1-30-2013

ARE GRADUATE STUDENTS RATIONAL? MACROECONOMIC DETERMINANTS OF GRADUATE ENROLLMENT AND COMPLETIONS IN THE BIOMEDICAL SCIENCES

John Clack

John William Clack

Candidate

Economics

Department

This thesis is approved, and it is acceptable in quality and form for publication:

Approved by the Thesis Committee:

Dr. Margaret Blume-Kohout, Chairperson

Dr. Kate Krause

Dr. David van der Goes

ARE GRADUATE STUDENTS RATIONAL? MACROECONOMIC DETERMINANTS OF GRADUATE ENROLLMENT AND COMPLETIONS IN THE BIOMEDICAL SCEINCES

by

JOHN WILLIAM CLACK

B.S., PUBLIC POLICY, OLIVET NAZARENE UNIVERSITY, 2007

M.A., ECONOMICS, UNIVERSITY OF NEW MEXICO, 2012

THESIS

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Master of Arts in Economics

The University of New Mexico
Albuquerque, New Mexico

December, 2012

DEDICATION

For Dr. William F. Wright; he believed in education and in me.

ACKNOWLEDGEMENTS

I would like to express deep appreciation to my committee chair, Professor Meg Blume-Kohout. Her experience and knowledge in the field of economic policy research helped to overcome difficulties in the data collection process and lead me to more valid results. Her careful guidance at every step made palatable a daunting research process.

I would also like to thank my committee members, Professor Kate Krause and Professor David van der Goes. Their input has been invaluable in the refining of this thesis into a final product.

Finally, I would like to thank my family. Uncle Bill encouraged me to take the first step into graduate school and enthusiastically supported me throughout. Mom and dad leant me their perseverance through encouragement.

ARE GRADUATE STUDENTS RATIONAL?
MACROECONOMIC DETERMINANTS OF GRADUATE ENROLLMENT AND COMPLETIONS IN THE BIOMEDICAL SCIENCES

by

John William Clack

B.S., Public Policy, Olivet Nazarene University, 2007
M.A., Economics, University of New Mexico, 2012

ABSTRACT

The National Institutes of Health (NIH) budget doubling from 1998 through 2003 stimulated demand for biomedical scientists, increasing both relative wages and employment. However, because research doctorates in these fields may take six years or more to complete, there is a substantial lag in the labor supply response to changing market conditions. Rational expectations models assume that prospective graduate students can forecast their expected future wages, taking into account other students' likely responses and thus also future employment levels. However, prior research on student enrollment and degree completion in science and engineering fields suggests that market conditions at the time of enrollment are taken as proxies for future conditions. Previous studies also suggest that graduate student enrollment and PhD completions may be responsive to changes in availability and mechanism of financial support. This thesis uses instrumental variables estimation on time-series data including biomedical scientists' wages and employment, bachelor's degrees and PhD completions, and NIH and private industry research funding, to examine responsiveness of labor supply to changing market conditions, and particularly to changes in NIH funding levels. We find that graduate student enrollment and PhD completions are highly responsive to NIH

financial support, to current trends in job availability at time of enrollment, and to expected earnings.

Table of Contents

I. INTRODUCTION

The number of science and engineering (S&E) PhDs awarded annually in the United States has been rising consistently from the 1960s to the present. Life sciences specifically have grown to see more than 11 thousand degrees granted in 2010, from fewer than three thousand graduates in 1966.

From 1998 through 2003, Congress effectively doubled the total combined budget for the U.S. National Institutes of Health (NIH), from $18.3 billion to $36.4 billion in constant 2010 dollars. The resulting increase in demand for biomedical scientists was largely reflected in dramatically increased demand for postdoctoral researchers; temporary appointments that have become a near-ubiquitous career waypoint for freshly minted PhDs. This is in contrast to the academic market of the 1970s (Stephan, 2012). The rapid growth in demand for postdocs was almost entirely met by an influx of international PhDs moving to the United States specifically to fill the need (Garrison, 2005). However, the increase in demand for labor apparently had no effect on more permanent, tenured and tenure-track faculty positions (Blume-Kohout, 2012).

Economists are coming to believe that the life science labor market may be saturated – providing a career path where scientists forfeit significant earnings for tenuous career prospects (Stephan, 2012). Given that trainees have become a primary engine of research in the lab-focused biomedical education system, universities may be motivated to overproduce scientists. Biomedical students graduating in the late seventies had an approximate 30% chance of receiving a tenure-track position by 1985. This dropped to 20% for students graduating in the late eighties and surveyed in 1995 (Committee on Dimensions, 1998). Where increased biomedical science investment was expected to create long-term job opportunities within these fields, it may not be doing so. For the group of elite 1992-1994 National Research Service Award (NRSA) winners it yielded little career assistance (Levitt, 2010). As of 2010, more than a decade after their graduations, only approximately 25% of those students had achieved tenure at a university. That trend has continued and the biomedical labor market still struggles to absorb fully-trained biomedical scientists into full-time and tenure-track faculty research positions (Stephan, 2012).

The changing life science labor market may alter job prospects at every level. While established end-career prospects may motivate students to matriculate, funding opportunities also determine enrollment. Garrison et. al. (2005) found that the number of research assistants funded by the NIH held steady from 1995 through 2000 - the beginning of the NIH budget doubling. In contrast, we found that the number of RAs was already increasing in that time period.

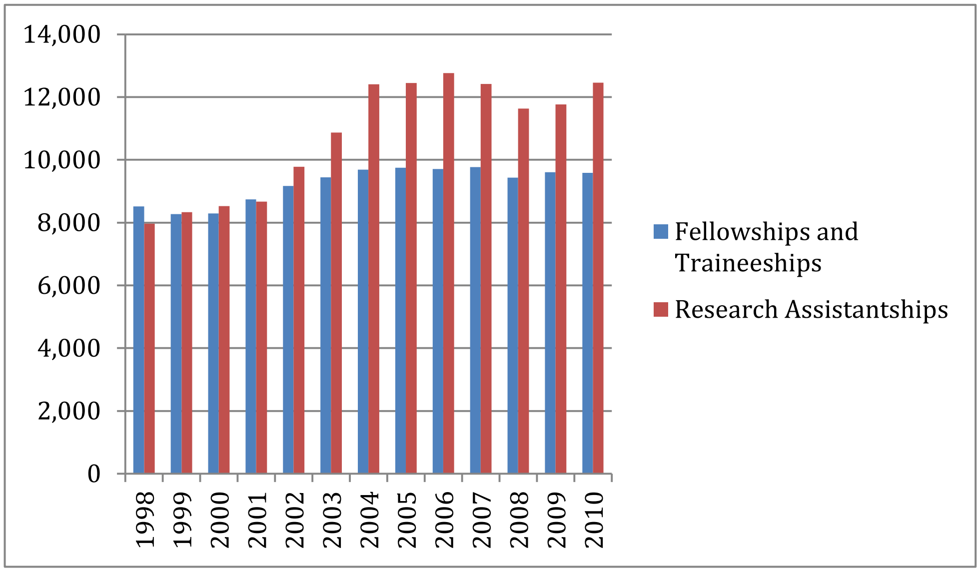

Figure 1: Students funded by NIH – Data from NIH Data Book, 2012 and NCSES

The appearance of declining employment rates in the biomedical sciences elicits concerns over institutional stewardship of the future biomedical workforce. Do universities care about the ultimate placement of their graduates? Do funding institutions realize that the current system may be preparing the majority of graduate students, given that most are hoping for a tenured career in academia, for undesired career outcomes? The NIH Advisory Committee to the Director (ACD) Biomedical Research Workforce Working Group recently issued a report addressing the NIH's understanding of these issues. NIH administration seems to recognize the current state of the life science labor market and the task force ultimately recommends limits on the number of funding years available to each student and a move from RA grants to training grants: the better to guide and homogenize training quality across institutions (Tilghman, 2012). At the institutional level it may be more difficult to coordinate enrollment based on the job market. Institutions are given to supplementing or withdrawing graduate student funding counter-cyclically to the National Institutes of Health (Ehrenberg, 1993).

Present concerns about the over-production and the quality of biomedical scientists come on the heels of a pervasive historical fears that the United States is under-producing in the sciences. *Rising Above the Gathering Storm*, a report from the National Research Council, paints a global landscape in which the United States is falling behind other nations in scientific education, research, and productivity (2005). Should we fail to remain competitive, the report argues, the United States will de-facto sacrifice economic and political status to other nations.

