

*Diversity and Knowledge Creation: An Empirical Investigation from U.S. Schools*

Sreenath Majumder, Ph.D.

Stacie Koch

**Abstract**

Existing literature documents the effect of ethnic, racial, and linguistic fractionalization on growth, government quality, and on various other political outcomes. In general this line of research documents a negative impact of diversity on various economic and political outcomes. In this paper, we investigate the impact of ethnic diversity on knowledge creation. In particular we look at the student ethnic and racial diversity in American graduate schools and its impact on faculty publications. We are using National Science Foundation and National Research Council data sets across 6 years. Our Ordinary Least Squares (OLS) regression results indicate that ethnic/racial diversity does have a positive impact on faculty publications in the natural sciences, where as for social science fields, the results shows no statistically significant impact.

## **Introduction**

The underlying question of the research presented here is “do interactions with people from different backgrounds generate new ideas?” This study attempts to answer this question by focusing on whether or not an individual program’s ethnic/racial composition of students has any impact on scholarly activity. Scholarly activity is defined here as average number of faculty publications. This research looks at the racial composition in 2002-2004 and faculty publications for 2006. This lag accounts for the time to actually write the paper and then having it submitted for publications. The assumption is that students in 2002-2004 will be assisting the faculty in their research and papers that are published in 2006. For the racial/ethnic composition index, a higher number means that there is a greater likelihood of having students from many different ethnic/racial backgrounds. For students thinking about graduate school, the racial diversity of a program may not have been an influential factor, but after the research has been presented, perhaps it will be one. The hypothesis is that a program’s racial/ethnic diversity is positively associated with scholarly activity. This research has shown that in the natural sciences, racial/ethnic diversity does have a positive impact on scholarly activity; however, the opposite is true for social sciences.

## **Literature Review**

Several studies have been done regarding how racial/ethnic diversity affects various parts of society. In general, the conclusion has been that diversity has negative impacts on economic growth and other economic and social aspects (Easterly and Levine 1997; Alesina, Easterly and Baquir 1999; Alesina et al 2003). Less research has been conducted regarding racial/ethnic diversity among graduate students. Kivlighan (2008) carried out a study involving diversity

among PhD graduates and their research activity, mainly article publications. The results showed that African American graduates from a more ethnic and racial diverse graduate program published an average of 1.47 more articles than those from a homogenous program. In addition, international students from a diverse program had an average of 1.23 more articles published compared to an average of 0 for those from a homogenous program. Kivlighan only focused on one school, though, while the current study focuses on this question on a larger scale. In addition, research exists about diversity in undergraduate education. The results of these studies show that diversity has positive effects on students' perceptions of their undergraduate experience. For example, higher diversity showed a positive relationship with students' awareness and satisfaction with the college or university (Chang, 2001, Chang, Denson, and et al., 2006). Hurtado (2001) found that students that study with others from a different racial/ethnic background reported growth in civic, job-related, and learning outcomes (Hurtado, 2001). Not much additional research has been done focusing on ethnic/racial diversity and knowledge creation; however, so this research looks to minimize that gap.

## **Models**

This research uses the Ordinary Least Squares (OLS) multiple regression model for testing the impact of student diversity, percentage of faculty with grants, average number of Ph.D. graduates between 2002 and 2006, percentage of students with grants in 2002-2004, and percentage of first year students with full financial support on average faculty publications. The main variable of interest is the student ethnic/racial diversity.

Thus, my regression is presented as the following:

$$fac\_pub_{2006,i} = \alpha + \beta ethnic_{t,i} + \gamma X_i + u_i$$

Where  $\beta ethnic_{t,i}$  is the racial/ethnic composition of students in an individual department,  $i$ , in the year  $t$ , which spans from 2002 through 2004.  $\gamma X_i$  includes the other variables mentioned above and  $u_i$  is the error term.

