STATISTICAL TERMINATION OR FEWER SELF-IDENTIFIED STUDENTS: WHAT IS CAUSING THE DECLINE IN AMERICAN INDIAN AND ALASKA NATIVE COLLEGE ENROLLMENTS?

Jeffrey D. Burnette
Rochester Institute of Technology

Jason T. Younker
University of Oregon

David P. Wick
Rochester Institute of Technology

PRELIMINARY DRAFT DO NOT CITE WITHOUT PERMISSION

1. INTRODUCTION

There is a long history of speculation that American Indians and Alaska Natives are a vanishing people that will eventually become indistinguishable from the mass population (Dippie 1991; Iverson 1998, pp.16-7). In 2007, the U.S. Department of Education (ED) announced new guidelines for collecting and reporting race and ethnicity data that require individuals be non-Hispanic, single-race American Indian and Alaska Native to count toward AIAN totals (U.S. Department of Education 2007). Those who also identify as Hispanic and AIAN are classified as Hispanic, while those who select another racial category are amalgamated into an alternate category, ``two or more races''. Prior to this, individuals selected the single option that best reflected their identity. Since the 2010 mandatory implementation date for this change, AIAN undergraduate enrollment has decreased by 26% decrease to 128,078 students (U.S. Department of Education 2018). At the same time, the newly created category of two or more races has grown by 105% and is now 4.5 times larger than the AIAN total. Similarly, the other category likely to collect AIAN students, Hispanic or Latino, has

also seen growth, 26%. While it is likely that these two categories have siphoned off part of AIAN undergraduate enrollment, their growth outstrips the decline in AIAN enrollment by such an amount that it is not at all possible to attribute the majority of either's expansion to a reclassification of AIANs.

The purpose of this paper is to examine AIAN undergraduate enrollment and determine if the trend exhibited in data published by the National Center of Education Statistics (NCES) is the result of fewer AIANs attending college or from the change in race and ethnicity protocols. Until now, most research has criticized the ED for its new guidelines but has limited its analysis to data on or before the mandatory implementation date. To our knowledge, this is the first study to explore the potential under-count in higher education using panel data. We use data from the Integrated Postsecondary Education Data System (IPEDS) and American Community Survey (ACS) to calculate the percentage of AIAN undergraduate enrollment for the period 2010 to 2016 and perform a difference-in-differences (DD) analysis. Overall, we find evidence that the decrease in AIAN undergraduate enrollment from 2010 to 2016 is a consequence of the process and procedures currently used by the NCES.

The next section details the history of enumerating AIANs in the United States and within higher education. This is followed by a review of the literature examining the impact of the ED's 2007 guidelines for reporting race and ethnicity and a description of our methodology. Finally, we conclude with a summary and contextualization of our results as well as a discussion of potential impacts from distorted IPEDS data.

2. BACKGROUND AND LITERATURE REVIEW

2.1 Enumerating American Indians and Alaska Natives in the U.S.

The United States began enumerating its population in 1790. Indians living in the general population were first counted as a stand-alone category in 1860, eleven years after the federal agency overseeing Indian Affairs was moved from the War Department to the newly created Department of the Interior. The reliability of data concerning the Indian population at this time as been questioned for many reasons, including interstate variation associated with enumerator specificity.[1]

In 1872, Francis A. White, Superintendent of the 1870 Census, raised a question that the federal government is still trying to answer; which individuals should be counted as part of the American Indian population? Essentially, he decided the fundamental the issue comes down to two overlapping sub-questions: should the government count only those who maintain their tribal relation and how do you appropriately classify individuals with mixed parentage? (U.S. Census Bureau 1872, pp. 19) Then as now the government-to-government relationship between federal and tribal governments caused confusion for enumerating American Indian identity. If the government is concerned with upholding its federal trust responsibility an accurate assessment of the issues and challenges faced by tribal citizens becomes paramount. Similarly, if policymakers are developing initiatives for those who descend from but are not tribal citizens it necessitates data that accurately depicts their conditions. The issue

[1] The tables from California are more detailed by comparison to other states. Overall, California is estimated to contain 40% of the total population (44,020) while Mississippi reports a total of two Indians (Jobe 2004, pp. 70).

becomes even more convoluted by mixed parentage and differential tribal enrollment requirements. Superintendent White, struggled only with the first part of this issue and focused on blood purity as well as the impact on categorical totals from using superior versus inferior blood or matrilineal versus fraternal descent for codification. Eventually, he decided that Indians of mixed race would be classified based upon their level of assimilation. Those who "are found residing with whites, adopting their habits of life and methods of industry, such persons are to be treated as belonging to the white population." Where, on the other hand, they are found in communities composed wholly or mainly of Indians, the opposite construction is taken. Subsequently a schedule was developed for the Special Census of Indians (not taxed) in 1880 to determine the proper assignment of mixed-race individuals (Jobe 2004). This was followed by the first full scale report of the Indian population; the Census of 1890, included both Indians not taxed (those living on a reservation or unsettled area) and Indians taxed. It also produced statistical summaries and descriptions of each tribe by state along with other socioeconomic information (U.S. Census Bureau 2018).

The passage of the Indian Citizenship Act of 1924 brought about a new issue concerning the enumeration of American Indians. Because the United States now declared that all Indians residing within the United States were now citizens it now became necessary for census takers to count all Indians, taxed and not taxed[2] and distinguish between Indians and Mexicans. Enumerators in Arizona, New Mexico and

[2] The 1930 Census was the last to deduct Indians not taxed from the population total for the purposes of proportioning Congressional representation.

California[3] were instructed to "take special care to differentiate between Mexican laborers and Indians." because "Some Mexican laborers may endeavor to pass themselves as Indians." Furthermore, enumerators were informed to rely upon community members since "persons residing in the region should have no difficulty in differentiating between the two types." The classification of mixed-race American Indians was still an issue and instructions for American Indians were dependent upon their other group of descendancy. When mixed with white, they were to be considered Indian "except where the percentage of Indian blood is very small, or where the individual was regarded as a white person in the community where he lives." However, "A person of mixed Indian and Negro blood should be returned as a Negro unless then Indian blood predominates and the status as an Indian is generally accepted in the community." (U.S. Census Bureau 1937, pp.1). The subjectivity of these rules, their application and the varied levels of training for enumerators led to substantial opportunity for undercounts for the Indian population.[4]

Federal assimilation policy in the 1950s left an indelible mark on Native nations through House Concurrent Resolution 108 (HCR-108) and the urban relocation program. Passed in 1953, HCR-108 officially heralded the termination era that ended federal recognition for approximately 110 tribes and bands, 11,500 people or 3% of tribal citizens (Wilkinson and Biggs 1977). Whereas, the urban relocation was a voluntary

[3] Arizona, New Mexico and California are the only states at this time that contain at least 4% of both the Mexican and Indian population. Oklahoma contained 28% of the Indian population less than 1% of the Mexican population. Whereas, 48% of Mexicans lived in Texas which was home to less than 0.5% of all Indians.