This paper extends research which has previously primarily focused on undergraduates, into the market for PhD-trained workers. Whether graduate students, and specifically those in life sciences, choose graduate training in response to current or future expected future labor market conditions has yet to be firmly established. Degrees in biomedical science appear so specialized, and require such a sacrifice of early career earnings, that it begs the question whether students who have chosen to study in these fields are doing so out of a natural affinity for the subject or the non-pecuniary benefit, rather than simply a desire to maximize their lifetime earnings. Other graduate degrees, such as the MBA, require a shorter time investment in education to produce greater expected lifetime earnings. Furthermore, since pursuit of a PhD is a long-term career investment, often in excess of five years, market conditions at graduation may be vastly different from market conditions at the time of enrollment. In an ideal labor market, students considering a graduate degree would make informed decisions based on their best understanding of their future career prospects including inherent non-pecuniary benefits.

While changes in graduate and undergraduate enrollment are broadly a function of population growth, there are important policy questions to be addressed in short-term enrollment trends. It is generally accepted that the university has acted within the labor market as a refuge from down economies for undergraduates. A workforce facing high unemployment rates or low wages will return to the university (Betts and McFarland, 1995). While we are measuring the macroeconomic determinants of graduate biomedical matriculation, we also take into account other influences, such as the availability of funding. Previous research has found that interest rates affect the attractiveness of school to potential students (Dellas and Koubi, 2003) and the sources of funding available to a student affect his or her ability to complete a degree (Ampaw, 2010). Cohort size has also proven to be important. An influx of foreign students into a given field will result in lower wages and employment rates within that field (Borjas, 2006). Students, undergraduates especially, have proven themselves to be vulnerable to business cycles and to be rational actors in the labor market (Dellas and Sakellaris, 2003). When the opportunity-cost of education is low, that is when unemployment rates are high or wages are low, people choose to pursue education in higher numbers. Student demographics also have some effect on matriculation. Male students with high undergraduate GPAs are more responsive to business cycles than are students with low GPAs and females (Bedard and Herman, 2008).

This thesis analyzes recent trends in the market for biomedical science PhDs and evaluates the extent to which graduate student enrollments reflect rational foresight regarding future market conditions. To a first approximation, one might expect current, short-run demand shifts for biomedical scientists to affect current new PhD enrollments, as such shifts may affect students' perceptions of longer-run career prospects. On the other hand, the particular policy shift examined here—the NIH budget doubling—was announced in advance, so rational agents might have anticipated that early trends in wages and job growth would not persist.

Using time-series data from a variety of nationally representative surveys including counts of first-time graduate student enrollments, PhD completions, estimates of biomedical scientists' and alternative occupations' salaries and employment rates, and both NIH and biopharmaceutical industry R&D expenditures over the period 1998 through 2010, we evaluate the relative importance of various incentives which draw students towards a career in biomedical sciences. In particular, we assess how changes in NIH R&D expenditures influence both demand and supply in the market for biomedical sciences workforce. We find that the increase in demand during the budget doubling period, reflected both in higher relative wages and in job growth for biomedical sciences occupations, appears to have lured students to enroll in graduate programs. However, students who entered in 1998 were graduating just as the expansion came to an end, and since 2006 the number of PhDs produced each year by U.S. biomedical sciences programs has generally exceeded the growth in U.S. jobs.

Empirically, Ryoo and Rosen (2004) found fluctuations in demand for professional engineers were substantially driven by changes in R&D and defense-related public expenditures. Along those lines, in this paper we evaluate responsiveness of demand for biomedical scientists to changes in NIH appropriations and industrial biopharma R&D.

I.A. Summary of Findings

We find that graduate student enrollment and PhD completions are responsive to expected earnings and to employment rates in biomedical fields. Enrollment and completions are especially responsive to NIH funding levels. Because NIH R&D is influential on the supply side and the demand side, it lacked the exogeneity necessary to identify market effects.

II. EMPIRICAL APPROACH

In the analysis that follows, we combine data from various government data sources on wages, employment, and education. Due to data limitations and strong evidence of a structural break after the NIH budget doubling concluded, our time series here is limited to 12 years, with some models using fewer observations due to inclusion of lagged variables. Nonetheless, our results are consistent with expectations from previous literature, and demonstrate a valid solution to the unique empirical challenge posed. The approach demonstrated thus should be useful to future analyses, when more data become available. Our estimations employ first-differenced OLS models in order to mitigate autocorrelation, and instrumental variables estimation to avoid simultaneity bias.

II.A. Data

Bureau of Labor Statistics Occupational Employment Statistics

Average salary estimates and employment statistics for biomedical scientists and alternative career fields were calculated for years 1999 through 2010 using data from the Bureau of Labor Statistics' (BLS) Occupational Employment Statistics (OES) survey. The Bureau of Labor Statistics, an agency within the U.S. Department of Labor, collects OES data semiannually from 200,000 business establishments. Wage estimates in the OES are produced by combining current-period survey data with that collected from the previous two surveys. That is, estimates reported for May 2012 would be calculated using data from the May 2012, November 2011, and May 2011 surveys. Due to occupation code reclassifications after 1998, this dataset is our limiting series. All wage estimates are inflated to constant 2010 dollars using the chained consumer price index.

To generate the biomedical scientist average wage time-series, we followed the occupation classification from the recent NIH Advisory Committee to the Director (ACD) task force report. These occupations are detailed in Table 1. The average wage for each year is a weighted average, where the weights are determined by the number of people in that occupation code, relative to the total sum of employees across all biomedical sciences occupations.

In Figure 2 we see that biomedical salaries began to grow more quickly just as the NIH budget doubling came to an end. There was also a spike in biomedical science compensation around 2001, in the middle of the NIH program of expansion.

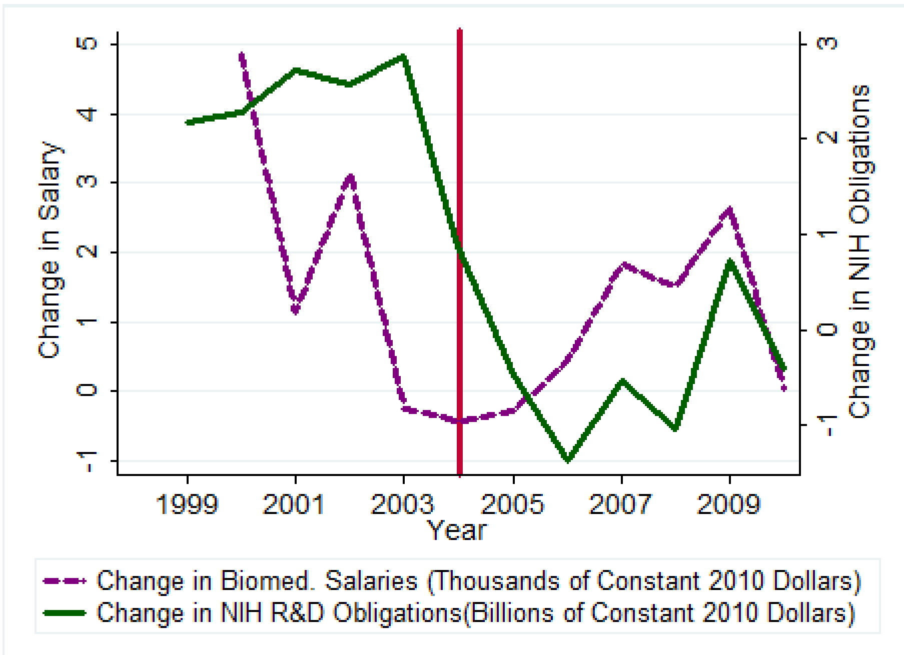

Figure 2: Biomedical Salary and NIH Obligations

Table 1. Biomedical Scientist Occupation Codes for BLS OES Data	
Occupation Code	Occupation Title
11-9121	Natural Sciences Managers
17-2031	Biomedical Engineers
19-1021	Biochemists and Biophysicists
19-1022	Microbiologists
19-1023	Zoologists and Wildlife Biologists
19-1041	Epidemiologists
19-1042	Medical Scientists, except Epidemiologists
19-4021	Biological Technicians
25-1042	Postsecondary Biological Sciences Teachers

The first wage series is the weighted average salary for individuals holding bachelor's degrees in biological sciences or chemistry, who have not earned (and are not currently earning) a graduate degree. We constructed this series by combining data from the National Science Foundation's Survey of Doctoral Recipients, and the American Community Survey, as described below.

NSF Survey of Doctoral Recipients (SDR)

The Survey of Doctoral Recipients (SDR) is a longitudinal survey of individuals who have received a doctorate degree in science, engineering, or health fields from a U.S. institution. Collected by the National Science Foundation every two to three years, it follows a sample of individuals from the time they receive their PhD in a science, engineering, or health-related field, until they reach age 75.

First, we used publicly-available SDR data to determine which bachelor's degree fields are associated with earning PhDs in life sciences fields. For surveys conducted in 1999 through 2008, we find more than 80% of life sciences PhDs earned bachelor's degrees in biological sciences or chemistry. Health-related majors such as nursing were also represented among the PhDs, but those majors less frequently chose to pursue research PhDs. As such, we felt the biological sciences and chemistry majors' alternative careers would be most representative of the alternatives a prospective PhD student might consider.