The ethnicity/racial index is found from the following equation:

$$ethnic = 1 - \sum_{i=1}^n (S_i)^2$$

$$0 \leq ethnic \leq 1$$

$ethnic = 0$ , the program is completely homogenous

$ethnic = 1$ , the program is completely heterogeneous

This number gives the probability that any two students picked randomly in a particular program are from two different racial groups.

In order to counter the problem of endogeneity, where student racial/ethnic diversity and faculty publications might be jointly determined, the following Instrumental Variable regression is used:

Second stage regression:

$$fac\_pub_{2006,i} = \alpha + \beta ethnic_{t,i} + \gamma X_i + u_i$$

First stage regression:

$$fac\_pub_{2006,i} = \alpha + \beta ethnic\_county_{2000,i} + \gamma X_i + v_i$$

The instrument in this regression for student ethnic/racial diversity is county racial/ethnic diversity based on the same model as student racial/ethnic diversity where 1 is completely heterogeneous and 0 is completely homogenous. By mapping out each county a university or college is in, I was able to match it with the county diversity data. This is a valid instrument as

diversity of the county can affect student diversity; however, it is unlikely to affect the average number of faculty publications.

### **Data and Sources:**

The data used in this study were originally from the National Science Foundation (NSF) and the National Research Council (NRC). The NSF had information about student race and ethnicity for individual departments as well as the number of students in each department with grants and other funding. Nine different ethno-racial groups were identified in the NSF data: Foreign, Hispanic, Native American, Asian American, Black, Pacific Islander, White, Multi non-Hispanic, and unknown. The NRC had information on individual programs' average time to complete their degree, average faculty publications, and percent of faculty with grants. By combining this data, we were able to test the impact of diversity on scholarly activity for natural and social sciences.

There are a different number of observations among each regression as the data was not compiled all together when I received it, nor was all the information there for every single variable. Some information was available for a certain variable in one year but not for the next. In addition, the number of entities for the natural sciences and the social sciences differs quite a bit. This difference is accounted for by the fact that some schools have more natural science programs than social sciences. The actual universities and colleges in these two groups are the same, however. The natural sciences division includes departments such as biological sciences, chemistry, computer science, mathematics, physics, environmental sciences, and anatomy. The social science division includes economics, sociology, psychology, anthropology, geography, and political science.

The dependent variable is the average number of faculty publications while the primary independent variable of interest is student racial/ethnic diversity. Other control variables include percentage of faculty with grants, average number of Ph.D. graduates between 2002 and 2006, and the percent of first year graduate students with full financial support.

Descriptive statistics for both natural sciences (table one) and social sciences (table two) are shown in the appendix.

## **Methodology and Results**

Results for the natural sciences show statistical significance regarding the impact of ethnicity/race on faculty publications for all three years. Student race/ethnic diversity has a positive and significant effect on faculty publications in 2006 through each regression. For 2002, column (1), the coefficient of interest is 1.27, meaning a one percent increase in ethnic/racial diversity leads to a 1.27 unit increase in average faculty publications. For 2003 and 2004, columns (3) and (5), the coefficients of interest are both 1.19, meaning a one percent increase in ethnic/racial diversity leads to a 1.19 unit increase in average faculty publications. These are all significant at the 99 percent level. Adding the control variables decreases the impact of diversity on publications, but the results are still significant. For instance, column (2) shows that a one percent increase in diversity in 2002 leads to a 0.41 unit increase in average faculty publications. Similar results were found for 2003 and 2004. As expected, percentage of faculty with grants in 2006 and percentage of students with grants both have significant positive impacts on faculty publications in 2006. This can be explained by the fact that if a faculty member has a grant, they are more able to conduct research and get their research publicized. Students with grants must have shown promise academically, thus they may have original ideas and they are able to

conduct their own research or assist faculty in theirs. Percentage of faculty with grants has a larger impact on faculty publications than student diversity, which can also be expected. All three years show similar results in this area; a one percent increase in faculty with grants results in approximately a 1.55 unit increase in the average number of faculty publications. The percentage of students with full financial funding also has a strong positive and significant relationship with faculty publications. Average Ph.D. graduates from 2002-2006 has a significant positive effect on faculty publications, however the relationship is weak compared to the other variables. From table three in the appendix, the results show 99 percent significance for all variables except ethnicity/race in 2002 (regression 6) which is significant at 95 percent.