[4] A detailed discussion of undercounts of the American Indian and Alaska Native population is provided by Carol Lujan (1990)

program it has also been referred to as individual termination because federal services were limited to those living on reservations (Burt 1986). Both policies were meant to assimilate American Indians into the general population and reduce the federal trust responsibility as individuals who were classified as AIAN one day were no longer considered AIAN by the government the next.

The switch from enumerators determining race to self-identification occurs with the Census of 1960 and with it came a pronounced growth in the American Indian population.[5] While part of the rise from 1950 to 1960 may be attributed to an undercount[6] of the 1950 population, scholars found this new trend to be associated with increased identification and began investigating different sub-populations within the larger AIAN group (Jeffrey S. Passel 1976; Jeffrey S. Passel and Berman 1986; Snipp 1986; Harris 1994; Eschbach 1993; Jeffrey S.Passel 1997).

In response to an undercount of the Hispanic population[7] during the 1970 Census, policy makers decided to introduce a separate question concerning ethnicity to the 1980 Census. At this time, the overlap between the Hispanics and AIANs was relatively small[8]. In contrast the change allowing individuals to report multiple race

[5] The increase in the American Indian population from 1950 to 1960 was 48%. This was higher than what had taken place in the previous 50 years. From 1900 to 1950 the Indian population grew by 45% (Passel 1986, pp.82).
[6] U.S. Tribal enrollment in 1952 exceeds the Census enumeration of 357,499 in 1950 by 13%. By comparison, 1952 Tribal enrollment comprises 77% of the 1960 census total. Furthermore, Tribal enrollments in 1981 consist of 65% of the 1980 Census total and 69% of those enumerated in the 2000 Census are enrolled. (Thornton 2005)
[7] High end estimates for the undercount were just under 14%, with those descending from Puerto Rican parentage or birth slightly lower than those whose parents were from Mexico. By comparison, the undercounts for blacks and whites was 7.7% and 1.9%, respectively (Siegel and Passel 1979, pp.34).
[8] Mathew Snip (1986} finds that only 3% of the 1,126,760 reporting AIAN as their race also identify with Hispanic ancestry while 75,600 individuals reporting AIAN ancestry also identify as Hispanic.

categories made in 2000 made an impact on AIAN totals not seen since the change to racial self-identification.[9] The demographic change from 1990 to 2000 is documented in Lieber and Ortyl (2014) and Liebler, Bhaskar and Porter (née Rastogi) (2016) continues this work by describing the changes from 2000 to 2010.

2.2 Enumerating Race and Ethnicity in Higher Education

Data collection in Higher education has a much shorter history than that of the Census. When surveys were initially collected in 1965 by the Higher Education General Information Survey (HEGIS) schools were asked to report the predominant race of the student body and were limited to either white or black. The first time HEGIS collected race and ethnicity information from institutions it used self-identified data for those graduating during the 1976-1977 academic year.[10] In 1985, the U.S. Department of Education (ED) began the two-year phase-in of IPEDS but still kept reporting optional. It isn't until 1992 that it become mandatory for all institutions that receive federal student financial aid (Title IV) funds to complete surveys conducted by IPEDS as part of its postsecondary institution data collection effort. (Aliyeva, Cody and Low 2018).

The largest change to race and ethnicity reporting in IPEDS came in 2008, schools were allowed to implement the multiple race reporting that was established during the 2000 Census. Notice of the new reporting rules was delivered in 2007 via the Federal

[9] The total number of individuals reporting an AIAN identity increased by 110% from 1990 to 2000. This percentage increase is reduced to 26% if the 2000 total is limited to just those who report as single-race AIAN (Ogunwole 2002, pp.5). The difference is less drastic from 2000 to 2010. The growth in single-race AIANs is 18% while for those identifying as AIAN alone or in combination it is 27% (Norris, Vines and Hoeffel 2012, pp.4).

[10] The categories used for 1976 are the largely the same used today: non-resident aliens, Hispanics, and non-Hispanic: whites, blacks, AIANs and Asians or Pacific Islanders.

Register in "Final guidance on maintaining, collecting, and reporting racial and ethnic data to the U.S. Department of Education" and involved a response to solicited comments. Concerns of respondents primarily revolved around potential of undercounts for single-race reporting categories due to the ED's tabulation instructions. Institutions were required to aggregate any individual that expressed Hispanic ethnicity into Hispanic totals, regardless of whether that individual also communicated a racial identity. Similarly, individuals selecting two or more race categories no longer have any association with any of their self-identified races. Instead, they are reclassified under the moniker "two or more races" (U.S. Department of Education 2007). Furthermore, the ED discounted the warnings of scholars by explaining that while totals in single-race categories may decrease, proportional representation for these groups was unlikely to change due to the comparative small number individuals expected to be assigned two or more races.

2.3 Literature on distortions in IPEDS race data and the ED's 2007 Final Guidance

One strand of research has concentrating on issues of aggregation of racial categories in higher education has focused on differences between racial groups. Thomson (2011) concentrates on the differential experiences (parental education, graduation rates, perceptions and goals) of individuals and how it correlates with race and ethnicity. He also references Thomson (2007) which finds that the ED's new protocols significantly impact the percentage of representation of certain racial groups. Meanwhile, Teranashi, Lok and Nguyen (2013) examines heterogeneity within the

Asian population and argues that combining all individuals into a single Asian category removes important variation that distinguishes individuals from one another.

Despite the ED mandating that institutions apply the new race and ethnicity protocols for 2010, most research investigating the misrepresentation of racial data in IPEDS has not included data after 2010. Broh and Minicucci (2008) provides a scathing critique of the ED's new guidelines and collected data at the Consortium on Financing Higher Education's 31-member institutions in 2007 to demonstrate their point. They compare enrollment percentages for each race and ethnicity group when using IPEDS rules to those they recommend. Their preferred classification system results in higher percentages for Asians, blacks and whites at the expense of the two or more category and a slight reduction to Latino enrollments. Rose (2012) conducts an analysis of AIAN undercounts for primary and secondary schools in New York state (NYS) by comparing 2010 NCES totals for individual school districts to those from the NYS School Report Card. He finds evidence that these two data sources are inconsistent with one another and that there seems to be no discernible pattern of under or over counts for NCES data. Byrd, Dika and Ramial (2013) investigate the impact of classifying all non-U.S. citizens into the non-resident alien category and ignoring their self-reported racial identity prior to the change in ED protocols. Their study compared the proportional representation for each racial category from IPEDS and the ACS to find data in the areas of engineering as well as biological and biological sciences misrepresents racial and ethnic populations by lumping non-citizens into a single category that ignores their race.