American Community Survey Public Use Microdata Sample

We use the 2009 American Community Survey (ACS) Public Use Microdata Sample (PUMS) to identify occupations associated with bachelor's degrees in biological sciences and chemistry. We first calculated the survey-weighted share of all biological sciences and chemistry Bachelor's-degree holders in each occupation code, as well as the share of all S&E-degree holders in each occupation code. Then, we merged these calculated shares from ACS with BLS OES wage data by 4-digit Standard Occupational Classification system code (SOC), and used these shares to estimate a weighted average salary by year for bachelor's degree holders across these alternative occupations.

NSF-NIH Survey of Graduate Students and Postdoctorates in Science and Engineering

We use the NSF-NIH Survey of Graduate Students and Postdoctorates (GSS) to estimate the number of students entering U.S. biomedical sciences degree programs each year. The GSS is an annual survey of departments and other degree-granting units at U.S. academic institutions, granting degrees in S&E fields. The survey collects data on part- and full-time enrollment, student demographics, and students' sources of financial support. For this analysis, we use counts of first-time full-time students, total students enrolled, and students whose primary source of financial support is NIH funding, including NIH-funded research assistantships, traineeships, or fellowships across institutions offering PhDs in biological, medical, and other life sciences.

Figure 3 demonstrates how enrollments have risen throughout the NIH budget doubling years, flattening out as the NIH expansion came to an end, and reaching a sharp peak in 2008. Enrollments rose consistently throughout most of the 12 years in our data set, only beginning to decline as the NIH began funding more students. This may indicate that biomedical graduate students do indeed have cobweb expectation.

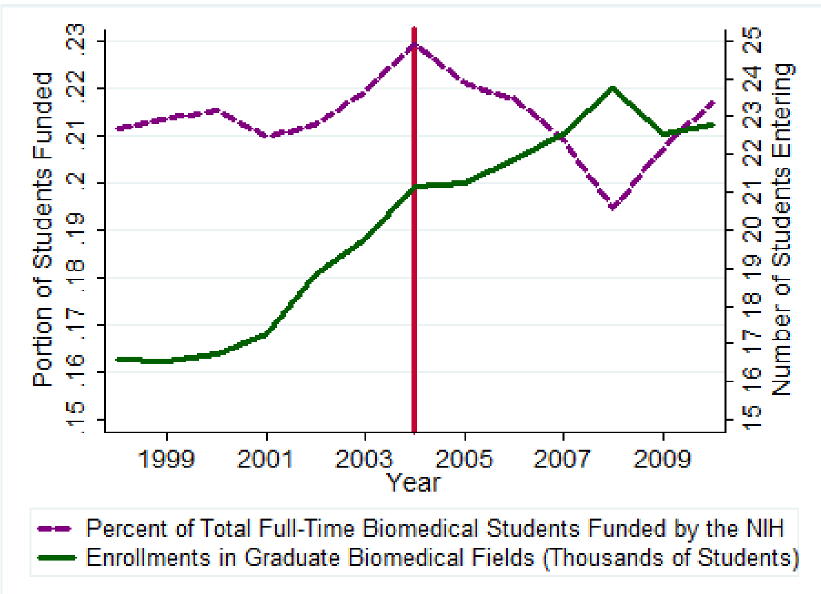

Figure 3: Enrollments and Availability of Funding

Integrated Postsecondary Education Data System (IPEDS) Completions

Counts of the number of students graduating each year from U.S. institutions with Bachelor's degrees in biological sciences and chemistry-related fields, as well as in all

S&E fields, were obtained from the National Center for Education Statistics' IPEDS Completions Survey. This data is collected annually in the Spring from all higher education institutions within the United States and Washington D.C. that participate in federal student financial aid programs.

Survey of Earned Doctorates

The Survey of Earned Doctorates (SED) has been used to collect statistics on the complete population of students graduating with PhDs each year, from 1957 to present. This survey includes information on race, gender, citizenship, as well as degree characteristics. We use the SED data to determine the number of completions in biomedical science fields each year. Figure 4 shows a similar interaction to Figure 3 except that completions are rising consistently as NIH funding becomes scarcer. This may demonstrate that as NIH funding dries up, students become more motivated to complete their degrees and graduate.

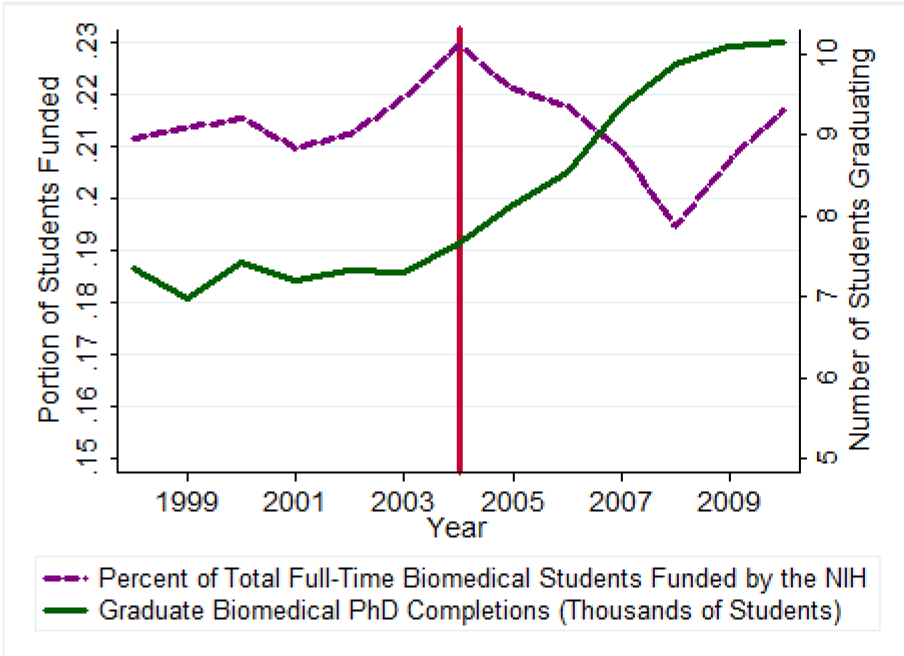

Figure 4: Graduates and Availability of Funding

Macroeconomic Data Series

In addition to the survey microdata sources noted above, we also use four macro-level data series to evaluate possible demand shifters. NIH R&D obligations by year were obtained from the NSF Survey of Federal Funds for R&D, adjusted to constant 2010 dollars using the Biomedical Research and Development Price Index (BRDPI).

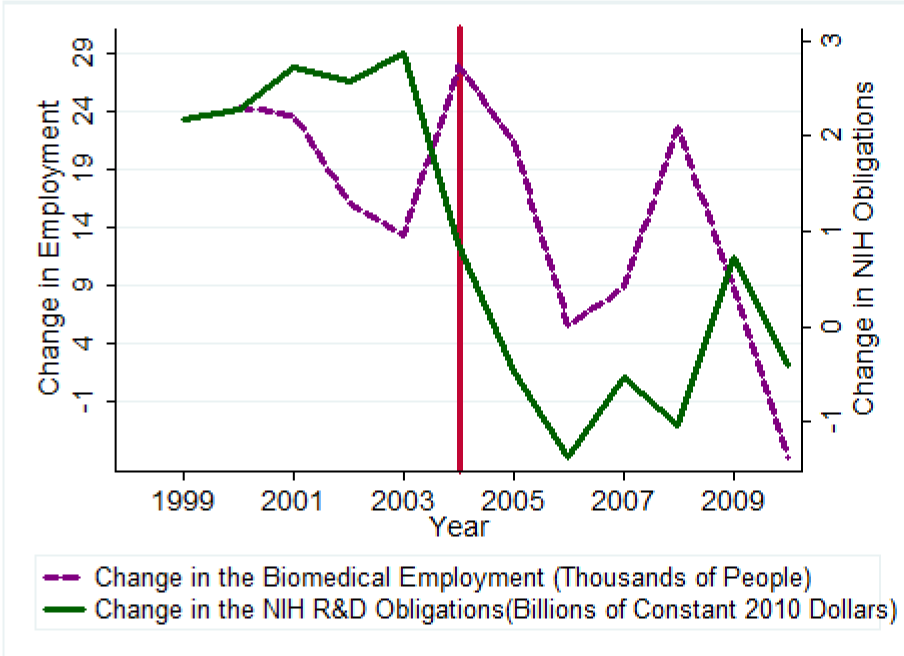

Figure 5: Employment and NIH Obligations

Real U.S. Gross Domestic Product (GDP) estimates in chained 2010 dollars are calculated using U.S. Bureau of Economic Analysis (BEA)[1] data. In addition, we construct two alternative estimates of annual biopharmaceutical industry expenditures. The first is Pharmaceutical Research and Manufacturers Association (PhRMA) members' reported expenditures on domestic R&D adjusted to constant 2010 dollars using the BRDPI. This series includes all R&D expenditures in the US by the pharmaceutical trad association's members, including both US and foreign-owned firms. The second uses BEA R&D satellite accounting data for pharmaceutical and medicine manufacturers (NAICS code 3254), available for years 1998 through 2007, and extrapolates that series

---

[1] Data available online at http://www.bea.gov/newsreleases/general/rd/2010/xls/1998_2007_rd_data_2010RDSA.xls

through 2010 based on annual changes in PhRMA-reported expenditures. In contrast with the PhRMA series, the BEA data excludes foreign-owned firms, but includes all US domestic industry-performed and industry-funded R&D.

