal of Political Economy*, Vol. 87, No. 5, October.

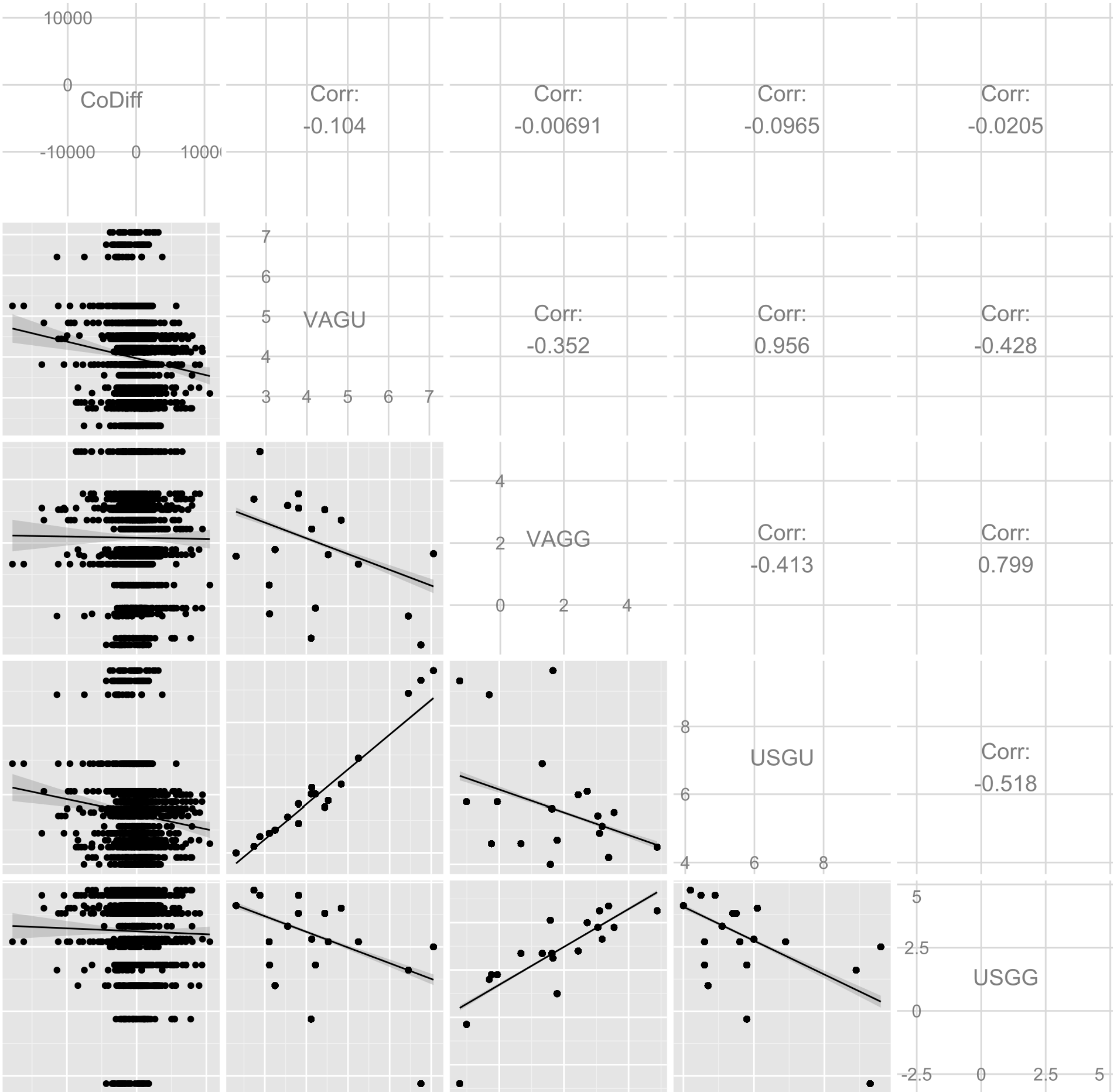
Summary Data, in Graphs

The correlation data and graphs between by year (1993-2013) between the Virginia unemployment rate, the US unemployment rate, US Gross Domestic Product growth, and Virginia Gross State Product Growth:

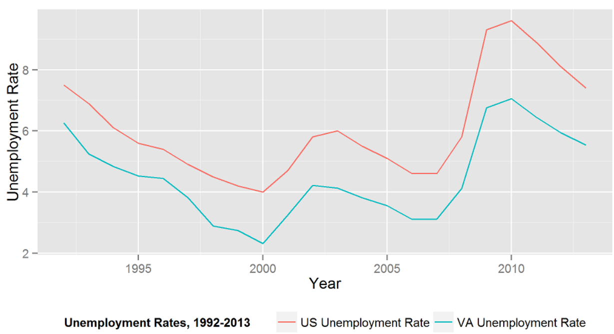


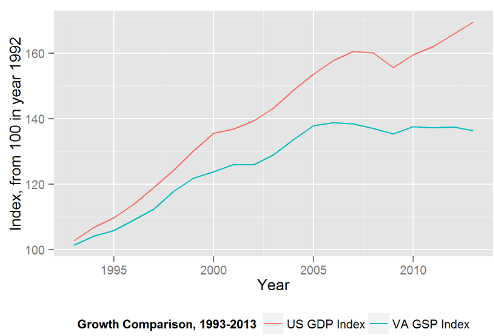
The correlation graphs between differences between graduation cohort median real wages (in 2009 dollars), holding experience constant, and Virginia and US unemployment and growth during the year of graduation.

- CoDiff - Median real wage cohort differential, controlling for work experience and degree level, in 2009 dollars
- VAGU - Virginia unemployment rate at graduation
- USGU - US unemployment rate at graduation
- VAGG - Virginia Gross State Product growth rate at graduation
- USGG - US Gross Domestic Product growth rate at graduation:



The following graphs are more or less self-explanatory:

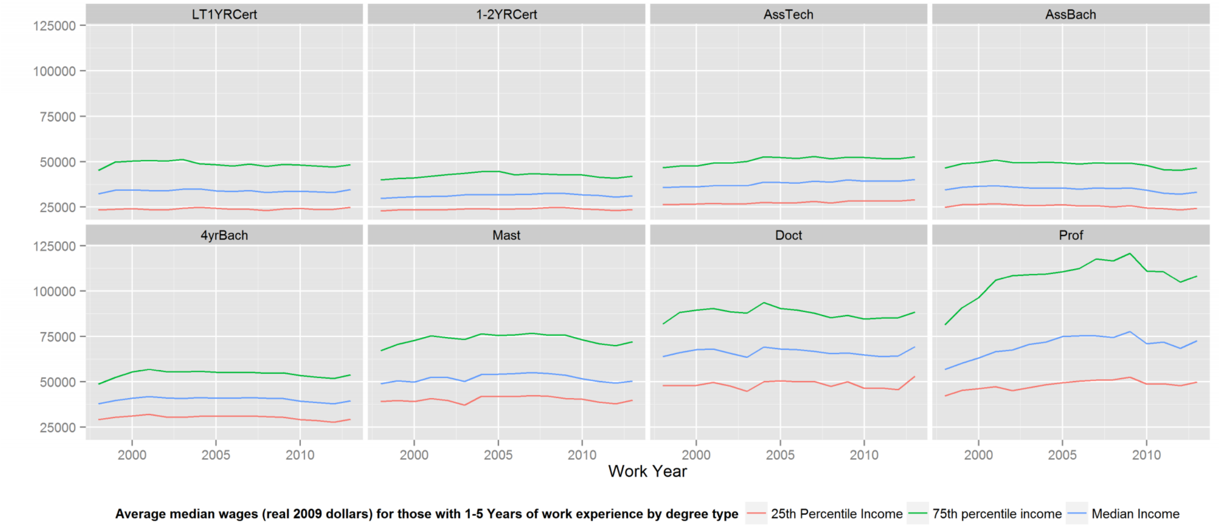


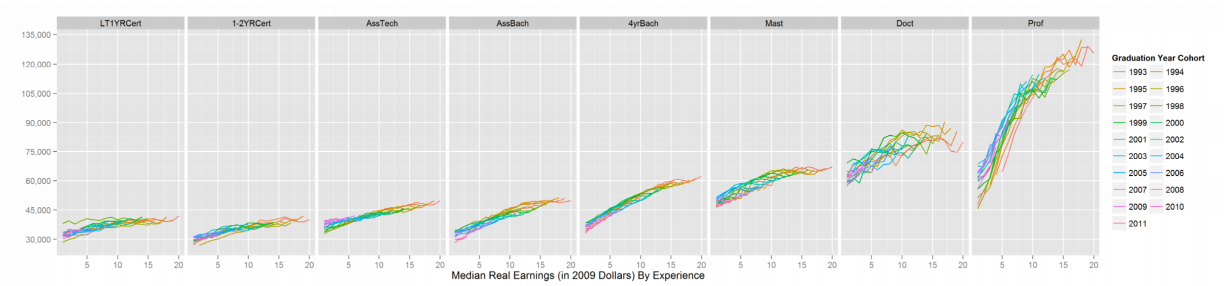


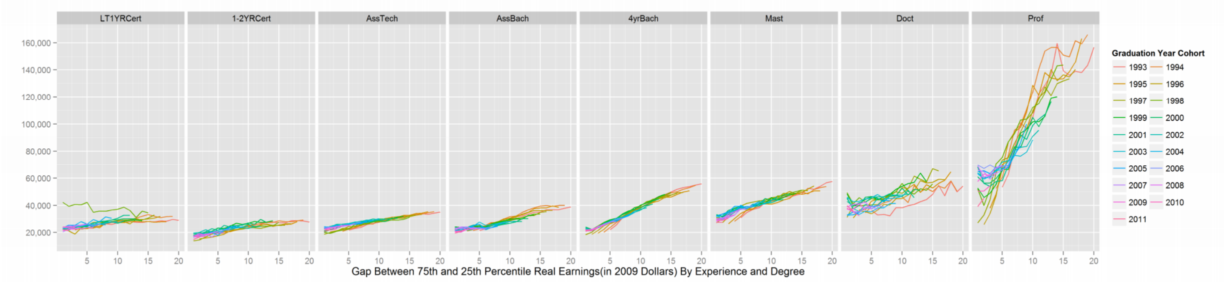
In most of the following graphs, there are the following degree types:

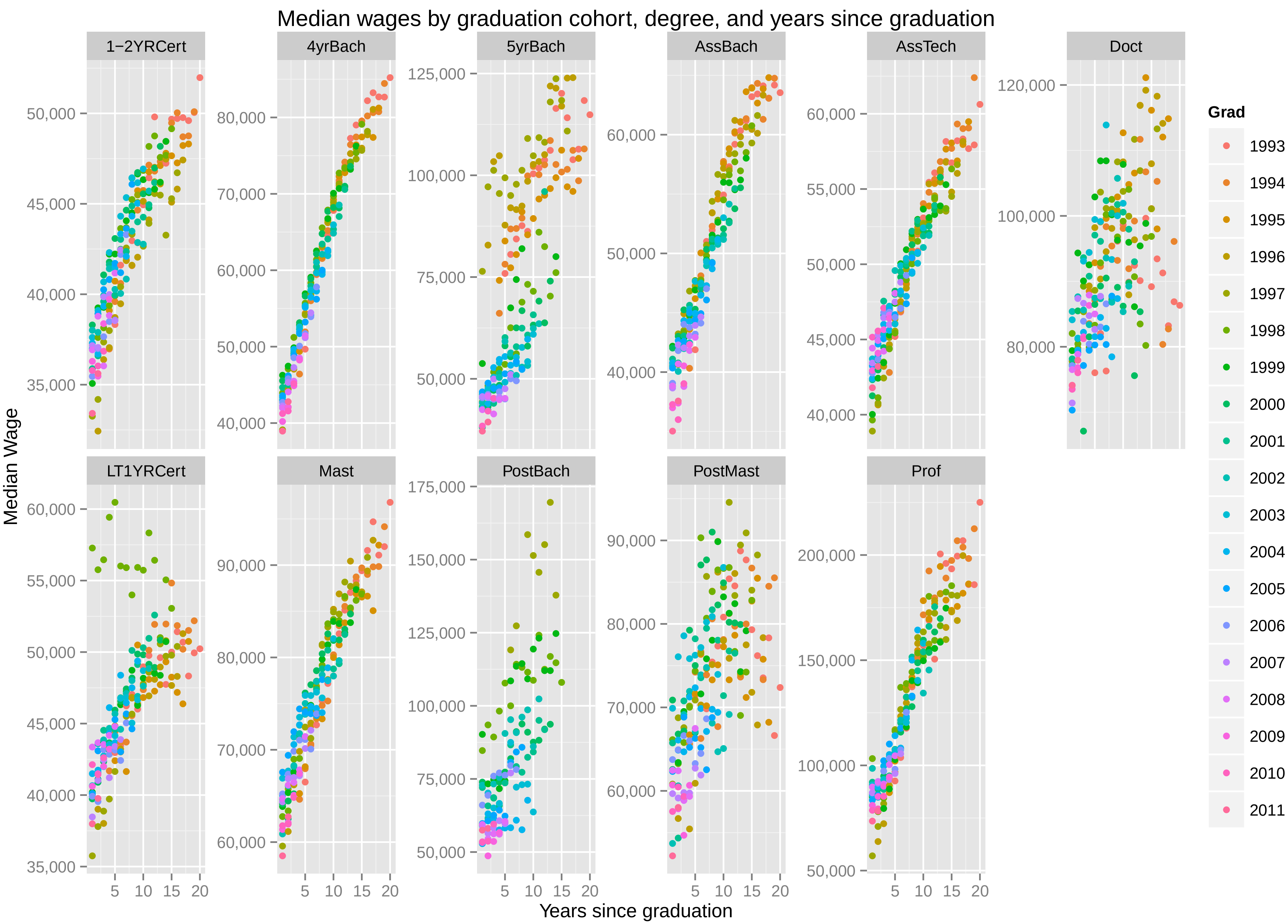
- LT1YR Certificate - A post high school certificate program that takes less than one year to complete
- 1-2YRCert - A post high school certificate program that takes one to two years to complete
- AssTech - An associate's degree with mostly technical class credits
- AssBach - An associate's degree with mostly college class credits
- Mast - Any master's degree
- Doct - Any doctorate's degree
- Prof - Professional degrees, such as law degrees and doctor of medicine degrees