Using ACS and IPEDS data from 2000 to 2016, Burnette, Younker and Wick (2018) conducts an exploratory difference-in-differences analysis to determine if there is evidence of undercount for AIAN undergraduates. A limitation of that study is the relatively small number of observations due to the use of annual national data. Consequently, there are a total of seventeen observations. Nonetheless, evidence of a significant undercount of AIANs in IPEDS data is found to range from 10% to 40% depending upon the selection of ACS control group. This current study converts both datasets into seventeen state level panels to increase the number of observations and estimate the intertemporal nature of the policy change and validate the robustness of those earlier results.

3. DATA AND METHODS

3.1 Data

Data for this study covers the 2000-2016 timeframe and comes from the ACS 1-year estimates and IPEDS Fall Enrollment (EF) surveys. A benefit of the ACS is race and ethnicity definitions are unchanged during the entire period and contain information at a more granular level than IPEDS. It also asks individuals about their college enrollment status and is designed to take advantage of information like race, sex, age and Hispanic origin to show characteristics of the population at the county level. Given this and that the ACS is not meant to produce population counts, we aggregate data to the state level and using person weights and concentrate on the percentage of undergraduates and graduates[11] who identify as AIAN instead of enrollment totals for

[11] Undergraduate and graduate totals are limited to U.S. citizens.

a total of 51 observations per year.[12] Lastly and most importantly, using the ACS provides an additional value; as long as the Census Bureau successfully constructs a representative sample, changes in enrollment patterns associated with societal causes or trends for a particular racial group should be reflected in the changes in that group's proportional enrollment. Consequently, IPEDS and ACS data should both contain any potential disproportionate reduction in AIAN enrollment related to the Great Recession or other downward pressures.

Race and ethnicity are reported using a few different variables in the ACS. All those from Hispanic, Latino or Spanish origin also provide information concerning their race. Individuals select from a number of categories the option(s) that apply or write in the appropriate category. These can then be rolled up into one of the 7-classification categories used by the ED.[13] In addition, race information is reported: using partially collapsed categories that are mutually exclusive but provide more detail than IPEDS, and all-inclusive categories. These all-inclusive categories include those who select AIAN, Asian, black or African American or white alone or in combination with some other race as part of that group's total. Finally, individuals in the ACS may also be assigned to their one-race bridged option. In 2000, the Census moved from a single box race option to a check all that apply approach. Consequently, Liebler and Halpern-Manners modified a technique used by the National Center for Health Statistics that

[12] Washington D.C. is included as its own observation. We also exclude data for Puerto Rico because our primary focus is the American Indian and Alaska Native population.

[13] IPEDS reports the following categories: American Indian or Alaska Native, Asian, black or African American, Hispanic or Latino, white, unknown and non-resident alien.

creates an individual's most likely single-race response using a number of covariates and developed an algorithm that produces that variable for ACS data.

We generate three different categories for AIANs from ACS data to serve as control groups for IPEDS data. Because the one-race bridge option is designed to replicate the data reporting process that was used by IPEDS before 2008 and can be applied over the entire time frame we anticipate it to provide the best counterfactual estimate. In contrast, our other categories potentially serve as a lower and upper bound. Non-Hispanic single-race AIANs should underestimate the size of the impact from switching reporting protocols at a national level. This is because prior to 2008 it excludes those individuals whose predominant identity is AIAN and possess a Hispanic or other racial identity and after the change IPEDS data should transition to mirror these values. Similarly, the all-inclusive AIAN category may overestimate the impact of ED's new regime at the national level because it will likely include many individuals that would not have identified as AIAN under the old protocols and would be further biased if individuals are more likely to identify as AIAN in combination now than prior to 2008.

The total number and percentage of undergraduate students who identify as AIAN from IPEDS data is depicted in figure 1. From 2000 to 2009 total enrollment increased on an annual basis. Whereas, a constant decrease began with the mandatory implementation date for the new protocols. The percentage of AIAN undergraduates follows a slightly different pattern. While again there is a steady decrease post 2009, the

period before remained relatively steady with AIANs comprising approximately one percent of the undergraduate population.

ACS estimates of the percentage of AIAN undergraduates are also displayed in figure 1. Overall, each exhibits a bit more between year variation; this is likely the result of sample variation that is often associated with minority populations. The percentage of non-Hispanic, single-race and one-race bridged AIAN undergraduates seems to follow a similar pattern as that from IPEDS but with slightly smaller decreases during the 2010 to 2016 timeframe. Overall, the representation of AIANs has decreased from 2007 to 2016 for most AIAN identity groups. The lone exception occurs when using the inclusive definition (any AIAN) that classifies an individual as AIAN if she reports an AIAN identity alone or in combination with some other identity. In that instance, the average year-to-year change from 2007 to 2016 is 1.1%. In contrast, the largest reduction for this period occurs using IPEDS data, -2.7% and is followed in succession by non-Hispanic, single-race, - 2.3%, and one-race bridged AIANs, -0.4%.

It is possible that the difference between the percentages for the any AIAN category and other groups may be the continued result of changes in self-identification patterns. If a previously unidentified AIAN individual becomes more likely to self-report an AIAN identity she would be now be classified as Any AIAN instead of just the aforementioned group. A similar but counter situation also may be helping to drive the decrease in the percentage of non-Hispanic, single-race undergraduate students. Consequently, these control groups are included for robustness and our primary focus is on the results obtained from the one-race bridged category.

[INSERT FIGURE 1 HERE]

A somewhat similar story is depicted in figure 2 but for graduate students. Again, IPEDS totals for AIANs decrease once the ED's policy becomes mandatory but because the population of graduate students is much smaller the effects seem a bit muted. This is especially the case when considering the percentage of AIAN graduate students, which 0.17 percentage points, from 2007 to 2016 as compared to the drop of 0.28 percentage points experienced by AIAN undergraduates.