Figure 6 shows the interaction between employment and pharmaceutical R&D obligations, showing that employment consistently trails the pharmaceutical industry by approximately one year.

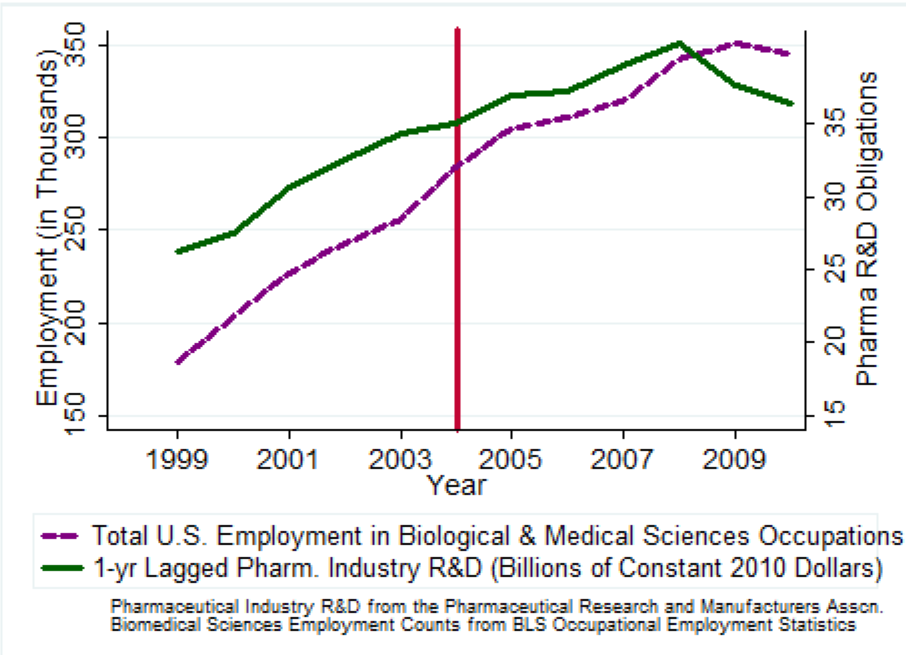

Figure 6: Employment and Pharma R&D Obligations

Table 2. Summary of Variables		
	Mean	**Standard Deviation**
Employment in Biomedical Fields	280568	58300.55
Average Annual Biomedical Workforce Salary	74667	4056.571
First-Time Full-Time Graduate Students Entering Biomedical Sciences PhD Programs	20121.2	2644.293
Full-Time Biomedical Sciene Graduate Students Supported by the NIH	17270.5	2342.308
Gross Domestic Product (in Billions of Dollars)	13404.2	1090.24
Total NIH R&D Obligations (in Millions of Dollars)	28638.7	4274.858
Total Life Sciences Broad Field U.S. PhD Completions	8262.77	6979
Total Bio/Med/Chem Bachelor's Degrees Conferred this Year in the U.S.	171639	30958.31
Average Annual Salary for BioChem Bachelor's Grads	53789.6	2359.66
PhRMA Member Companies R&D Expenditures (in Millions of Dollars)	35016.9	4620.01
Pharmaceutical R&D Investment	51217.9	15734.58
All Salary, Investment, and Gross Production data are represented in constant 2010 dollars.		

II.B. Models

As discussed by Freeman (1976) and others, markets for highly-skilled labor are subject to substantial "production lag," where labor supply is largely predetermined by entry into training programs several years prior. We use both cobweb and forward-looking analyses in order to determine whether students are reacting to present-day conditions or attempting to forecast the future job market. The cobweb-type models make the assumption that present-day, time *t* enrollment is determined by present-day, time *t* market conditions. Participants in this labor market are committing to participation in a labor market for which they have no more indication of wage rates than the present wage rate (Hoy, 2001). In contrast, forward-looking models assume that students in time t are attempting, with some success, to predict wages and employment at the time they will be entering the market, time $t+d$. Here, we begin by estimating the supply of new entrants into biomedical sciences graduate programs, as follows:

$$GradEnter_t = \alpha_1 BioSalary_{t+d} + \alpha_2 AltSalary_{t+d} + \delta GradEnter_{t-1} + \varepsilon_1 \quad (1)$$

Equation (1) asserts that the number of students entering biomedical sciences graduate programs in a given year should be determined largely by expected salaries for completed biomedical sciences PhDs d years hence, $BioSalary_{t+d}$, where d is the time delay between admission and completion to reduce effects of any exogenous trend in enrollment, we include the first lag of the dependent variable. In addition to considering their expected future wages if they go on and complete a biomedical sciences PhD, prospective students should also consider the opportunity cost of choosing to attend graduate school, instead of pursuing some alternative career path accessible to those who have earned only a Bachelor's degree. For simplicity, in equation (1) we ignore the opportunity cost associated with years spent in graduate school, so include only the expected salaries for those alternative careers d years hence, $AltSalary_{t+d}$. In addition, due to data limitations, we are not able to use average wages by experience. Where ideally wage expectations could be represented by expected lifetime earnings, we are forced to use annual wage data. Ryoo and Rosen (2004) assume that over a short time series the wage profile by age and experience will be stable. If that assumption can be borrowed then this simplification should not affect our results.

Finally, note that in our empirical analyses, all variables are log-transformed. By using logged variables in most of our equations, we demonstrate the percent change in the dependent variable as a result of a percent change in independent variables. This gives a constant elasticity formulation where independent variable coefficients will give the elasticity of the dependent variable under the assumption that all other variables are held constant.

If graduate students have rational expectations regarding future market conditions, then students' expectations at time *t* for $BioSalary_{t+d}$ and $AltSalary_{t+d}$ would, on average, equal the true values of each variable at time *t+d*. However, rational prospective students may also consider strategic labor supply responses by other prospective students and existing PhD scientists. To assess this possibility, in equation (2) we add the stock of biomedical scientists employed at time *t+d*:

$$GradEnter_t = \beta_1 BioSalary_{t+d} + \beta_2 AltSalary_{t+d} + \beta_3 BioEmp_{t+d} + \delta GradEnter_{t-1} + \varepsilon_2 \tag{2}$$

If students have "cobweb" expectations, then their expectation for $BioSalary_{t+d}$ would simply be the current salary for biomedical scientists, $BioSalary_t$, and likewise their expectation for $AltSalary_{t+d}$ would be $AltSalary_t$. Likewise, their expectations regarding future job growth may depend on current changes in employment levels for biomedical scientists, *BioEmp*. We test these alternative assumptions about students' expectations empirically.

One difficulty with the formulations in equations (1) and (2) above, as discussed by Ryoo & Rosen (2004), is that they do not control for exogenous year-to-year changes in cohort size that could affect the number of students completing college, and thus eligible to enter PhD programs (i.e., supply shifters). We therefore also consider the following *relative* supply model:

$$\frac{GradEnter_t}{BachDeg_t} = \gamma_1 \frac{BioSalary_{t+d}}{AltSalary_{t+d}} + \gamma_2 \frac{BioEmp_{t+d}}{TotEmp_{t+d}} + \varepsilon_3 \qquad (3)$$

In this specification, the dependent variable represents the share of students who graduated with bachelor's degrees in biology or chemistry in year *t* and went on to enter graduate programs in biomedical science fields in the following academic year. The explanatory variable is the *relative* financial prospects at graduation, in year *t+d*, for a student who completes a PhD in biomedical sciences, versus the wages paid for those majors' alternative career paths.

If we presume that all workforce outcomes for biomedical sciences PhDs are considered by prospective students—including non-science-related and non-research jobs—then the correct measure for students' expected income at graduation is the average entry-level wage for a biomedical sciences PhD, regardless of his or her occupation, as we calculate from the SDR data.

On the other hand, if new PhDs are taking non-research and non-science-related jobs due to an excess labor supply that depresses entry-level wages in biomedical sciences research occupations, then only the wages in those specific positions are relevant. This latter specification is somewhat more attractive for our market model, as these wages are the relevant prices for the demand side of the market. However, in that case, the rational expectations labor supply model should also include a measure of the *stock* of biomedical sciences researchers as well. We explore each of these specifications empirically, below.