Table four shows the results of the social sciences with the dependent variable being average number of faculty publications in 2006. These results show that racial/ethnic composition of the department does not have any significant impact on faculty publications in either direction. In fact, none of the results were significant. This can partially be explained by the idea that faculty teaching in the social sciences do not rely on students to help with their studies unlike in the natural sciences. Curiously, the percentage of faculty with grants did not significantly impact the average number of faculty publications either.

The IV results (shown in tables five and six) are similar to the OLS results for both the natural and social sciences. Using all of the same control variables, results for the natural sciences show that student diversity in 2002 has a stronger, statistically significant and positive impact on faculty publications. In fact, a one unit increase in student diversity results in a 3.01 unit increase in the average number of faculty publications. Similarly, for diversity in 2003 and 2004, a one unit increase in diversity yields a 2.45 and 2.34 unit increase in faculty publications respectively. The standard T-test shows that ethno-racial diversity of the county is a valid and

significant instrument for student ethno-racial diversity as well. This is true for both the natural and social sciences. However, as with the OLS results, IV results were not statistically significant for social sciences. There were sign changes for the when using the IV regression for the social sciences compared to the OLS results, but again, those were not significant.

The differences in social sciences and natural sciences may be attributed to the fact that faculty in the natural sciences simply publish more articles compared to faculty in the social sciences. According to Lawrence (2002), faculties in the natural sciences report the largest number of publications both per year and throughout their career as a whole. In fact, he also found that faculties in the natural sciences devote more time to research than faculty in social sciences (Lawrence, 2002).

## **Conclusion**

Using the Ordinary least Squares multiple regression, the results from the data above show that higher racial/ethnic diversity among graduate students in 2002-2004 positively affects average faculty publications in 2006. In addition, percentage of faculty with grants in 2006, the average number of Ph.D.s completed in 2002-2006, percentage of students with grants, and the percentage of first year students with funding all have positive and significant impacts on average faculty publications. For natural sciences, the original hypothesis is confirmed for all regressions. The hypothesis is rejected in social sciences, though. None of our results were statistically significant.

Further research needs to be done to account for the name of the school. This would control for a school like Harvard as compared to a smaller, less well-known school. We have data on the type of institution each school is (Research University I or Research University II in



addition to Carnegie code) which might be useful in controlling for the name of the school. This is important as according to Lawrence (2002), faculties in research institutions spend more time on research whereas faculties in smaller schools, such as liberal arts colleges, spend more time devoted to teaching. This, of course, can have drastic differences in the results that were found in this paper.

## Works Cited

- Chang, Mitchell. "The Positive Educational Effects of Racial Diversity on Campus." *ERIC*. N/A.N/A (2001): n. page. Web. 11 Dec. 2011.  
<[http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?\\_nfpb=true&\\_&ERICExtSearch\\_SearchValue\\_0=ED456198&ERICExtSearch\\_SearchType\\_0=no&accno=ED456198](http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED456198&ERICExtSearch_SearchType_0=no&accno=ED456198)>.
- Chang, Mitchell, Denson Nida, et al. "The Educational Benefits of Sustaining Cross-Racial Interaction among Undergraduates." *Journal of Higher Education*. 77.3 (2006): 430-55. Print.
- Denson, N., & Chang, M.J. (2009). Racial diversity matters: The impact of diversity-related student engagement and institutional context. *American Educational Research Journal*, 46(2), 322-353.
- Dundar, Halil, and Darrell R. Lewis. "Determinants Of Research Productivity In Higher Education." *Research In Higher Education* 39.6 (1998): 607-631. *Academic Search Premier*. Web. 23 Mar. 2012.
- Hurtado, Sylvia. "Linking Diversity and Educational Purpose: How Diversity Affects the Classroom Environment and Student Development." *ERIC*. N/A.N/A (2001): n. page. Web. 11 Dec. 2011.  
<[http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?\\_nfpb=true&\\_&ERICExtSearch\\_SearchValue\\_0=ED456199&ERICExtSearch\\_SearchType\\_0=no&accno=ED456199](http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED456199&ERICExtSearch_SearchType_0=no&accno=ED456199)>.