The data from the ACS is takes a noticeable drop in 2002 for all AIAN categories and then rebounds in 2003. The 2002 reduction in the percentage of AIANs also can be seen to have occurred for most undergraduate groups. In addition, starting in 2008 the percentage of inclusive AIANs begins to oscillate around 1.28% and finally makes a large jump up without an offset in 2014. Percentages for other AIAN identities from the ACS are more consistent year-to-year during the 2003 to 2014 timeframe with an uptick for 2015.

[INSERT FIGURE 1 HERE]

3.2 Difference-in-Differences

Institutions in higher education are required by the ED to maintain the detailed race and ethnicity data they collect on their students for at least three years. Outside access to this data is limited and may not be possible. In the absence of a direct comparison, it is necessary to generate a second-best counterfactual and compare it to what is predicted to have happened had there been no change. The difference-in-differences (DD) methodology has served as a frequent tool to evaluate the impact of policies in

these circumstances. The selection of a control group is of paramount importance; it must follow a parallel path to the treatment group prior to the policy and have no reason to diverge other than the policy being evaluated. Zhang (2018) uses the reliability of ACS data and its ability to create a counterfactual for IPEDS data to implement a DD analysis to assess the impact of the Post-9/11 GI Bill on college enrollment.

In the most general form we employ the following ordinary least squares regression equation:

$$E_{it} = \beta_0 + \beta_1 A_i + \beta_2 P_t + \beta_3 (A_i \text{ x } P_t) + \theta_x F_x + \epsilon_{it} \qquad (1)$$

Where:

E_{it} - is the percentage of AIAN enrollment in dataset *i* at time *t*.

A_i – is a dataset dummy variable equal to 1 if data comes from the IPEDS and 0 otherwise.

P_t – is a policy dummy variable equal to 1 if data comes on or after the mandatory implementation date, 2010, and 0 otherwise.

F_x – is a set of fixed effect dummies representing effect x.

In this specification, β_3 is the coefficient of interest because it estimates the differential impact between ACS and IPEDS data during the post policy period. This requires that we control for the average difference between the two datasets, β_1, and the average difference between the pre and post policy periods, β_2. In the next section we use this equation and incorporate both state and time fixed effects in order to replicate the national level results of Burnette, Younker and Wick (2018) using state level data. We also adopt an alternate specification that allows the calculation of β_3 on an annual basis that includes state but not time fixed effects. In all cases, we weight each

observation by the total (undergraduate or graduate) AIAN student population for that state from either IPEDS or ACS as appropriate.

Another consequence from the data collection process used by ED and performing DD for state level values is that the size of β_3 is expected to be time dependent. Unless students decide in mass to update their race and ethnicity information at their school a significant number of students counted in IPEDS totals from 2010 to 2016 are likely to have completed their information under the old race and ethnicity classification regime. Furthermore, as students withdraw, transfer or graduate from each college the percentage of students who submitted using the old protocols will continue to decrease until nearly all have completed their information using the new system. Similarly, the first year that a college implements to ED's new policy the only students likely to be included are new entrants. Because of this, we include neither time fixed effects nor a trend variable in models that estimate β_3 on an annual basis.

4. RESULTS

4.1 Average Policy Effects, 2010-2016

Table 2 depicts the results of regression equation (1) with robust errors for the three ACS control groups (non-Hispanic, single-race AIANs, one-race bridged AIANs and Inclusive AIANs) without fixed effects, with state fixed effects in addition to both state and time fixed effects.

Table 1 supports the results for undergraduates from Burnette, Younker and Wick (2018) while taking advantage of state level data and reports the estimates for graduate students. The estimates for undergraduates are consistent and significant

across control group; the reduction in the percentage of AIAN undergraduates calculated using IPEDS data from 2010 to 2016 is more severe than is expected from ACS data. In contrast, the results for AIAN graduate students is less definitive as only the one-race bridged and Any AIAN categories return significant results for a post 2009 reduction in IPEDS data. Both undergraduate and graduate estimates display the trend we hypothesized concerning the control group's degree of AIAN identification; smaller impacts are associated for non-Hispanic AIANs and largest when any AIAN identity is expressed.

[INSERT TABLE 1 HERE]

4.2 Annual Policy Impacts, 2008-2016

The results of modifying equation (1) is to estimate a different policy dummy for each year from 2008 to 2016 are displayed in Table 3. In this instance one dummy variable is equal to 1 when the year is 2008 or after while all other dummy variables are equal to 1 only for that particular year. This approach allows IPEDS data to be impacted during the ED's optional implementation period while all other coefficients measure the marginal effect relative to 2008. In this specification, we only present only the results for state fixed effects because including time fixed effects results in an over specification that eliminates annual impacts.

The coefficients obtained from the DD regression analysis, presented in table 2, help provide a more nuanced understanding of the impact on AIAN undergraduate enrollment from the new policy. Most notable is that the estimates of the interaction between IPEDS data and individual years are very similar in size and become

significant at some point after the mandatory implementation date. This occurs because the coefficient for the dummy variable after 2007 picks up the average change in both ACS and IPEDS data that occurs for those who identify only as AIAN and when the control group is Any AIAN identification. Data for IPEDS and srnH percentage of AIAN undergraduates is 0.07 percentage points lower from 2008 to 2016 than it is from 2000 to 2007. In contrast, when using any AIAN this value is 0.20 points higher overall. However, it is necessary to combine these estimates with those that measure the average difference between IPEDS and ACS data after 2007. In total the net estimate is slightly negative across all control groups, -0.0151, -0.051 and -0.0133 for only, one-race bridged and any AIAN respectively.

The year-to-year estimates for undergraduates are remarkably similar and gives confidence that the year-to-year estimates of the policy effect only collect the impact from changing protocols. The largest difference across control groups equal to 0.0076 and occurs in 2012. Furthermore, the estimates are consistent with the idea that the timing of data collection from students causes a time dependent effect as students withdraw, transfer or graduate from their educational institution. All year-to-year estimates increase in size at a decreasing rate and become significantly different from zero after 2012. By 2016, the estimated reduction in the percentage of undergraduate AIANs is 0.24 percentage points. A strikingly large number considering that according to IPEDS data the percentage of AIANs in 2016 is 0.77% of the total undergraduate population.