Again, ignoring for simplicity any additional signals a student may receive during his or her *d* years of graduate training that might affect expectations of future salaries, the number of biomedical sciences PhDs supplied to the market at time t, $Grad_t$, is then determined mainly by $GradEnter_{t-d}$:

$$Grad_t = \phi_1 GradEnter_{t-d} + \sum_{\lambda=0}^{d} \nu_\lambda NIHFunded_{t-\lambda} + \varepsilon_4 \quad (4)$$

In equation (4) above, we expect the term ϕ_1 will be less than one, reflecting usual attrition from doctoral programs. We also add additional terms to assess whether increases in the proportion of students funded by NIH research assistantships, traineeships, or fellowships impacts timely PhD completions. The outcome variable $Grad_t$ shows the influx of new PhDs into the labor supply that year. Like equation (3), this equation can also be expressed in relative terms, to avoid supply-side shift effects like changes in cohort size.

We can interpret ν_λ as the impact of changes in the availability of NIH support for students in each year of their PhD program on their probability of completing the PhD within *d* years. A significant negative coefficient would imply that an increase in NIH support for students during that year of a given cohort's graduate training actually decreases the six year completion rate for the cohort. When first-differencing the variables as in the empirical analyses below, the coefficient represents the effect of acceleration in NIH spending in any given year on the rate of PhD completion.

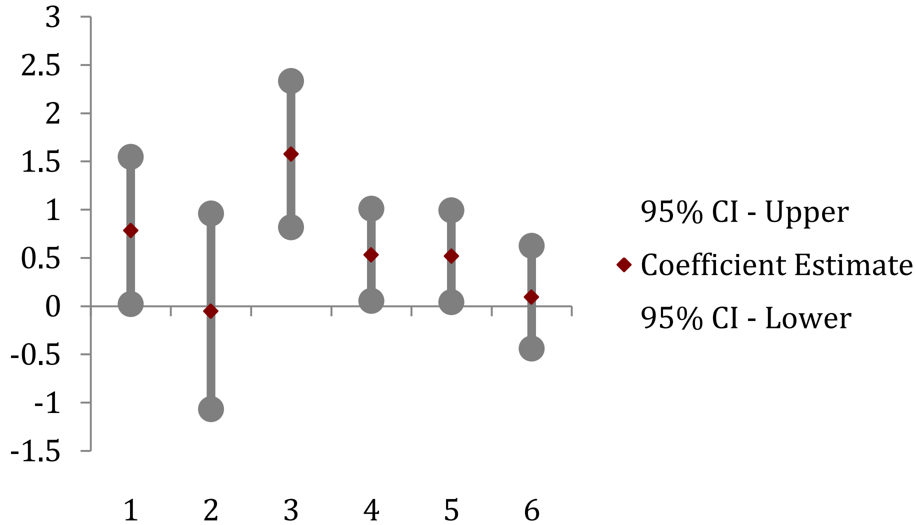

Figure 7: Impacts of changes NIH graduate student support on 6-year PhD completions in years 1 through 6 - 95% Confidence Interval

Adding relative wages both at entry and at graduation provides an alternative estimate of the effects of wages on the number of students who ultimately are eligible to enter the workforce. To preserve degrees of freedom given the relatively short time series we have available, in this model we consider only the percentage of biomedical sciences graduate students funded by NIH six years prior to graduation, as a supply-side attractor:

$$\frac{Grad_t}{BachDeg_{t-d}} = \delta_1 \frac{BioSalary_{t-d}}{AltSalary_{t-d}} + \delta_2 \frac{BioSalary_t}{AltSalary_t} + \delta_3 \frac{NIH\ FundedStudents_{t-d}}{TotalStudents_{t-d}} + \varepsilon_5 \qquad (5)$$

In the empirical analysis, we also investigate variations on equation (5) that are exclusively cobweb (i.e., include only salary and number of jobs at time *t-d*) or exclusively forward-looking (i.e., include only salary and number of jobs at time *t*).

Modeling Demand for Biomedical Scientists

In recent years approximately 70 percent of new PhD biomedical scientists have taken postdoctoral research/training positions (Stephan, 2012). Many of these postdoctoral positions are funded by NIH extramural research and training grants, but some are in industry (e.g., at biopharmaceutical firms) and in government. We therefore represent the (inverse) demand function for biomedical scientists as follows:

$$\frac{BioSalary_t}{AltSalary_t} = \theta_1 \frac{BioEmp_t}{TotEmp_t} + \theta_2 \frac{NIHRnD_t}{GDP_t} + \theta_3 \frac{PharmaRnD_t}{GDP_t} + \varepsilon_6 \qquad (6)$$

In this representation, the dependent variable is the log relative wage for biomedical scientists, and $BioEmp_t$ is the total number of biomedical scientists employed at time *t*. The demand-shifters are total NIH obligations for R&D in year *t*, $NIHRnD_t$, representing demand for postdoctoral workers in academia and government, and estimated annual pharmaceutical and biotechnology industry R&D expenditures, $PharmaRnD_t$, to represent demand in industry. As usual, we expect the sign on θ_1 will be negative, reflecting that an increase in labor market supply will, all else equal, reduce PhDs' market wages.

Finally, we can also estimate the relative demand function directly, with quantity demanded as the dependent variable:

$$\frac{BioEmp_t}{TotEmp_t} = \rho_1 \frac{BioSalary_t}{AltSalary_t} + \rho_2 \frac{NIHRnD_t}{GDP_t} + \rho_3 \frac{PharmaRnD_t}{GDP_t} + \varepsilon_7 \tag{7}$$

II.C. Econometric Estimation

If an exogenous shock to wages or employment in one period affects unobserved factors in later periods, there may exist autocorrelation in the error terms. While there are many ways of testing for autocorrelation, here we follow Ryoo and Rosen (2004) in using the Durbin-Watson tests.

The Durbin-Watson test, or *d* statistic, uses the OLS residuals to detect first-order autocorrelation. The d statistic is calculated as:

$$d = \frac{\sum_{t=2}^{T} (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{T} \hat{u}_t^{\,2}} \approx 2(1 - \hat{\rho})$$

The term $\hat{\rho}$ is the sample correlation coefficient, estimated from the residuals:

$$\hat{\rho} = \frac{\sum (\hat{u}_t \hat{u}_{t-1})}{\sum \hat{u}_t^2}$$

Because the correlation coefficient is bounded between -1 and 1, where -1 is perfect negative autocorrelation and 1 is perfect positive correlation, the *d* statistic will be approximately 2 when there is no autocorrelation, will approach zero when autocorrelation is strongly positive, and will approach 4 when autocorrelation is strongly negative. Durbin and Watson (1951) derived a table of critical values, that is, lower and upper bounds for an acceptable *d* statistic, given various combinations of number of explanatory variables and sample size. The closer the *d* statistic is to 2, the greater the probability that we fail to reject the null hypothesis of no autocorrelation.

The original Durbin-Watson test assumes strict exogeneity of the regressors, which precludes its use in autoregressive models like that in equation (1). In models such as that, where lagged values of the dependent variable are included among the explanatory variables, the d statistic will be biased towards 2 and thus we may fail to reject the null of no autocorrelation, even when autocorrelation is present. Durbin (1970) presented a more general alternative test that permits lagged dependent variables, and also assesses presence of higher-order autocorrelation.

Instrumental Variables (IV) Estimation

Structural market models, such as the supply and demand equations we estimate here, are characterized by jointly (simultaneously) determined prices and quantities. If estimated independently, without taking into account the information provided by other equations in the system, simultaneous equations will yield inconsistent results. Specifically, if the error term in one equation is correlated with an explanatory variable in the other equation, simultaneous equations bias will occur.

Empirically, one approach to estimating simultaneous equations is two-stage least squares (2SLS) IV. To implement 2SLS IV estimation for the labor supply equation, we need to find one or more instruments that are highly correlated with biomedical scientists' wages, but are otherwise uncorrelated with unobserved factors affecting the number of students enrolling in (or completing) PhD programs in biomedical sciences. Ryoo and Rosen (2004) employ the third and fourth lags of public and private defense R&D spending relative to total GDP, which they argue reflect changes in demand that only affect supply of bachelor's-degree engineers through their prospective wages. For us, NIH R&D as a share of total GDP would be analogous, but empirically we find that public (NIH) R&D funding is in fact a strong predictor of supply as well as demand, and therefore does not assist in resolving the identification problem. This relationship between our nominal "demand shifter" NIH funding and graduate student enrollments can be seen very clearly in Figure 8, below.