- John M. Parente, et al. "Racial And Ethnic Diversity In The Classroom: Does It Promote Student Learning?." *Journal Of Higher Education* 72.5 (2001): 509-531. *Academic Search Premier*. Web.
- Kivlighan Jr, Dennis, M. . "Compositional Diversity and the Research Productivity of PhD Graduates." *Journal of Diversity in Higher Education*. 1.1 (2008): 59-66. Print.
- Lawrence, Janet H. "Faculty Performance of Research and Scholarship." *Encyclopedia of Education*. online. The Gale Group Inc., 2002. Web. 10 Apr 2012.  
<<http://www.encyclopedia.com/utility/printdocument.aspx?id=1G2:3403200228>>.
- Nettles, Michael, T. "Success in Doctoral Programs: Experiences of Minority and White Students." *American Journal of Education*. 98.4 (1990): 494-522. Print.
- Toutkoushian, Robert, K., and John, C. Smart. "Do Institutional Characteristics affect Student Gains from College?." *Review of Higher Education*. 25.1 (2001): 39-61. Print.
- Umbach, Paul. "The Contribution Of Faculty Of Color To Undergraduate Education." *Research In Higher Education* 47.3 (2006): 317-345. *Academic Search Premier*. Web.

## Appendix

Table One: Natural Sciences

<b>Variable Name</b>	<b>Observation</b>	<b>Mean</b>	<b>SD</b>	<b>Min</b>	<b>Max</b>	<b>Unit</b>
Faculty Pub	1431	1.81	1.21	0.03	9.16	Fraction
Ethnic	1423	0.49	0.14	0	0.80	Fraction
% Faculty with grants in 2006	1431	0.19	0.18	0	1.00	Fraction
Avg. PhD. Completed 2002 – 2006	1431	5.92	5.67	1.00	60.4	Fraction
% Students with Grants	1423	0.39	0.26	0	1.00	Fraction
% of First Year Students with Funding	1431	0.93	0.18	0	1	Fraction
Ethnicity of County	1419	0.42	0.19	0.03	0.78	Fraction

Table Two: Social Sciences

<b>Variable Name</b>	<b>Observation</b>	<b>Mean</b>	<b>SD</b>	<b>Min</b>	<b>Max</b>	<b>Unit</b>
----------------------	--------------------	-------------	-----------	------------	------------	-------------

Faculty Pub	626	0.79	2.64	0.01	48.19	Fraction
Ethnic	625	0.49	0.15	0	0.80	Fraction
% Faculty with grants in 2006	626	0.44	0.23	0	1.00	Fraction
Avg. PhD. Completed 2002 – 2006	626	6.17	4.40	1	26.8	Fraction
% Students with Grants	625	0.13	0.16	0	1	Fraction
% of First Year Students with Funding	626	0.80	0.28	0	1	Fraction
Ethnicity of County	626	0.43	0.20	0	1	Fraction

Table Three: Natural Sciences

Dependent Variable: Average number of Faculty Publications in 2006

	(1)	(2)	(3)	(4)	(5)	(6)		
<i>ethnic 2002</i>	<b>1.27***</b>	<b>0.41***</b>	<i>ethnic 03</i>	<b>1.19***</b>	<b>0.32***</b>	<i>ethnic 04</i>	<b>1.19***</b>	<b>0.36***</b>
	<b>(0.09)</b>	<b>(0.16)</b>		<b>(0.21)</b>	<b>(0.19)</b>		<b>(0.11)</b>	<b>(0.18)</b>