The estimates for graduate students possess many of the same patterns as those for undergraduates but in several instances lack significance. The only year-to-year estimate found to be significantly different from zero when using any AIAN as the control group is that for 2016. However, in all other areas the results for AIAN graduates are remarkably similar. Again, the coefficients grow in size and are significant when using non-Hispanic, single-race and one-race bridged AIANs as comparison groups. In addition, the year-to-year estimates for non-Hispanic, single-race AIANs obtain significance first and is followed by those for one-race bridged in the subsequent year. The starkest similarity is in the consistency of the estimates across control groups. Again, the largest difference between the year-to-year coefficients across groups is less than 0.01.

[INSERT TABLE 2 HERE]

4.3 Average Undercounts from Annual Estimates by Year, 2010-2016

In table 3, the adjusted total and percentage of AIAN enrollment is presented with estimates of the number individuals that are not classified as AIAN when using the new protocols but would have been under the old regime. The adjusted values for total and percent enrollment are calculated for all control groups based upon coefficients from the DD regression and then averaged for each year. Overall, we do find a downward trend in total AIAN undergraduate enrollment starting in 2011 with a slight reduction in the proportion of undergraduate AIANs. However, this reduction is far smaller than that observed in IPEDS data. At the same time, the size of the estimated undercount for AIAN undergraduates increases with each passing year. By

2016, we estimate there are 37,996 AIAN undergraduates not classified as AIAN and truly staggering number when considering that is just under 30% of the total number of AIAN undergraduate students according to IPEDS.

[INSERT TABLE 3 HERE]

A similar set of values is displayed in table 4 for AIAN students enrolled in a graduate or professional degree program. Again, the total number of AIANs begins a modest but steady decline in 2011. However, there is no discernable reduction in the proportional representation of AIAN graduate students from 2008 to 2016; the adjusted percentage of enrollment is constant at 0.65%. The undercount of AIAN graduate students follows the same trend as for undergraduate students as it increases from year to year but at a decreasing rate. Lastly, the size of the undercount as a percentage of the total from IPEDS, is incredibly consistent with that obtained for undergraduates for all years; the 2016 value for both undergrad and graduate students is approximately 30%.

[INSERT TABLE 4 HERE]

5. CONCLUSION AND DISCUSSION

The results from using ACS data to examine the existence of an undercount of AIAN college students are alarming. There is evidence that the "Final guidance on maintaining, collecting and reporting racial and ethnic data to the U.S. Department of Education" has caused IPEDS data to significantly reduce the total number and proportion of American Indians and Alaska Native college students enrolled at both the undergraduate and graduate level by as much as 30% of the reported value. In addition, the size of the AIAN college student undercount has grown in magnitude

each year but appears to be leveling off. We believe this is likely due the persistence of students who submitted their race and ethnicity information under the previous regime and that the percentage of AIAN students will settle into a new steady state level as these students either update their information or transition out of their respective institution. Finally, we find that the choice of econometric model is important. Separating the data into pre and post implementation periods that presumably calculated an average effect led to undercount estimates that were dependent upon the selection of control group. In contrast, the inclusion of dummy variables that concentrated the effect for that particular year with other variables that differentiated the post period and data source yielded estimates that were independent from the control group chosen.

When considered within the context of the history of federal assimilation policies targeting Native American tribal nations especially within the termination era these results become deeply disturbing. In the 1950s, the U.S. government explicitly declared that it would no longer recognize the sovereignty of individual tribal nations as it sought to assimilate them within the greater societal collective. In 2007, the U.S. Department of Education announced that it would allow individuals to express the complexity of their racial and ethnic identity more fully by allowing them to select all the categories that apply. At the same time, they constructed a reporting mechanism that eliminates the racial identity of individuals who respond they possess a connection with more than one community as they are placed into a category they may not have even selected. Future research that investigates whether students within higher

education know how their race and ethnicity data are processed and if their responses are dependent upon the method of aggregation would be useful to either validate or refute the ED's current data collection strategy.

The act of reducing racial and ethnic identity to a form that can be used for quantitative analysis necessitates that certain details are lost. Because of the unique government-to-government relationship that exists between sovereign Native American tribal nations and the United States federal government American Indian there are additional complexities and responsibilities. These have not been fully accounted for in this analysis due to its reliance on self-reported data. Our analysis is not able to differentiate between individuals who are enrolled citizens of tribal nations and those whose identity is more consistent with the social construction of race associated with other racial groups. In order to more accurately determine the impact of the ED's final guidance on tribal citizens it will be necessary to access data that includes specific information on tribal enrollment status.

The undercount of AIANs also has policy ramifications because data distortions can prevent policy makers from identifying their optimal choice and decisions concerning resource allocation are often driven by population size and the expected return on investment. The perception that AIANs are vanishing from higher education or are relatively small can limit institutional investment. Research that demonstrates the importance of a critical mass of students and faculty in determining student success can make it difficult for an institution to justify the amount of resources necessary to recruit and maintain a successful AIAN student population when other historically

marginalized communities with larger numbers have already met or are closer to the critical mass threshold. Corporations interested in developing a diverse talent pool of applicants with postsecondary degrees may make a similar calculation and determine that the costs of finding and hiring a qualified AIAN candidate exceeds their acceptable parameters. A necessary component for the development of evidence-based programs and policies is accurate data. The bias we have found in IPEDS data for AIANs calls into question the reliability of IPEDS data. Consequently, there is a need for research that examines if these results are limited to AIANs or potentially forebode the experience for other racial groups.

FIGURE 1
Total and Percentage of Undergraduate of American Indians and Alaska Natives