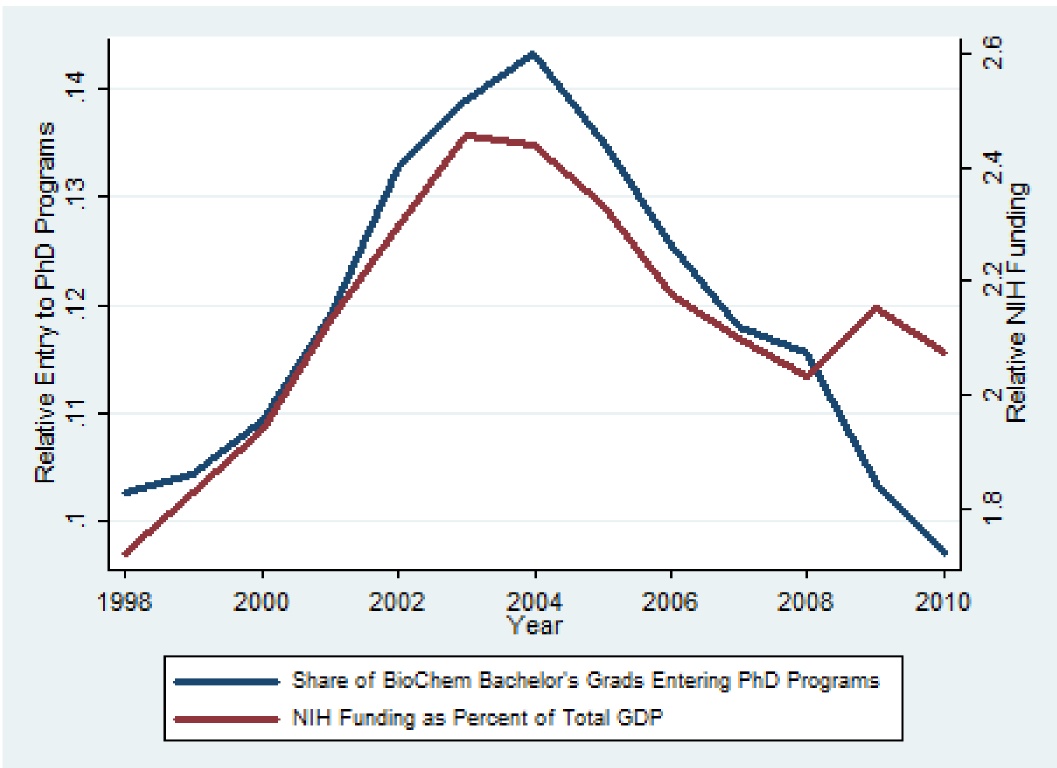

Figure 8: Relative Enrollment and Relative NIH Funding

In the analyses that follow, we therefore instrument for wages using only the third and fourth lags of pharmaceutical industry R&D expenditures, relative to GDP. The relevance of industry R&D expenditures to market wages is tested via the partial F-statistic for the excluded instruments in the first stage regression. To maximize the number of usable observations in our time series given restrictions on available data, we employ the PhRMA estimated R&D expenditures as instruments, and use the BEA pharmaceutical R&D estimates as the explanatory variable for our demand equations.

Overid Testing

In addition to the relevance condition—that industry R&D expenditures must be correlated with the endogenous regressor, market wage—industry R&D expenditures must also satisfy the exogeneity condition. That is, industry R&D expenditures can only be correlated with graduate student enrollment and completions via the price mechanism (wages). When the system is overidentified, meaning the number of exogenous instrumental variables exceeds the number of problematic endogenous variables, we can use the Sargan or Hansen tests of overidentifying restrictions ("overid tests") to assess whether evidence supports exogeneity of the instruments.

Overid tests essentially construct alternative models using subsets of the instruments so that the system is "just-identified," meaning there are only as many instruments as there are regressors, and then compare the residuals. If the overid test statistic exceeds its critical value, this can be interpreted either as evidence that one or more of the instruments is correlated with the error term (so is not exogenous after all), or possibly that the omitted instruments actually belong in the second-stage regression and our equation is misspecified (see Davidson and McKinnon (2004), p. 336-338).

Using Hansen's J-statistic, we are testing the null hypothesis that all of the instrumenting variables are valid. If the critical value associated with this test is large relative to the chi-squared random variable, and significant at the .05 or .10 significance level, then we reject the null hypothesis.

III. RESULTS

In this section we summarize the results of econometric regressions used in this paper. We have used time-series data to create OLS and IV models which provide some insight into the responsiveness of graduate students to macroeconomic inputs. The conclusions resulting from these estimations, combined with descriptive statistics, help to map the realities of the biomedical workforce.

Models of the Supply of Biomedical Scientists

We begin with basic OLS regression to see if graduate students in biomedical fields express rational expectations in the labor market. Models (1) and (2) in Table 3 are dynamic autoregressive models, with the log of the number of students entering PhD programs as the dependent variable. All equations are in a log-log format, giving us the elasticity of any given variable.

The dynamic ordinary least squares model (1) has for an outcome variable the log of entering PhD students. Predictor variables are the six-year leads of wages and employment. This regression finds no significant effects on enrollment for either of the labor market variables. Only the coefficient on the lagged AR(1) dependent variable is significant, with a coefficient exceeding one, indicating an unstable, or non-stationary, process.

Lacking useful results from the forward-looking dynamic model, we test cobweb expectations in model (2). In this regression logged wages and employment are measured in the present time period *t*. PhD enrollment is found to be significantly correlated with employment. A 1% increase in biomedical jobs will yield an approximately .45% increase in graduate enrollment ($p<.001$).

For model (3) we add the use of relative variables to our OLS cobweb model in order to control for any cohort effects. It is possible that population trends may be increasing PhD enrollments, instead of labor market trends. For this reason we divide the dependent variable by the number of students graduating from the university with a degree in biology or chemistry. In addition, we relativize wages by dividing by an opportunity-cost salary and employment by dividing by employment in non-PhD S&E careers fields.

Model (3) yields no significant results and is revealed by the Durbin-Watson test to have high autocorrelation. We attempt to mitigate autocorrelation in model (4) by returning to a dynamic cobweb model and first-differencing all variables. The autoregressive variable in this model is the only significant variable, revealing that a 1% increase in logged relative enrollment in the previous period will increase logged relative enrollment in the present period by .83%.

Table 3. Rational versus Cobweb Expectations in PhD Enrollments

	(1)	(2)	(3)	(4)
	OLS Dynamic Forward	OLS Dynamic Cobweb	OLS Relative Cobweb	OLS Relative Dynamic First-Differenced Cobweb
Wages	-0.148	-0.973	-1.942	1.178
	(0.903)	(0.686)	(2.884)	(0.749)
Employment	0.0674	0.445***	0.603	-0.0883
	(0.0661)	(0.129)	(0.200)	(0.308)
Graduate First-Time Enrollment, t-1	1.197***	0.164		0.829***
	(0.156)	(0.258)		(0.202)
Observations	12	12	12	11
Durbin-Watson	2.088	2.118	0.373	2.299
Durbin's alternate	0.0405	0.0551	29.59	1.157
Durbin's alternate p-value	0.846	0.821	0.000616	0.318

*** p<0.01, ** p<0.05, * p<0.1
Outcome variable for models (1) and (2) is the log of the number of students entering PhD programs in biomedical sciences. Outcome variable for models (3) and (4) is the log of the ratio of the number of students entering PhD programs to the number of Bachelor's degrees earned in biological sciences and chemistry.
Wages and employment variables for model (1) are six-year leads, reflecting expected market conditions at time of graduation with perfect foresight (rational expectations). Models (2)-(4) use current wages and employment levels (cobweb expectations). Models (3) and (4) use the log of the ratio of wages for biomedical sciences vs. alternative career fields, and the log of the ratio of employment in biomedical sciences occupations vs. alternative career fields.
In model (4), all variables are first-differenced to mitigate autocorrelation.
All equations are estimated using OLS, and fully robust standard errors are reported in parentheses below each coefficient estimate.

Autocorrelation seems to be resolved by first-differencing all of the variables. However, the explanatory variables of wage and employment remain insignificant. Least there be simultaneity from the two independent variables, we run a first-differenced cobweb model, model (1) in table 4, with only log relative wage as the independent variable. This regression provides us with a unit-elastic result for wage ($p<.1$).

Using the same variables as model (1), we switch to IV regression for model (2). This should help address any endogeneity that may have been an issue in the wage variable. IV regression reveals a wage elasticity of 3, ($p<0.1$). When we add lagged variables of wage and employment in model (3), now calculating PhD completions instead of entrants, we find a similar elasticity on wage which is also significant. This indicates that labor supply is highly responsive to changes in wage.

Model (4) exchanges the present-day wage calculation for a lagged NIH funding variable – the effect of logged, relative, first-differenced NIH funding of students in time period t-6 on present-day logged, relative, first-differenced completions. Employment at time t-6 is very significant ($p<.001$), indicating that a 1% change in the first-difference of employment results in an .87% change in the first-difference of completions. The lagged wage variable is not significant at all – indicating that students may care more about their prospect of getting a job than they do about the wage they will earn. The NIH funding variable is also very significant ($p<.001$), and it is negative. A 1% increase in the change in NIH funding at the time of enrollment actually has a 1.4% negative effect on the change in completions. This may be because an increased number of students being funded by the NIH could have a negative effect on the quality of student.