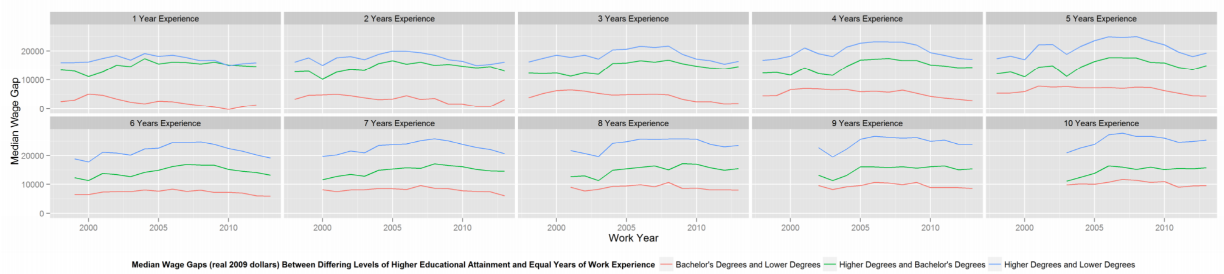
Here is the trend of real wages over time, controlling for experience. It seems that real wages were more or less flat, at least for those with 1-5 years experience (the only years of experience that can be compared across all of the wage years of the dataset), for all degree types except professional degrees during the period of the study.



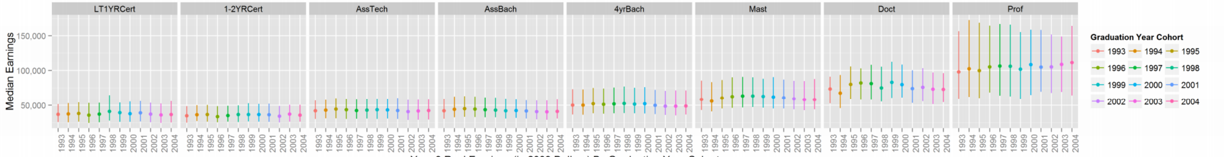




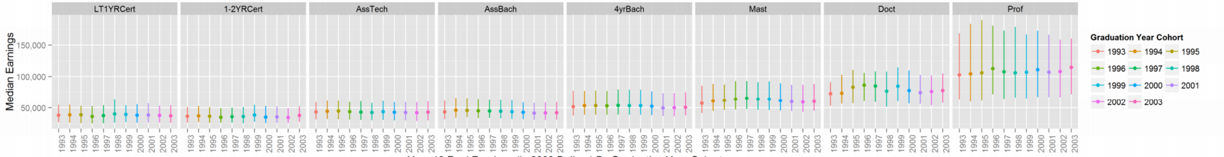

Median wages by graduation cohort, degree, and years since graduation
1−2YRCert
4yrBach
5yrBach
AssBach
AssTech
Doct
LT1YRCert
Mast
PostBach
PostMast
Prof
Median Wage
Years since graduation
Grad
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011



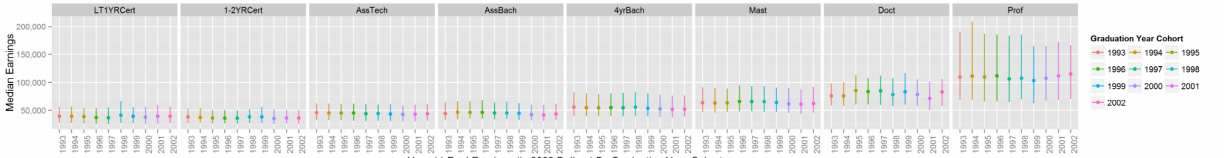
The full set of earnings by years of experience, degree type, and graduation year. The bottom of the small lines correspond to the 25th percentile of wages for the year of work and graduation cohort in question. The top corresponds to the 75th percentile, and the point is the median wage:



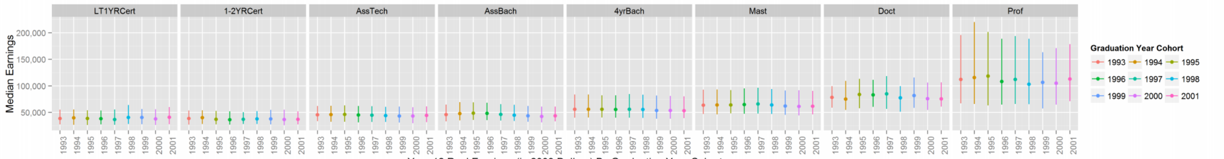
Year 9 Real Earnings (in 2009 Dollars) By Graduation Year Cohort



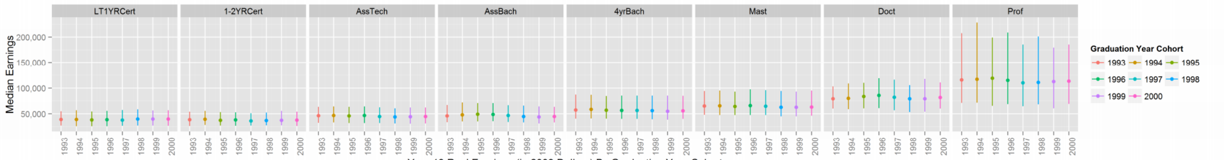
Year 10 Real Earnings (in 2009 Dollars) By Graduation Year Cohort