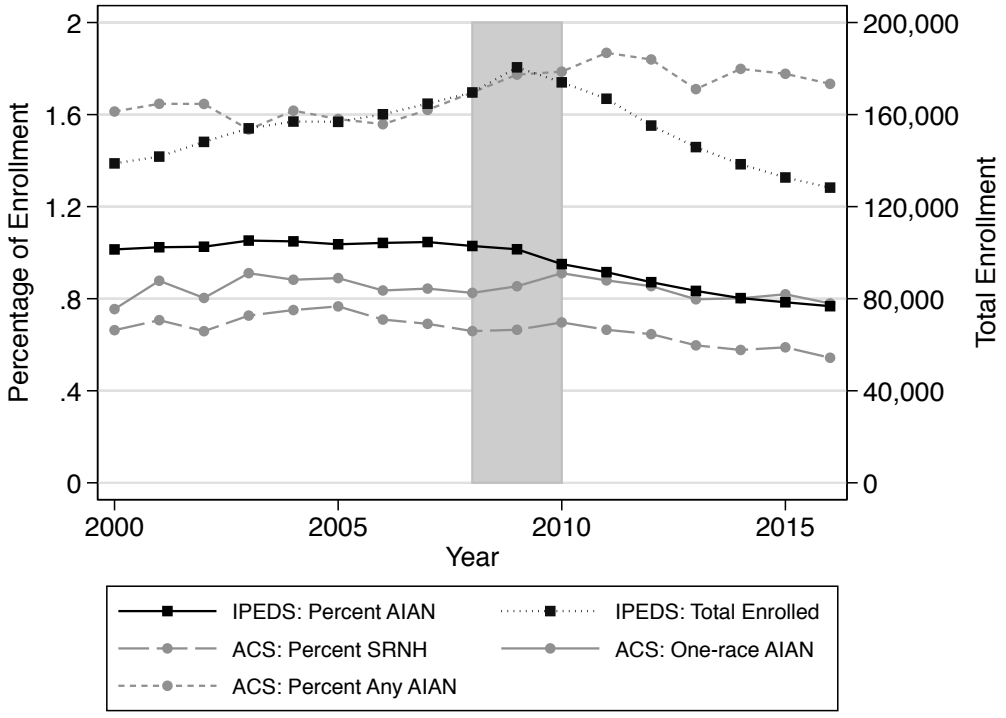

FIGURE 2
Total and Percentage of American Indian and Alaska Native Graduate Students

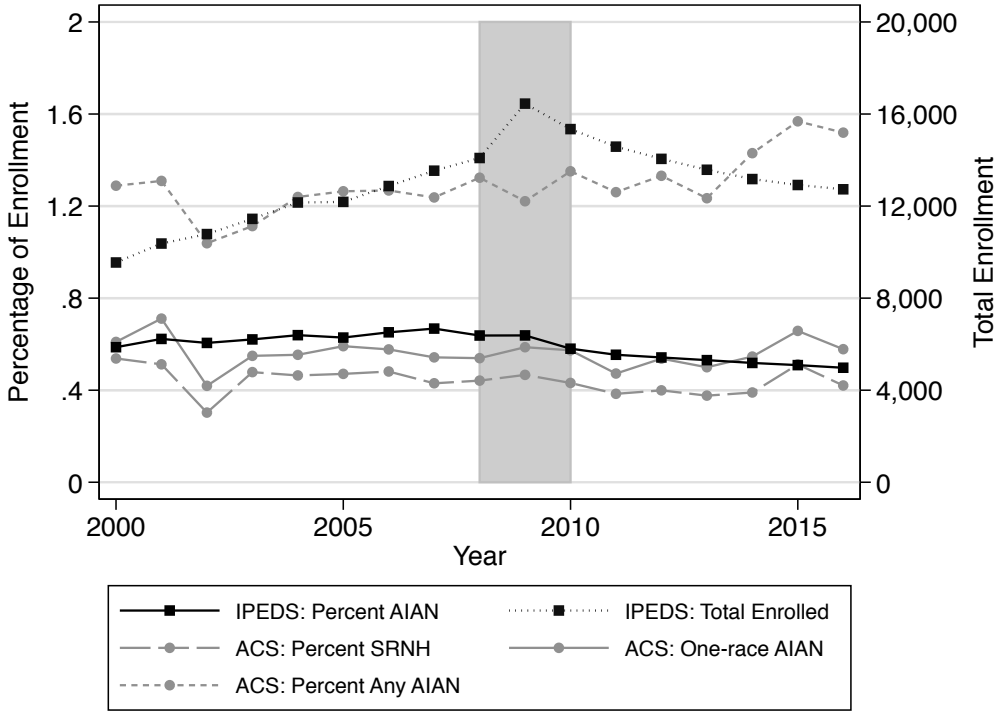

Table 1 – DD Results - 2010-2016: Average Policy Effect on Enrollment Percentage,

	Undergraduate Students			Graduate Students		
	Only AIAN	One-race Bridged	Any AIAN	Only AIAN	One-race Bridged	Any AIAN
IPEDS	0.3245***	0.1787***	-0.6084***	0.1841***	0.0832***	-0.5708***
	(0.0228)	(0.0248)	(0.0331)	(0.0206)	(0.0236)	(0.0361)
After 20009	0.0353	0.1288*	0.2311*	-0.0438	0.0141	0.1478
	(0.0604)	(0.0631)	(0.0957)	(0.0744)	(0.0879)	(0.1454)
IPEDS x After 2009	-0.1164***	-0.1841***	-0.3507***	-0.0489	-0.0728*	-0.2294***
	(0.0300)	(0.0342)	(0.0499)	(0.0291)	(0.0333)	(0.0527)
State Fixed Effects	Y	Y	Y	Y	Y	Y
Time Fixed Effects	Y	Y	Y	Y	Y	Y
N	1734	1734	1734	1734	1734	1734
R-squared	0.92	0.93	0.93	0.73	0.72	0.72

[CAPTION] Note: * - p<0.05, ** - p<0.01, *** - p<0.001, standard errors in parenthesis.
Source: Author's calculations from IPEDS and ACS

Table 2 – DD Results: Annual Policy Effect on Enrollment Percentage, 2010-2016

	Undergraduate Students			Graduate Students		
	Only AIAN	One-race Bridged	Any AIAN	Only AIAN	One-race Bridged	Any AIAN
IPEDS	0.3220***	0.1852***	-0.5669***	0.1812***	0.0797**	-0.5629***
	(0.0265)	(0.0276)	(0.0358)	(0.0246)	(0.0285)	(0.0444)
After 2007	-0.0673**	0.0025	0.1995***	-0.0419	-0.0189	0.1330**
	(0.0244)	(0.0271)	(0.0393)	(0.0285)	(0.0332)	(0.0508)
IPEDS x After 2007	0.0522	-0.0176	-0.2128***	0.0479	0.0289	-0.1164
	(0.0490)	(0.0496)	(0.0620)	(0.0374)	(0.0419)	(0.0668)
IPEDS x Year 2009	-0.0184	-0.0176	-0.0163	0.0015	0.0020	0.0028
	(0.0526)	(0.0563)	(0.0711)	(0.0312)	(0.0345)	(0.0589)
IPEDS x Year 2010	-0.0820	-0.0806	-0.0770	-0.0564*	-0.0551	-0.0518
	(0.0510)	(0.0575)	(0.0771)	(0.0275)	(0.0329)	(0.0615)
IPEDS x Year 2011	-0.1149*	-0.1133	-0.1086	-0.0830**	-0.0818**	-0.0787
	(0.0501)	(0.0583)	(0.0817)	(0.0259)	(0.0315)	(0.0617)
IPEDS x Year 2012	-0.1561**	-0.1542**	-0.1485	-0.0950***	-0.0933**	-0.0897
	(0.0500)	(0.0592)	(0.0854)	(0.0255)	(0.0322)	(0.0642)
IPEDS x Year 2013	-0.1889***	-0.1873**	-0.1814*	-0.1088***	-0.1071***	-0.1040
	(0.0503)	(0.0598)	(0.0873)	(0.0257)	(0.0320)	(0.0643)
IPEDS x Year 2014	-0.2167***	-0.2157***	-0.2101*	-0.1199***	-0.1180***	-0.1159
	(0.0481)	(0.0576)	(0.0874)	(0.0259)	(0.0335)	(0.0655)
IPEDS x Year 2015	-0.2313***	-0.2303***	-0.2244*	-0.1280***	-0.1260***	-0.1229
	(0.0479)	(0.0577)	(0.0879)	(0.0263)	(0.0343)	(0.0665)
IPEDS x Year 2016	-0.2443***	-0.2437***	-0.2376**	-0.1377***	-0.1365***	-0.1333*
	(0.0480)	(0.0581)	(0.0897)	(0.0259)	(0.0341)	(0.067)
State Fixed Effects	Y	Y	Y	Y	Y	Y
N	1734	1734	1734	1734	1734	1734
R-squared	0.92	0.93	0.93	0.73	0.72	0.72