Table 4. Cobweb vs Rational Expectations on PhD Enrollment and Completions

	(1)	(2)	(3)	(4)
	OLS Cobweb First-Differenced Enrollments	IV Cobweb First-Differenced Enrollments	IV Mixed First-Differenced Completions	IV Cobweb First-Differenced Completions
Relative Wage, t	0.974*	3.167*	3.596**	
	(0.434)	(1.827)	(1.812)	
Relative Employment, t				
Relative Wage, t-6			0.346	1.006
			(1.421)	(0.778)
Relative Employment, t-6			0.409	0.874***
			(0.283)	(0.0782)
Relative Enrollment, t-1	0.793***	0.601**		
	(0.156)	(0.261)		
Percent of Students with NIH Funding, t-6				-1.492***
				(0.351)
Observations	11	11	5	5
Durbin-Watson	2.163			
Durbin's alternate	0.214			
Durbin's alternate p-value	0.656			
First-Stage F-stat	OLS	0.222	0.417	4.193
Partial R-squared			0.5973	0.4804
Hansen's J-statistic		3.030	0.0541	0.316
Hansen's J p-value		0.220	0.816	0.574

Outcome variables for models (1) and (2) are first-differenced logs of the ratios of first-time graduate student enrollment in biomedical sciences PhD programs, over total Bachelor's degrees in biological sciences and chemistry that year. Outcome variables for models (3) and (4) are logs of the ratio of PhDs completed over the sixth lag of Bachelor's degrees.

Table 5, model (1) once again raises the issue of simultaneity. The dynamic OLS regression produces strong significance on each of the four variables included, the negative coefficient on present-day log relative wage has returned. This implies that an increase in the relative wage for biomedical scientists of 1% decreases PhD completions by 2%, ($p<.001$). Two other variables in the model are more predictable, with employment and NIH Funding both being highly significant. Students will complete their PhD more quickly with the possibility of a job and they will complete it in a less timely manner if they enrolled at a time of NIH funding expansion.

Model (2) returns to IV regression, using the exact same variables from model (1), but the first-stage partial F-statistic is unimpressive. The instrumented variable in this model is the lagged relative PhD completions variable – indicating that instruments do a better job of predicting quantity than price.

IV models (3) and (4) provide very strong results. This is most likely a result of the reduction in the number of observations, in addition to the time period over which they span. NIH obligations and industry R&D are the instruments for employment in this model. The observations in these two models cover the years 2006 through 2010 – a period which follows the NIH budget expansion years when Congress was no longer emphasizing biomedical research. Whereas the NIH was a significant driver the labor market in the past, when budget growth stopped other variables were given the opportunity to become influential – including industry R&D. The partial F-statistic for models (3) and (4) are very good compared to previous estimations.

Both models indicate that scientist supply is very responsive to changes in wages. Model (3) controls for the availability of NIH student support at matriculation, producing a significant coefficient nearly identical to that from table 4, model (4). The story remains that students who enroll in the midst of NIH support growth are less likely to graduate in a timely manner. Model (4) finds something unique to previous NIH support coefficients in this paper. Students who are in the middle of their biomedical studies when the NIH begins an expansion are more likely to graduate. For a 1% increase in the change in the proportion of NIH students funded in year 3 of their PhD program, there will be a 1.2% increase in the six-year completion rate ($p<.001$).

Table 5. Effects of NIH Funding for Graduate Students on PhD Completion Rates

	(1)	(2)	(3)	(4)
	OLS Forward Dynamic	IV Forward Dynamic	IV Mixed First-Differenced	IV Mixed First-Differenced
Log Wage, t	-2.063**	1.494	2.912***	4.157***
	(0.758)	(1.631)	(0.900)	(0.350)
Log Employment, t	0.463***	1.183***		
	(0.0695)	(0.190)		
Log Percent of Students with NIH Funding, t-6	-2.423***	-3.715***	-1.453***	
	(0.494)	(0.476)	(0.262)	
Log Percent of Students with NIH Funding, t-3				1.235***
				(0.0857)
Log Employment, t-6			0.721***	0.463***
			(0.112)	(0.0381)
Log Relative Completions, t-1	0.793***	0.682***		
	(0.156)	(0.144)		
Observations	12	10	5	5
Durbin-Watson	2.065			
Durbin's alternate	0.0578			
Durbin's alternate p-value	0.817			
First-Stage F-stat	OLS	1.013	3.353	7.855
Partial R-squared		0.9558	.9821	.9192
Hansen's J-statistic		0.0486	0.0169	1.057
Hansen's J p-value		0.826	0.897	0.304

Outcome variables for models (1) and (2) are the log of absolute PhD completions from biomedical PhD programs in a given year. Outcome variables for models (3) and (4) are the same, with the addition that these variables have been first-differenced. The log of wages is differenced in equations (3) & (4), as is the Logged Percent of Students with NIH funding in time period t-6.

Models of the Demand for Biomedical Scientists

Table (6) model (1) presents results from an inverse (log-log) demand function where the first-difference of wages is determined by first-differenced values for employment, NIH R&D, and the lagged value of Industry R&D. We are able to use the third and fourth lags of pharmaceutical R&D investment as instruments for employment because the observations span a time period after the NIH budget doubling had come to an end. The overid test fails to reject the null hypothesis, allowing that the variables may be exogenous. By this specification we find that wages are negatively affected by employment rates so that a 1% increase in employment yields a .76% decrease in wages at time t ($p<.001$). This corroborates economic theory.

We also calculate a traditional demand function. Model (2) relates the elasticity of biomedical science labor demand with respect to wages. Every coefficient is highly significant with an elasticity of -1.2% for the wage variable. This indicates that a 1% increase in wages results in a 1.2 percent decrease in biomedical labor demand ($p<.001$).

While both models had ubiquitously significant coefficients and a strong partial first-stage F-stat, it is worthwhile to note that the coefficients on *NIH R&D Obligations* are larger than the coefficients for *Industry R&D* in both the invers demand function and the demand function. The NIH variable is significantly more influential in both models, with a value of (.7, $p<.001$) in the inverse function and (.9, $p<.001$) in the regular demand function. Industry R&D only had an elasticity of demand of less than .3 in both models ($p<.001$).

Table 6. Demand Functions for Biomedical Sciences Labor		
	(1)	(2)
	Inverse Demand	Demand Function
Employment, t	-0.768***	
	(0.0273)	
Wage, t		-1.299***
		(0.0402)
NIH R&D Obligations, t	0.695***	0.906***
	(0.0578)	(0.0553)
Industry R&D, t-1	0.226***	0.295***
	(0.0180)	(0.0191)
Constant	0.0369***	0.0481***
	(0.00207)	(0.00110)
Observations	7	7
First-Stage F-stat	80.17	597.2
Partial R-squared	0.9836	0.9845
Hansen's J-stat	1.546	1.471
Hansen's J p-value	0.214	0.225
Instruments	L4D.lrgind L3D.lrgind	L4D.lrgind L3D.lrgind
The outcome variable for model (1) is the logged first-difference of wages within biomedical sciences. The outcome variable in model (2) is the log of first-differenced biomedical employment. All independent variables in equations (1) and (2) are first-differenced and logged. NIH R&D Obligations at time t and Industry R&D at time t-1 are both relative variables – each divided by GDP.		

IV. CONCLUSIONS

IV.A. Policy Implications

This research finds that graduate biomedical students take wages and employment levels into consideration when they are deciding whether or not to pursue a PhD. We find that the most significant influence on matriculation is the level of National Institutes of Health funding available. A matriculated student is not, however, a future graduate. NIH funding ironically lures life science students into graduate school while simultaneously decreasing their chances of success. It may be that the allure of extra funding opportunities is too strong – drawing insufficiently prepared students into biomedical science.

In the first-differenced cobweb OLS model, we found positive unit-elastic wage significance upon enrollment. A one percent increase in wages would result in a one percent increase in enrollments. Using IV cobweb, first-differenced models, we found some significance in relative wages. We find robust evidence that elasticity of labor supply (as measured by PhD enrollment and completions) with respect to relative wages, whether at time of entry to the PhD program or at time of completion, has an elasticity of about 3. This implies that labor supply is highly responsive: a 1% increase in relative wages would result in a 3% increase in relative employment. It is using IV cobweb first-differencing that we first see negative effect of NIH funding on completions. An inverse demand function reveals a very significant, negative effect of employment on salary. In a regular demand function, NIH R&D obligations are revealed to be very significant and unit elastic.