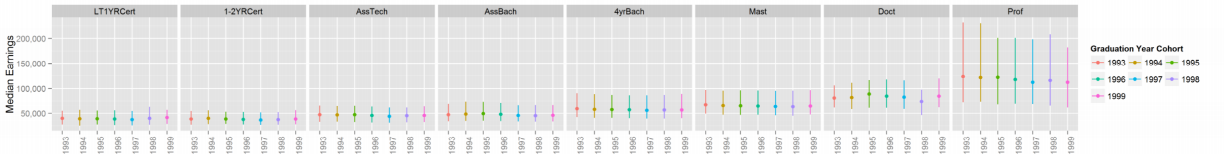
Year 11 Real Earnings (in 2009 Dollars) By Graduation Year Cohort



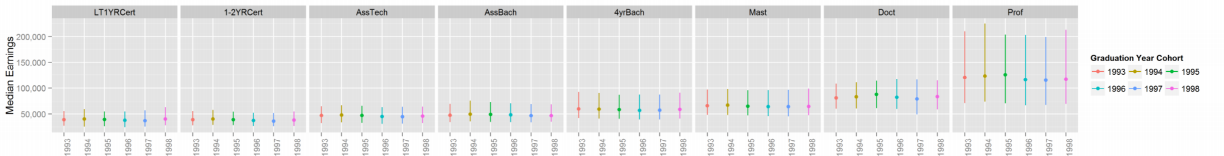
Year 12 Real Earnings (in 2009 Dollars) By Graduation Year Cohort



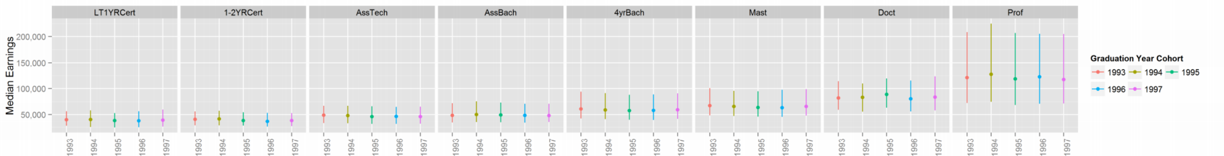
Year 13 Real Earnings (in 2009 Dollars) By Graduation Year Cohort



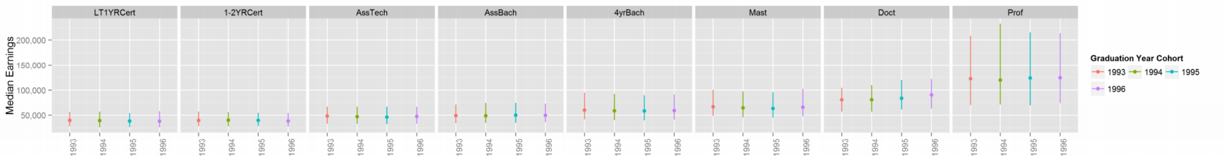
Year 14 Real Earnings (in 2009 Dollars) By Graduation Year Cohort



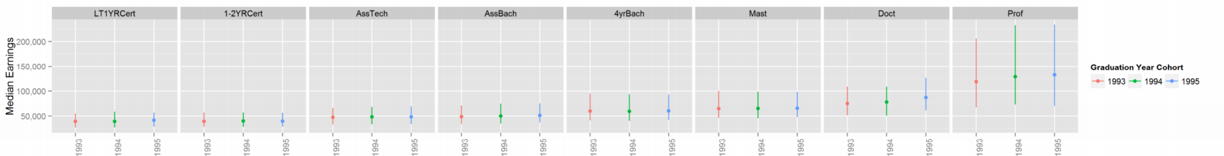
Year 15 Real Earnings (in 2009 Dollars) By Graduation Year Cohort



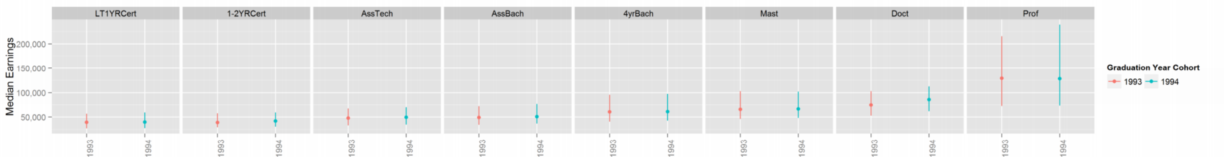
Year 16 Real Earnings (in 2009 Dollars) By Graduation Year Cohort



Year 17 Real Earnings (in 2009 Dollars) By Graduation Year Cohort



Year 18 Real Earnings (in 2009 Dollars) By Graduation Year Cohort



Year 19 Real Earnings (in 2009 Dollars) By Graduation Year Cohort

Tables Appendix

Table 1: Full Base Regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VAGU	-0.27 (0.02)	—	—	—	—	—	—	—
USGU	—	-2.06 (0.02)	—	—	—	—	—	—
VAGG	—	—	3.57 (0.02)	—	—	—	—	—
USGG	—	—	—	4.27 (0.02)	—	—	—	—
VAGUL	—	—	—	—	3.07 (0.02)	—	—	—
VAGUF	—	—	—	—	—	-2.97 (0.02)	—	—
USGUL	—	—	—	—	—	—	1.19 (0.02)	—
USGUF	—	—	—	—	—	—	—	-4.5 (0.02)
Observations:	4608775	4608775	4608775	4608775	4608775	4608775	4608775	4608775
k:	129	129	129	129	129	129	129	129
residual sd:	50.67	50.61	50.41	50.31	50.56	50.57	50.65	50.38
R-Squared:	0.11	0.112	0.119	0.123	0.114	0.114	0.111	0.12

Year dummies for 1999-2013, degree dummies, and all of their interactions were not included in the table, nor was the intercept. Degree names and numbers are listed in the regressions sections of the paper. All coefficients are scaled by 100. This was the regression run (EI refers to the economic indicator run. Capital B refers to a range of factors.):
$ln(SW_i) = \beta_0 + \beta_1 EI_g + \mathcal{B}_1 Yr_t + \mathcal{B}_2 Deg_d + \mathcal{B}_3(Deg_d * Yr_t)$