[CAPTION] Note: * - p<0.05, ** - p<0.01, *** - p<0.001, standard errors in parenthesis.
Source: Author's calculations from IPEDS and ACS

Table 3 – Undercounts of American Indian and Alaska Native Undergraduates

	IPEDS Values		Adjusted Values		Estimated Undercount	
Year	Total Enrollment	Percent Enrollment	Total Enrollment	Percent Enrollment	Total Enrollment	As Percentage of IPEDS Total
2008	169,607	1.03%	-	-	-	-
2009	180,531	1.01%	-	-	-	-
2010	174,004	0.95%	185,972	1.02%	11,968	6.9%
2011	166,880	0.92%	184,706	1.01%	17,826	10.7%
2012	155,222	0.87%	179,874	1.01%	24,652	15.9%
2013	145,874	0.83%	175,854	1.01%	29,980	20.6%
2014	138,420	0.80%	172,864	1.01%	34,444	24.9%
2015	132,666	0.78%	168,884	1.00%	36,218	27.3%
2016	128,262	0.77%	166,258	0.99%	37,996	29.6%

[CAPTION] Source: Author's calculations from IPEDS and ACS

Table 4 – Undercounts of American Indian and Alaska Native Graduate Students

	IPEDS Values		Adjusted Values		Estimated Undercount	
Year	Total Enrollment	Percent Enrollment	Total Enrollment	Percent Enrollment	Total Enrollment	As Percentage of IPEDS Total
2008	14,088	0.64%	-	-	-	-
2009	16,451	0.64%	-	-	-	-
2010	15,344	0.58%	17,070	0.65%	1,726	11.2%
2011	14,582	0.55%	17,004	0.65%	2,422	16.6%
2012	14,051	0.54%	16,734	0.65%	2,683	19.2%
2013	13,578	0.53%	16,585	0.65%	3,007	22.1%
2014	13,173	0.52%	16,446	0.65%	3,273	24.8%
2015	12,916	0.51%	16,380	0.65%	3,464	26.8%
2016	12,729	0.50%	16,482	0.64%	3,753	29.5%

REFERENCES

Aliyeva, Aida, Christopher A. Cody, and Kathryn Low (2018). "History and Origins of Survey Items for the Integrated Postsecondary Education Data System (2016-17 Update).(NPEC 2018-023)". In: U.S. Department of Education. Washington, DC: National Post-secondary Education Cooperative. Retrieved October 2, 2018 from https://nces.ed.gov/pubs2018/NPEC2018023.pdf.

Byrd, W. Carson, Sandra L. Dika and Letticia T. Ramial (2013) "Who's in STEM? An exploration of race, ethnicity and citizenship reporting in a federal education dataset." In: Equity & Excellence in Education, 46.4 pp. 484-501. issn: 1066-5684. doi:10.1080/10665684.2013.838485

Broh, C Anthony and Stephen Minicucci (2008). "Racial identity and government classification: A better solution". In: Association for Institutional Research, Seattle, WA, May 28.

Burt, Larry W (1986). "Roots of the Native American urban experience: Relocation policy in the 1950s". In: American Indian Quarterly, pp. 85–99.

Burnette, Jeffrey D., Jason T. Younker and David P. Wick (2018) "The Statistical Elimination of American Indians and Alaska Natives from Higher Education". Manuscript submitted for publication

Dippie, Brian W. (1991). The vanishing American: white attitudes and U.S. Indian policy. University Press of Kansas.

Eschbach, Karl (1993). "Changing identification among American Indians and Alaska natives". In: Demography 30.4, pp. 635–652. issn: 1533-7790. doi:10.2307/2061811.

Harris, David (1994). "The 1990 Census count of American Indians: What do the numbers really mean?". In: Social Science Quarterly 75.3, pp. 580–693.

Iverson, Peter (1998). We are still here: American Indians in the twentieth century. Harlan Davidson.

Jobe, Margaret M. (2004). "Native Americans and the U.S. Census: A brief historical survey". In: Journal of Government Information 30.1, doi:10.1016/j.jgi.2001.10.001.

Liebler, Carolyn A. and Andrew Halpern-Manners (2008). "A practical approach to using multiple-race response data: A bridging method for public use microdata" In: Demography 45.1, pp. 143-155. doi:10.1352/dem.2008.0004.

Liebler, Carolyn A., Renuka Bhaskar and Sonya R. (née Rastogi) Porter (2016). "Joining, leaving, and staying in the American Indian/Alaska Native race category between 2000 and 2010" In: Demography 53.2, pp. 507-40. doi: 10.1007/s13524-016-0461-2.

Liebler, Carolyn A. and Timothy Ortyl (2014). "More than one million new American Indians in 2000: Who are they?" In: Demography 51.3, pp. 1101-30. doi: 10.1007/s13524-014-0288-7.

Norris, Tina, Paula L. Vines, and Elizabeth M. Hoeffel (2012). The American Indian and Alaska Native Population: 2010. US Department of Commerce, Economics and Statistics Administration, US Census Bureau Washington, DC.

Ogunwole, Stella U. (2002). The American Indian and Alaska Native population: 2000. U.S. Department of Commerce, Economics and Statistics Administration, U.S. Census Bureau Washington, DC.