The findings around NIH R&D funding are the most significant results of this research. This contributes to previous literature in the field indicating that discretion in the use of NIH funding for student aid is vital. Figure 8 indicates that students are highly responsive to the availability of funding when deciding whether to enroll. Figure 7 shows that the early availability of NIH funds in a student's graduate career may delay or altogether prevent graduation. The NIH is right to have begun the research into graduate careers that began with the NIH Advisory Committee to the Director (ACD) Biomedical Research Workforce Working Group. Our results would indicate that the NIH can get the most from their human capital investments by restricting the years of graduate school in which NIH funding is available. If this funding were to be given to students beginning only in year 3 of their graduate school career, it might reduce excess enrollment of the unqualified while increasing completions. The NIH may, however, consider other strategies. If they wish to increase the overall quality of graduating students, the NIH may choose to continue funding students early in their careers – luring a greater number of applicants - under the assumption that only the brightest will complete, increasing the graduate pool from which scientists are made.

IV.B. Limitations

The most significant results from this paper have come from simple correlations. Figure 8 shows a very significant relationship between NIH funding and enrollment. The econometric regressions run in this paper are hampered by the restricted availability of data - caused by the career code reclassifications in Bureau of Labor Statistics data which occurred in 1998. Our results, however, are in line with those produced in similar research. It is possible that these methods could be applied with an expanded dataset in the future, yielding similarly significant results around the importance of wages and employment rates.

Ryoo and Rosen (2004) find a similar wage elasticity, between 2.5 and 4.5, in their research on engineering graduate students. Freeman (1974), Ryoo & Rosen (2004), and the research presented in this paper all found a greater relevance in cobweb expectation models. This would indicate that students may be less forward-looking in their decision making than they are responsive to present-day conditions.

IV.C. Summary and Directions for Future Research

This paper was unable to conduct an analysis of demographic responses to macroeconomic matriculation incentives. The data would indicate that race and gender have some effect on the types of occupations that students will ultimately work in. Future research should seek to understand the effect of race and gender on responsiveness to macroeconomic variables. It would also be ideal for these same analyses to be conducted with alternate wage data. The Survey of Doctoral Recipients could provide an alternate source of wage data for scientists working in biomedicine. Finally, a larger time series is needed in order to lend greater significance to these wage and employment results. In 15 or 20 years these analyses might be repeated, creating greater insight into the significance of wages and employment upon matriculation. In the meantime, the National Institutes of Health and other graduate student funding sources would do well to carefully examine the mechanisms and timelines by which they will assist potential and current students.

REFERENCES

1. **Ampaw, Frimpomaa D. .** "The Effect of Labor Market Conditions and Financial Aid on Doctoral Student Retention," North Carolina State University, 2010, 143.

2. **Augustine, N. R.** "Rising above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future." *Retrieved March*, 2005, *19*, pp. 2008.

3. **Bedard, Kelly and Herman, Douglas A.** "Who Goes to Graduate/Professional School? The Importance of Economic Fluctuations, Undergraduate Field, and Ability." *Economics of Education Review*, 2008, *27*(2), pp. 197-210.

4. **Betts, J. R. and McFarland, L. L.** "Safe Port in a Storm: The Impact of Labor Market Conditions on Community College Enrollments." *Journal of Human Resources*, 1995, pp. 741-65.

5. **Blume-Kohout, Margaret E. and Adhikari, Dadhi.** "Training the Biomedical Workforce: Does Funding Mechanism Matter?," *Southern Economic Association 82nd Annual Meetings, .* New Orleans, LA, 2012.

6. **Committee on Dimensions, Causes; Implications of Recent Trends in Careers of Life, Scientists and National Research, Council.** *Trends in the Early Careers of Life Scientists*. The National Academies Press, 1998.

7. **Dellas, Harris and Koubi, Vally.** "Business Cycles and Schooling." *European Journal of Political Economy*, 2003, *19*(4), pp. 843-59.

8. **Dellas, H. and Sakellaris, P.** "On the Cyclicality of Schooling: Theory and Evidence." *Oxford Economic Papers*, 2003, *55*(1), pp. 148-72.

9. **Durbin, J.** "Testing for Serial Correlation in Least-Squares Regression When Some of the Regressors Are Lagged Dependent Variables." *Econometrica: Journal of the Econometric Society*, 1970, pp. 410-21.

10. **Durbin, J. and Watson, G. S.** "Testing for Serial Correlation in Least Squares Regression. Ii." *Biometrika*, 1951, pp. 159-77.

11. **Ehrenberg, Ronald G.** "The Flow of New Doctorates." *Journal of Economic Literature*, 1992, *30*(2), pp. 830-75.

12. **Ehrenberg, R. G. and Mavros, P. G.** "Do Doctoral Students' Financial Support Patterns Affect Their Times-to-Degree and Completion Probabilities," National Bureau of Economic Research, 1995.

13. **Ehrenberg, Ronald G.; Rees, Daniel I. and Brewer, Dominic J.** "Institutional Responses to Increased External Support for Graduate Students." *The Review of Economics and Statistics*, 1993, *75*(4), pp. 671-82.

14. **Freeman, R. B.** "A Cobweb Model of the Supply and Starting Salary of New Engineers." *Industrial and Labor Relations Review*, 1976, pp. 236-48.

15. **Freeman, R. B.** "Employment Opportunities in the Doctorate Manpower Market." *Industrial and Labor Relations Review*, 1980, *33*(2), pp. 185-97.

16. **Freeman, Richard B.** "Fellowship Stipend Support and the Supply of Science and Engineering Students: Nsf Graduate Research Fellowships." *American Economic Review*, 2005, *95*(2), pp. 61-65.

17. **Garrison, H. H.; Stith, A. L. and Gerbi, S. A.** "Foreign Postdocs: The Changing Face of Biomedical Science in the Us." *The FASEB journal*, 2005, *19*(14), pp. 1938-42.

18. **Hamburger, Michael J.** "Interest Rates and the Demand for Consumer Durable Goods." *The American Economic Review*, 1967, *57*(5), pp. 1131-53.

19. **Hoy, M.** *Mathematics for Economics: International Student Edition*. Mit Press, 2001.

20. **Kennedy, P. E.** *A Guide to Econometrics, 5th Edition*. MIT Press, 2003.

21. **Levitt, David.** "Careers of an Elite Cohort of U.S. Basic Life Science Postdoctoral Fellows and the Influence of Their Mentor's Citation Record." *BMC Medical Education*, *10*(1), pp. 80.

22. **National Research Council, NRC.** *Research-Doctorate Programs in the Biomedical Sciences: Selected Findings from the Nrc Assessment.* National Academies Press, 2011.

23. **NIH.** "Biomedical Research Workforce Working Group Report," National Institutes of Health, 2012, 156.

24. **NSF.** *WebCASPAR Ingegrated Science and Engineering Recources*
25. National Science Foundation, 2008.

26. **Ryoo, J. and Rosen, S.** "The Engineering Labor Market." *Journal of political economy*, 2004, *112*(S1), pp. S110-S40.

27. **Smith, Courtney A.** "Millennial Students Who Go Directly to Graduate School: Influences on This Decision and the Characterization of Their Experience," Wright State University.

28. **Stephan, Paula.** *How Economics Shapes Science*. Harvard University Press, 2012.

29. **Stephan, Paula.** "Perverse Incentives." *Nature*, 2012, *484*(7392), pp. 29-31.

30. **Tilghman, Shirley; Rockey, Sally; Degen, Sandra; Forese, Laura; Ginther, Donna; Gutierrez-Hartmann, Arthur; Hrabowski, Freeman; Jackson, James; Joshua-Tor, Leemor; Lifton, Richard, et al.** "Biomedical Research Workforce Working Group Report," Bethesda, MD: National Institutes of Health, 2012.

31. **Wooldridge, J. M.** *Introductory Econometrics: A Modern Approach*. South Western, Cengage Learning, 2009.

APPENDIX TABLE A1. American Community Survey Undergraduate Field-of-Degree Codes for Biological Sciences and Chemistry	
Code	Field of Degree
Biological Sciences Fields	
3600	Biology
3601	Biochemical Sciences
3602	Botany
3603	Molecular Biology
3604	Ecology
3605	Genetics
3606	Microbiology
3607	Pharmacology
3608	Physiology
3609	Zoology
3611	Neuroscience
3699	Miscellaneous Biology and Epidemiology
2402	Biological Engineering
2404	Biomedical Engineering
4002	Nutritional Sciences
5102	Applied Biotechnology
Chemistry Fields	
5003	Chemistry
2405	Chemical Engineering
Health Related Fields	
6100	General Medical and Health Services
6102	Communication Disorders Sciences and Services
6103	Health and Medical Administrative Services
6104	Medical Assisting Services
6105	Medical Technologies Technicians
6106	Health and Medical Preparatory Programs
6107	Nursing
6108	Pharmacy
6109	Treatment Therapy Professions
6110	Community and Public Health