Table 2: Full Base Regressions with Experience Predictors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exp	2.45 (0.09)	2.44 (0.09)	2.5 (0.09)	2.5 (0.09)	2.52 (0.09)	2.43 (0.09)	2.5 (0.09)	2.41 (0.09)
I(Exp 2)	-0.07 (0.01)	-0.07 (0.01)	-0.07 (0.01)	-0.07 (0.01)	-0.07 (0.01)	-0.07 (0.01)	-0.07 (0.01)	-0.07 (0.01)
VAGU	-0.74 (0.02)	—	—	—	—	—	—	—
USGU	—	-0.6 (0.02)	—	—	—	—	—	—
VAGG	—	—	0.41 (0.02)	—	—	—	—	—
USGG	—	—	—	0.41 (0.02)	—	—	—	—
VAGUL	—	—	—	—	-0.58 (0.02)	—	—	—
VAGUF	—	—	—	—	—	-0.69 (0.02)	—	—
USGUL	—	—	—	—	—	—	-0.52 (0.02)	—
USGUF	—	—	—	—	—	—	—	-0.59 (0.02)
Observations:	4608775	4608775	4608775	4608775	4608775	4608775	4608775	4608775
k:	145	145	145	145	145	145	145	145
residual sd:	48.99	48.99	48.99	48.99	48.99	48.99	48.99	48.99
R-Squared:	0.168	0.168	0.168	0.168	0.168	0.168	0.168	0.168

See Table 1 for further description. This was the regression run:
$ln(SW_i) = \beta_0 + \beta_1 EI_g + \beta_2 Exp_i + \beta_3 Exp_i^2 + \mathcal{B}_1 Yr_t + \mathcal{B}_2 Deg_d + \mathcal{B}_3(Deg_d * Yr_t) + \mathcal{B}_4(Exp_i * Deg_d) + \mathcal{B}_5(Exp_i^2 * Deg_d)$

Table 3: Full Base Regressions with Experience Predictors and Experience - Economic Indicator Interactions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exp	2.81 (0.13)	2.82 (0.14)	2.81 (0.1)	3 (0.1)	3.08 (0.12)	2.47 (0.13)	3.33 (0.13)	2.17 (0.14)
I(Expˆ2)	-0.1 (0.01)	-0.1 (0.01)	-0.09 (0.01)	-0.1 (0.01)	-0.11 (0.01)	-0.08 (0.01)	-0.13 (0.01)	-0.06 (0.01)
VAGU	-0.62 (0.06)	—	—	—	—	—	—	—
Exp:VAGU	-0.08 (0.02)	—	—	—	—	—	—	—
VAGU:I(Expˆ2)	0.01 (0)	—	—	—	—	—	—	—
USGU	—	-0.53 (0.05)	—	—	—	—	—	—
Exp:USGU	—	-0.06 (0.02)	—	—	—	—	—	—
USGU:I(Expˆ2)	—	0.01 (0)	—	—	—	—	—	—
VAGG	—	—	0.99 (0.05)	—	—	—	—	—
Exp:VAGG	—	—	-0.18 (0.01)	—	—	—	—	—
VAGG:I(Expˆ2)	—	—	0.01 (0)	—	—	—	—	—
USGG	—	—	—	0.91 (0.04)	—	—	—	—
Exp:USGG	—	—	—	-0.18 (0.01)	—	—	—	—
USGG:I(Expˆ2)	—	—	—	0.01 (0)	—	—	—	—
VAGUL	—	—	—	—	-0.32 (0.06)	—	—	—
Exp:VAGUL	—	—	—	—	-0.12 (0.02)	—	—	—
VAGUL:I(Expˆ2)	—	—	—	—	0.01 (0)	—	—	—
VAGUF	—	—	—	—	—	-0.76 (0.06)	—	—
Exp:VAGUF	—	—	—	—	—	0 (0.02)	—	—
VAGUF:I(Expˆ2)	—	—	—	—	—	0 (0)	—	—
USGUL	—	—	—	—	—	—	-0.23 (0.05)	—
Exp:USGUL	—	—	—	—	—	—	-0.14 (0.02)	—
USGUL:I(Expˆ2)	—	—	—	—	—	—	0.01 (0)	—
USGUF	—	—	—	—	—	—	—	-0.74 (0.05)
Exp:USGUF	—	—	—	—	—	—	—	0.04 (0.02)
USGUF:I(Expˆ2)	—	—	—	—	—	—	—	0 (0)
Observations:	4608775	4608775	4608775	4608775	4608775	4608775	4608775	4608775
k:	147	147	147	147	147	147	147	147
residual sd:	48.99	48.99	48.99	48.99	48.99	48.99	48.99	48.99
R-Squared:	0.168	0.168	0.168	0.168	0.168	0.168	0.168	0.168

See Table 1 for further description. This was the regression run: $ln(SW_i) = \beta_0 + \beta_1 EI_g + \beta_2 Exp_i + \beta_3 Exp_i^2 + \mathcal{B}_1 Yr_t$
$+\mathcal{B}_2 Deg_d + \mathcal{B}_3(Deg_d * Yr_t) + \mathcal{B}_4(Exp_i * Deg_d) + \mathcal{B}_5(Exp_i^2 * Deg_d) + \mathcal{B}_6(Exp_i * EI_g) + \mathcal{B}_7(Exp_i^2 * EI_g)$

Table 4: Degree regressions for Virginia unemployment at graduation with experience

	Deg. 1	Deg. 2	Deg. 3	Deg. 4	Deg. 5	Deg. 6	Deg. 7	Deg. 8
VAGU	-1.08***	-0.69***	0.95***	-0.21**	-1.1***	-0.95***	-0.1	0.61***
	(0.129)	(0.124)	(0.067)	(0.095)	(0.032)	(0.049)	(0.177)	(0.205)
Exp	2.45***	2.93***	2.44***	4.53***	6.03***	3.05***	4.11***	7.36***
	(0.106)	(0.098)	(0.053)	(0.075)	(0.025)	(0.041)	(0.152)	(0.165)
I(Exp^2)	-0.07***	-0.07***	-0.04***	-0.08***	-0.14***	-0.06***	-0.12***	-0.18***
	(0.006)	(0.005)	(0.003)	(0.004)	(0.001)	(0.002)	(0.009)	(0.009)
								(*p<0.1; **p<0.05; ***p<0.01)
R^2	0.013	0.034	0.032	0.066	0.109	0.04	0.049	0.098
Residual sd	54.2	46.62	48.02	48	48.67	47.9	45.41	66.91
Observations	176093	153341	551236	343925	2337497	887395	57212	101924

Year dummies for 1999-2013 and the intercept were not included in the table. All coefficients are scaled by 100 for ease of presentation (including the residual s.d.). Degree names and numbers are as follows, with more detail in the data section: 1 - LT1YRCert, 2 - 1-2YRCert, 3 - AssTech, 4 - AssBach, 5 - 4yrBach, 6 - Mast, 7 - Doct, 8 - Prof
This was the regression performed: $ln(SW_i) = \beta_0 + \beta_1 VAGU_g + \beta_2 Exp_i + \beta_3 Exp_i^2 + \mathcal{B}Year_t$