Passel, Jeffrey S. (1976). "Provisional evaluation of the 1970 census count of American Indians". In: Demography 13.3, pp. 397–409. issn: 1533-7790. doi:10.2307/2060536.

— (1997). "The growing American Indian population, 1960–1990: Beyond demography". In: Population Research and Policy Review16.1, pp. 11–31. issn: 1573-7829. doi:10.1023/A:1005724610787.

Passel, Jeffrey S. and Patricia A. Berman (1986). "Quality of 1980 census data for American Indians". In: Social Biology 33.3-4, pp. 163–182. doi:10.1080/19485565.1986.9988637.

Rose, Richard A. (2012). American Indian race and ethnicity data: An historical analysis and comparison of NCES and NYS data for NYS school districts. Sage Graduate School.

Siegel, Jacob S and Jeffrey S Passel (1979). Coverage of the Hispanic Population of the United States in the 1970 Census: A Methodological Analysis. 82. U.S. Department of Commerce, Bureau of the Census.

Snipp, C. Matthew (1986). "Who are American Indians? Some observations about the perils and pitfalls of data for race and ethnicity". In: Population Research and Policy Review 5.3, pp. 237–252.

Thornton, Russell (2005). "Native American Demographic and Tribal Survival into the Twenty-first Century". In: American Studies46.3/4, pp. 23–38. issn: 00263079, 21536856. url: http://www.jstor.org/stable/40643888.

U.S. Census Bureau (1872). Compendium of the ninth census: 1870. U.S. Government Printing Office.

— (1937). Fifteenth census of the United States: 1930 - The Indian Population of the United States and Alaska. U.S. Government Printing Office.

— (2018). Censuses of American Indians. Retrieved September 28, 2018, from https://www.census.gov/history/www/genealogy/decennial_census_records/censuses_of_american_indians.html.

U.S. Department of Education (2007a). "Final guidance on maintaining, collecting, and reporting racial and ethnic data to the U.S. Department of Education". In: Federal Register 72.202, pp. 59266–279. Retrieved October 2, 2018 from https://title2.ed.gov/Public/TA/Guidance.pdf.

U.S. Department of Education, National Center for Education Statistics (2018). Trend Generator. Retrieved September 28, 2018, from https://nces.ed.gov/ipeds/trendgenerator/#/trend-table/2/3?trending=row&valueCode=1&rid=14&ridv=1&cid=47.

Wilkinson, Charles F. and Eric R. Biggs (1977). "The evolution of the termination policy". In: American Indian Law Review 5.1, pp. 139–184.

A. Appendix

Tables A.1 and A.2 display the results of performing the regression analysis used in earlier sections but for the period 2000-2009 and assuming there is a structural break in 2005. This is done to ensure that data prior to the mandatory implementation date follows a common path. Each table provides evidence that there is a significant difference between IPEDS and ACS data on average throughout the period. However, there is no other variable significantly different from zero. Consequently, we do not find enough evidence to reject the hypothesis that ACS control groups and IPEDS data are on a parallel path. This true for both AIAN undergraduate and graduate students.

Table A.1 – Test Results: Parallel Paths for Undergraduate Students

	Only AIAN	One-race Bridged	Any AIAN	Only AIAN	One-race Bridged	Any AIAN
IPEDS	0.3350***	0.1951***	-0.5697***	0.3347***	0.1945***	-0.5696***
	(0.0344)	(0.0373)	(0.0507)	(0.0342)	(0.0372)	(0.0505)
After 2004	0.0061	0.058	0.1195	0.0091	0.0193	0.0647
	(0.0666)	(0.0753)	(0.1090)	(0.0357)	(0.0395)	(0.0555)
IPEDS x After 2004	-0.0236	-0.0336	-0.0763	-0.0166	-0.0267	-0.071
	(0.045)	(0.0493)	(0.0661)	(0.0588)	(0.0641)	(0.0822)
IPEDS x Year 2006				0.0046	0.0047	0.0047
				(0.0583)	(0.0640)	(0.0773)
IPEDS x Year 2007				0.0069	0.0068	0.0072
				(0.0566)	(0.0624)	(0.0778)
IPEDS x Year 2008				-0.0104	-0.0106	-0.0098
				(0.0541)	(0.0602)	(0.0771)
IPEDS x Year 2009				-0.0289	-0.0285	-0.0263
				(0.0576)	(0.0665)	(0.0867)
State Fixed Effects	Y	Y	Y	Y	Y	Y
Time Fixed Effects	Y	Y	Y	N	N	N
N	1020	1020	1020	1020	1020	1020
R-squared	0.91	0.91	0.93	0.91	0.91	0.93

[CAPTION] Note: * - p<0.05, ** - p<0.01, *** - p<0.001, standard errors in parenthesis.
Source: Author's calculations from IPEDS and ACS

Table A.2 – Test Results: Parallel Paths for Graduate Students

	Only AIAN	One-race Bridged	Any AIAN	Only AIAN	One-race Bridged	Any AIAN
IPEDS	0.1731***	0.0731	-0.5496***	0.1728***	0.0727	-0.5502***
	(0.0326)	(0.0386)	(0.0642)	(0.0329)	(0.0396)	(0.0654)
After 2005	-0.0287	0.005	0.0094	0.0052	0.0085	0.0761
	(0.0709)	(0.0853)	(0.1552)	0.0388)	(0.0466)	(0.0715)
IPEDS x After 2005	0.0134	0.013	-0.047	-0.0019	-0.0043	-0.0688
	(0.0412)	(0.0482)	(0.0761)	(0.0445)	(0.0526)	(0.0871)
IPEDS x Year 2006				0.0221	0.0228	0.0246
				(0.0266)	(0.0321)	(0.0633)
IPEDS x Year 2007				0.0389	0.0408	0.0441
				(0.0374)	(0.0404)	(0.0660)
IPEDS x Year 2008				0.009	0.0127	0.0186
				(0.0275)	(0.0321)	(0.0616)
IPEDS x Year 2009				0.0106	0.0147	0.0213
				(0.0287)	(0.0342)	(0.0638)
State Fixed Effects	Y	Y	Y	Y	Y	Y
Time Fixed Effects	Y	Y	Y	N	N	N
N	1020	1020	1020	1020	1020	1020
R-squared	0.74	0.74	0.71	0.74	0.73	0.71

[CAPTION] Note: * - p<0.05, ** - p<0.01, *** - p<0.001, standard errors in parenthesis.
Source: Author's calculations from IPEDS and ACS