% of faculty with grants 2006	1.55*** (0.14)		1.53*** (0.15)		1.57*** (0.15)	
Average Ph.D. completed 02-06	0.08*** (0.01)		0.08*** (0.01)		0.08*** (0.01)	
% of students with Grants 2002	0.29*** (0.10)		0.31*** (0.09)		0.27*** (0.10)	
% of first year students with funding	0.68*** (0.12)		0.66*** (0.12)		0.66*** (0.12)	
$R^2$	0.02	0.29	0.02	0.29	0.01	0.19
Number of observations	1426	1425	1427	1426	1425	1424

Notes: Robust standard errors are in parenthesis.

\*\*\* Significant at 1%, \*\* Significant at 5%, \* Significant at 10%

Table Four: Social Sciences

Dependent Variable: Average number of Faculty Publications in 2006

	(1)	(2)	(3)	(4)	(5)	(6)		
<i>ethnic 2002</i>	<b>-0.07</b> (0.7)	<b>-0.02</b> (0.74)	<i>ethnic 03</i>	<b>0.16</b> (0.87)	<b>0.32</b> (0.97)	<i>ethnic 04</i>	<b>-0.49</b> (0.77)	<b>-0.31</b> (0.80)

% of faculty with grants 2006	0.39		0.29		0.10	
	(0.41)		(0.43)		(0.43)	
Average Ph.D. completed 02-06	-0.03		-0.03		-0.03	
	(0.02)		(0.03)		(0.02)	
% of students with Grants 2002	0.56		1.01		1.58	
	(0.92)		(0.99)		(1.06)	
% of first year students with funding	0.02		0.00		-0.05	
	(0.327)		(0.31)		(0.31)	
$R^2$	0.00	0.00	0.00	0.01	0.00	0.01
Number of observations	626	626	626	626	625	625

Notes: Robust standard errors are in parenthesis.

\*\*\* Significant at 1%, \*\* Significant at 5%, \* Significant at 10%

Table Five: IV Natural Sciences

Dependent Variable: Average number of Faculty Publications in 2006

	(1)	(2)	(3)
<i>ethnic 2002</i>	<b>3.01*</b>	<i>ethnic 03</i>	<b>2.45*</b>
	<b>(1.74)</b>		<b>(1.32)</b>
% of faculty with grants 2006	1.50***	1.56***	1.61***
	(0.16)	(0.16)	(0.15)
Average Ph.D. completed 02-06	0.07***	0.07***	0.08***
	(0.01)	(0.01)	(0.01)
% of students with Grants 2002	0.30***	0.27**	0.23**
	(0.11)	(0.11)	(0.11)
% of first year students with funding	0.71***	0.67***	0.65***
	(0.15)	(0.14)	(0.13)
$R^2$	0.21	0.24	0.25
Number of observations	1411	1412	1410

Notes: Robust standard errors are in parenthesis.

\*\*\* Significant at 1%, \*\* Significant at 5%, \* Significant at 10%



Table Six: IV Social Sciences

Dependent Variable: Average number of Faculty Publications in 2006

	(1)		(2)		(3)
<i>ethnic 2002</i>	<b>3.57</b>	<i>ethnic 03</i>	<b>3.06</b>	<i>ethnic 04</i>	<b>2.88</b>
	<b>(5.03)</b>		<b>(4.50)</b>		<b>(4.69)</b>
% of faculty with grants 2006	0.36		0.25		0.028
	(0.42)		(0.43)		(0.45)
Average Ph.D. completed 02-06	-0.03		-0.03		-0.03
	(0.03)		(0.03)		(0.03)
% of students with Grants 2002	0.83		1.32		1.87
	(1.15)		(1.24)		(1.33)
% of first year students with funding	-0.01		0.03		0.00
	(0.37)		(0.29)		(0.28)
$R^2$	0.00		0.00		0.00
Number of observations	625		625		624

Notes: Robust standard errors are in parenthesis.

\*\*\* Significant at 1%, \*\* Significant at 5%, \* Significant at 10%