Table 5: Degree regressions for US unemployment at graduation with experience

	Deg. 1	Deg. 2	Deg. 3	Deg. 4	Deg. 5	Deg. 6	Deg. 7	Deg. 8
USGU	-0.56***	-0.43***	0.82***	-0.47***	-0.88***	-0.82***	-0.24	0.43**
	(0.115)	(0.111)	(0.06)	(0.086)	(0.029)	(0.044)	(0.155)	(0.184)
Exp	2.52***	2.94***	2.48***	4.44***	6***	3.01***	4.06***	7.34***
	(0.108)	(0.099)	(0.053)	(0.076)	(0.026)	(0.042)	(0.154)	(0.168)
I(Exp^2)	-0.07***	-0.07***	-0.04***	-0.07***	-0.14***	-0.06***	-0.12***	-0.18***
	(0.006)	(0.005)	(0.003)	(0.004)	(0.001)	(0.002)	(0.009)	(0.009)
								(*p<0.1; **p<0.05; ***p<0.01)
R^2	0.013	0.034	0.032	0.065	0.109	0.04	0.049	0.098
Residual sd	54.21	46.63	48.03	47.99	48.67	47.9	45.4	66.96
Observations	176093	153341	551236	343925	2337497	887395	57212	101924

See table 4 and the data section for additional notes. This was the regression performed:
$ln(SW_i) = \beta_0 + \beta_1 USGU_g + \beta_2 Exp_i + \beta_3 Exp_i^2 + \mathcal{B}Year_t$

Table 6: Degree regressions for Virginia Gross State Product growth at graduation with experience

	Deg. 1	Deg. 2	Deg. 3	Deg. 4	Deg. 5	Deg. 6	Deg. 7	Deg. 8
VAGG	1***	0.32***	-0.54***	0.51***	0.5***	0.55***	-0.13	-0.43***
	(0.096)	(0.092)	(0.05)	(0.064)	(0.024)	(0.038)	(0.143)	(0.158)
Exp	2.42***	2.97***	2.46***	4.42***	6.06***	3.06***	4.14***	7.39***
	(0.106)	(0.099)	(0.054)	(0.076)	(0.026)	(0.041)	(0.152)	(0.167)
I(Exp^2)	-0.07***	-0.07***	-0.04***	-0.08***	-0.14***	-0.07***	-0.12***	-0.18***
	(0.006)	(0.005)	(0.003)	(0.004)	(0.001)	(0.002)	(0.009)	(0.009)
								(*p<0.1; **p<0.05; ***p<0.01)
R^2	0.014	0.034	0.032	0.066	0.109	0.04	0.049	0.098
Residual sd	54.19	46.62	48.02	47.99	48.67	47.9	45.39	66.94
Observations	176093	153341	551236	343925	2337497	887395	57212	101924

See table 4 and the data section for additional notes. This was the regression performed:
$ln(SW_i) = \beta_0 + \beta_1 VAGG_g + \beta_2 Exp_i + \beta_3 Exp_i^2 + \mathcal{B}Year_t$

Table 7: Degree regressions for US GDP growth at graduation with experience

	Deg. 1	Deg. 2	Deg. 3	Deg. 4	Deg. 5	Deg. 6	Deg. 7	Deg. 8
USGG	0.14	0.32***	-0.56***	0.47***	0.6***	0.54***	0.05	-0.72***
	(0.1)	(0.1)	(0.053)	(0.072)	(0.025)	(0.039)	(0.141)	(0.163)
Exp	2.62***	2.94***	2.48***	4.46***	6.03***	3.05***	4.09***	7.46***
	(0.107)	(0.1)	(0.054)	(0.076)	(0.026)	(0.042)	(0.153)	(0.168)
I(Exp^2)	-0.08***	-0.07***	-0.04***	-0.08***	-0.14***	-0.07***	-0.12***	-0.18***
	(0.006)	(0.005)	(0.003)	(0.004)	(0.001)	(0.002)	(0.009)	(0.009)
								(*p<0.1; **p<0.05; ***p<0.01)
R^2	0.013	0.034	0.032	0.066	0.108	0.04	0.05	0.099
Residual sd	54.21	46.62	48.02	47.98	48.67	47.91	45.4	66.92
Observations	176093	153341	551236	343925	2337497	887395	57212	101924

See table 4 and the data section for additional notes. This was the regression performed:
$ln(SW_i) = \beta_0 + \beta_1 USGG_g + \beta_2 Exp_i + \beta_3 Exp_i^2 + \mathcal{B}Year_t$

Table 8: Degree regressions for Virginia unemployment at graduation without experience

	Deg. 1	Deg. 2	Deg. 3	Deg. 4	Deg. 5	Deg. 6	Deg. 7	Deg. 8
VAGU	-1.69***	0.2	1.86***	2.54***	-0.53***	-1.37***	-1.31***	0.15
	(0.125)	(0.121)	(0.066)	(0.091)	(0.033)	(0.049)	(0.176)	(0.209)
								(*p<0.1; **p<0.05; ***p<0.01)
R^2	0.003	0.01	0.012	0.011	0.015	0.01	0.009	0.028
Residual sd	54.47	47.2	48.5	49.36	51.15	48.65	46.35	69.49
Observations	176095	153343	551238	343927	2337499	887397	57214	101926

See table 4 and the data section for additional notes. This was the regression performed:
$ln(SW_i) = \beta_0 + \beta_1 VAGU_g + \mathcal{B}Year_t$

Table 9: Degree regressions for US unemployment at graduation without experience

	Deg. 1	Deg. 2	Deg. 3	Deg. 4	Deg. 5	Deg. 6	Deg. 7	Deg. 8
USGU	-1.86***	-0.84***	0.51***	0.09	-2.78***	-2.38***	-2.31***	-2.67***
	(0.109)	(0.109)	(0.06)	(0.085)	(0.029)	(0.043)	(0.15)	(0.184)
								(*p<0.1; **p<0.05; ***p<0.01)
R^2	0.004	0.01	0.011	0.009	0.019	0.012	0.012	0.03
Residual sd	54.47	47.2	48.53	49.43	51.06	48.59	46.3	69.44
Observations	176095	153343	551238	343927	2337499	887397	57214	101926

See table 4 and the data section for additional notes. This was the regression performed:
$ln(SW_i) = \beta_0 + \beta_1 USGU_g + \mathcal{B}Year_t$

Table 10: Degree regressions for lagged Virginia unemployment at graduation without experience

	Deg. 1	Deg. 2	Deg. 3	Deg. 4	Deg. 5	Deg. 6	Deg. 7	Deg. 8
VAGUL	-0.07	2.03***	3.3***	4.84***	3.8***	1.16***	0.55***	4.67***
	(0.118)	(0.108)	(0.06)	(0.08)	(0.03)	(0.047)	(0.173)	(0.197)
								(*p<0.1; **p<0.05; ***p<0.01)
R^2	0.002	0.012	0.016	0.02	0.022	0.009	0.009	0.033
Residual sd	54.51	47.16	48.4	49.15	50.98	48.66	46.36	69.3
Observations	176095	153343	551238	343927	2337499	887397	57214	101926

See table 4 and the data section for additional notes. This was the regression performed:
$ln(SW_i) = \beta_0 + \beta_1 VAGUL_g + \mathcal{B}Year_t$

Table 11: Degree regressions for futured Virginia unemployment at graduation without experience

	Deg. 1	Deg. 2	Deg. 3	Deg. 4	Deg. 5	Deg. 6	Deg. 7	Deg. 8
VAGUF	-2.19***	-1.29***	0.22***	-0.6***	-3.95***	-3.11***	-2.96***	-3.48***
	(0.121)	(0.122)	(0.065)	(0.092)	(0.032)	(0.047)	(0.169)	(0.204)
								(*p<0.1; **p<0.05; ***p<0.01)
R^2	0.004	0.011	0.011	0.009	0.022	0.014	0.014	0.031
Residual sd	54.47	47.19	48.54	49.42	50.99	48.55	46.25	69.41
Observations	176095	153343	551238	343927	2337499	887397	57214	101926

See table 4 and the data section for additional notes. This was the regression performed:
$ln(SW_i) = \beta_0 + \beta_1 VAGUF_g + \mathcal{B}Year_t$

Table 12: Degree regressions for VA unemp. at and one year before/after graduation with experience

	Deg. 1	Deg. 2	Deg. 3	Deg. 4	Deg. 5	Deg. 6	Deg. 7	Deg. 8
VAGU	-1.51***	-1.35***	-0.17	1.28***	-0.9***	0.04	2.72***	-0.49
	(0.351)	(0.363)	(0.188)	(0.258)	(0.09)	(0.136)	(0.484)	(0.569)
VAGUL	-0.13	0.57**	0.41***	-0.76***	0.09	-0.42***	-2.27***	0.09
	(0.241)	(0.247)	(0.132)	(0.176)	(0.063)	(0.097)	(0.344)	(0.4)
VAGUF	0.71***	0.21	0.95***	-1.04***	-0.34***	-0.81***	-1.23***	1.35***
	(0.233)	(0.233)	(0.122)	(0.169)	(0.058)	(0.089)	(0.318)	(0.373)
Exp	2.53***	2.94***	2.5***	4.46***	6***	3***	4.06***	7.47***
	(0.108)	(0.099)	(0.054)	(0.076)	(0.026)	(0.042)	(0.154)	(0.168)
I(Exp^2)	-0.07***	-0.07***	-0.05***	-0.07***	-0.14***	-0.06***	-0.11***	-0.18***
10	(0.006)	(0.005)	(0.003)	(0.004)	(0.001)	(0.002)	(0.009)	(0.01)
								(*p<0.1; **p<0.05; ***p<0.01)
R^2	0.014	0.034	0.032	0.066	0.109	0.04	0.05	0.099
Residual sd	54.2	46.61	48.02	47.99	48.67	47.9	45.39	66.92
Observations	176091	153339	551234	343923	2337495	887393	57210	101922

See table 4 and the data section for additional notes. This was the regression performed:
$ln(SW_i) = \beta_0 + \beta_1 VAGU_g + \beta_2 VAGUL_g + \beta_3 VAGUF_g \mathcal{B}Year_t$

Table 13: Degree regressions for US unemp. at and one year before/after graduation with experience

	Deg. 1	Deg. 2	Deg. 3	Deg. 4	Deg. 5	Deg. 6	Deg. 7	Deg. 8
USGU	0.92***	-0.89**	-0.91***	-0.15	-0.22**	0.22*	0.57	-0.66
	(0.335)	(0.384)	(0.192)	(0.285)	(0.089)	(0.129)	(0.44)	(0.555)
USGUL	-1.53***	0.73***	0.98***	0.07	-0.31***	-0.63***	-0.98***	0.31
	(0.237)	(0.261)	(0.136)	(0.196)	(0.063)	(0.094)	(0.321)	(0.397)
USGUF	-0.32	-0.25	1.14***	-0.5***	-0.52***	-0.7***	-0.01	1.07***
	(0.219)	(0.243)	(0.124)	(0.181)	(0.057)	(0.084)	(0.29)	(0.359)
Exp	2.51***	2.88***	2.54***	4.39***	5.95***	2.95***	4.06***	7.48***
	(0.111)	(0.101)	(0.054)	(0.078)	(0.026)	(0.042)	(0.157)	(0.171)
I(Exp^2)	-0.07***	-0.07***	-0.05***	-0.07***	-0.14***	-0.06***	-0.12***	-0.18***
	(0.006)	(0.005)	(0.003)	(0.004)	(0.001)	(0.002)	(0.009)	(0.01)
						(*p<0.1; **p<0.05; ***p<0.01)		
R^2	0.013	0.034	0.032	0.066	0.109	0.04	0.05	0.099
Residual sd	54.19	46.62	48.02	48	48.67	47.9	45.38	66.95
Observations	176091	153339	551234	343923	2337495	887393	57210	101922

See table 4 and the data section for additional notes. This was the regression performed:
$ln(SW_i) = \beta_0 + \beta_1 USGU_g + \beta_2 USGUL_g + \beta_3 USGUF_g \mathcal{B}Year_t$

Table 14: Degree regressions for VA unemp. one year before and after graduation with experience

	Deg. 1	Deg. 2	Deg. 3	Deg. 4	Deg. 5	Deg. 6	Deg. 7	Deg. 8
VAGUL	-0.96***	-0.18	0.31***	-0.1	-0.42***	-0.4***	-0.68***	-0.19
	(0.143)	(0.146)	(0.081)	(0.107)	(0.037)	(0.057)	(0.201)	(0.237)
VAGUF	-0.1	-0.47***	0.86***	-0.38***	-0.79***	-0.79***	0.12	1.07***
	(0.146)	(0.149)	(0.078)	(0.108)	(0.037)	(0.056)	(0.2)	(0.237)
Exp	2.5***	2.94***	2.49***	4.48***	6***	3***	4.02***	7.45***
	(0.108)	(0.1)	(0.054)	(0.076)	(0.026)	(0.042)	(0.154)	(0.168)
I(Exp^2)	-0.07***	-0.07***	-0.05***	-0.08***	-0.14***	-0.06***	-0.11***	-0.18***
	(0.006)	(0.005)	(0.003)	(0.004)	(0.001)	(0.002)	(0.009)	(0.01)
						(*p<0.1; **p<0.05; ***p<0.01)		
R^2	0.013	0.034	0.032	0.065	0.109	0.04	0.049	0.099
Residual sd	54.21	46.64	48.01	48	48.67	47.89	45.4	66.92
Observations	176092	153340	551235	343924	2337496	887394	57211	101923

See table 4 and the data section for additional notes. This was the regression performed:
$ln(SW_i) = \beta_0 + \beta_1 VAGUL_g + \beta_2 VAGUF_g \mathcal{B}Year_t$

Table 15: Degree regressions for US unemp. one year before and after graduation with experience

	Deg. 1	Deg. 2	Deg. 3	Deg. 4	Deg. 5	Deg. 6	Deg. 7	Deg. 8
USGUL	-1***	0.21	0.43***	0.01	-0.44***	-0.49***	-0.64***	-0.04
2	(0.126)	(0.132)	(0.073)	(0.1)	(0.033)	(0.05)	(0.174)	(0.21)
USGUF	0.18	-0.69***	0.68***	-0.59***	-0.63***	-0.59***	0.29*	0.67***
4	(0.131)	(0.141)	(0.073)	(0.103)	(0.034)	(0.05)	(0.174)	(0.212)
Exp	2.51***	2.89***	2.54***	4.39***	5.96***	2.95***	4.1***	7.46***
6	(0.11)	(0.101)	(0.054)	(0.078)	(0.026)	(0.042)	(0.157)	(0.171)
I(Exp^2)	-0.07***	-0.07***	-0.05***	-0.07***	-0.14***	-0.06***	-0.12***	-0.18***
8	(0.006)	(0.005)	(0.003)	(0.004)	(0.001)	(0.002)	(0.009)	(0.01)
						(*p<0.1; **p<0.05; ***p<0.01)		
R^2	0.013	0.034	0.032	0.066	0.109	0.04	0.05	0.099
Residual sd	54.19	46.63	48.02	47.99	48.67	47.9	45.42	66.94
Observations	176092	153340	551235	343924	2337496	887394	57211	101923

See table 4 and the data section for additional notes. This was the regression performed:
$ln(SW_i) = \beta_0 + \beta_1 USGUF_g + \beta_2 USGUF_g \mathcal{B}Year_t$

Acknowledgements

There are several people and programs I would like to thank for helping support me during my time at Peking University HSBC Business School. The first is PHBS itself, for two years of scholarship and some excellent classes. Without the scholarship I would have been unable to have completed school because of the financial constraints of raising a family.

The second and most important is my wife - without her support over the last two years and her loving care for our daughter and myself, I would not have been able to devote as much time and energy as I needed to complete my degree. Thank you from the bottom of my heart; I love you so much.

The third is my mother, for always believing in me and supporting me no matter what. Great thanks go out to you as well, Mom; I love you so much, too.

I would also like to thank several excellent professors and my advisor - Zhang Yilin for her excellent math course and kind words, David Ong for his thought provoking courses and demand for original thought, and Domenico Tarzia for his excellent coverage of the field of research from behavioral economics and finance, and my advisor, Christopher Balding, for the freedom to take this thesis in whatever direction I so chose.

I also want to give thanks to fellow classmates Max Karnfelt for his sticking through multiple tough courses with me and giving feedback on many ideas over the last couple of years, and to David Garcia for recommending the program and acting as a sounding board when I had many questions as my graduate student career began.

Finally, I wanted to thank Stanford professors Trevor Hastie and Robert Tibshirani for their excellent textbook and free online course through Stanford Online, Statistical Learning. It put me on the path to learning more advanced predictive techniques and the programming language R.

北京大学学位论文原创性声明和使用授权说明

原创性声明

本人郑重声明： 所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不含任何其他个人或集体已经发表或撰写过的作品或成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本声明的法律结果由本人承担。

论文作者签名：

日期： 年 月 日

学位论文使用授权说明

（必须装订在提交学校图书馆的印刷本）

本人完全了解北京大学关于收集、保存、使用学位论文的规定，即：

- 按照学校要求提交学位论文的印刷本和电子版本；
- 学校有权保存学位论文的印刷本和电子版，并提供目录检索与阅览服务，在校园网上提供服务；
- 学校可以采用影印、缩印、数字化或其它复制手段保存论文；
- 因某种特殊原因需要延迟发布学位论文电子版，授权学校□一年/□两年/□三年以后，在校园网上全文发布。

（保密论文在解密后遵守此规定）

论文作者签名：

导师签名：

日期： 年 月 日

北京大学 攻读硕士学位研究生选题报告审核表

Evaluation form for research proposal of Graduation Thesis （本表一式一份存学校档案）

院、系：	专　　业：
姓　名：	研究方向：
学　号：	导师姓名：
拟定学位论文题目：	Major: Financial Economics
Department: PHBS	Research Topic:
Student Name:	Short and Medium Term Income Effects
Jeremy Michael Ward Schutte	of Graduating in a Bad Economy
Student No.: 1401213315	Advisor's Name: Christopher Balding

Title of your Graduation Thesis: Graduating in a Bad Economy: Some Short and Medium Term Income Effects of Entering the Work Force During Economic Downturns

本人陈述：选题来源、研究意义、国内外研究状况、主要研究内容、拟采取的研究方法、预期研究结果和论文写作计划等 Personal statement. It shall include your sources, research significance, research background (home & abroad), research content, research methodologies, expected results, timeline, etc.

I will use about two decades of Virginia higher degree cohort wage data, as well as some BLS wage data, and the appropriate economic employment and growth indices (state and national measures), to try and measure the short term and medium term wage effects of graduating into a poor economy.

Since the primary data is not by individual, but by distributions of workers by graduation year and major, I will need to treat the data as log-normal distributions of wages, and represent the individuals within through repeated simulations from the given distributions (which contain 25th percentile, 50th percentile, and 75th percentile income for each graduation-year / degree type / wage-year).

With full time work outside of school and two more classes remaining, I expect to finish within the required time frame, or perhaps a little ahead of time. I already have a good bit of the research done, including data collection, manipulation, and crude simulation, but I will need to collect additional data and do quite a bit more work to see what can be learned from the data, and figuring out how to estimate the standard errors of these simulated data is a challenge I am not sure I will overcome.

I expect to find a significant and persistent negative effect on wages from graduating at an inopportune time and already have preliminary results suggesting this. Although the research is not novel, it is significant to anyone who cares about the effects of a bad economy on their future wages, and who might choose whether or not to immediately enter the work force based on such information.

I have no research background outside of what I have done at PHBS, and studied online and offline at the same time through Stanford's elements of statistical learning course and other materials related to learning the R statistical and programming language.

指导教师对选题报告的意见 Advisor's comments:： None received