

A County Level Analysis of High School Dropout Rates Nationwide

Brandon Hallstrand

Undergraduate Student, Business Administration

Kunjan Upadhyay

Undergraduate Student, Business Administration

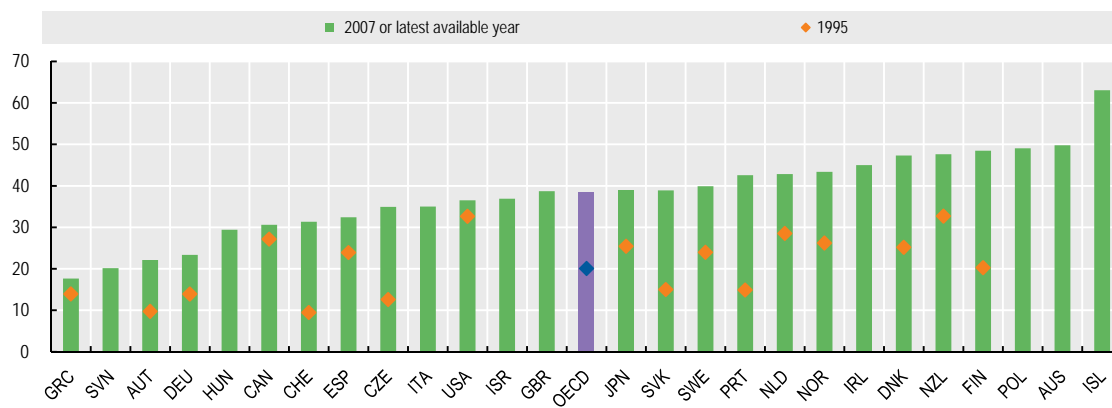
University of Wisconsin-Stout

A County Level Analysis of High School Dropout Rates Nationwide

In the past 20 years, education in the United States has seen many improvements, including the desegregation of public schools in the 1960s and increases in female graduation and education rates. However, this upward trend in national school attainment and graduation rates hides tremendous educational attainment disparities among U.S. regions, genders, and races. According to Laird, DeBell, and Chapman (2007), the national dropout rate declined substantially from 14% in 1980 to 8% in 2008, while the gaps between racial groups have remained. Even in 2008, the dropout rates by race ranged from 4.4% to 18.3% (Laird, DeBell, & Chapman, 2007). With the male dropout rate in 1980 being 15.1% and the female dropout rate 13.1% and the 2008 rates of 8.5% and 7.5% respectively, it is evident that educational attainment has improved for both genders. However, one interesting fact between the 1980 and 2008 rates is that males continue to show a higher dropout rate compared to females.

At the international level, the general educational system of the United States is being challenged as well. According to the Organization for Economic Co-operation and Development (OECD) (2010), the United States is currently ranked 16th amongst the 26 listed nations, a change from fifteen years ago when the United States were ranked 1st.

Figure 1: Educational Attainment Rates within OECD Countries 2010



Source: Organization for Economic Cooperation and Development, 2010.

Prior studies investigated high school dropout rates and educational attainment in specific areas such as the central region of the United States or Chicago public schools (e.g., Randel, Moore, & Blair, 2008; Allensworth & Easton, 2001). We feel that these previous studies do not sufficiently address the high school dropout rates within the United States as a whole. This paper will contribute to the literature by analyzing the high school dropout rates across the United States using county-level data. We will use economic as well as social factors to explain the prevalent high school dropout rates in the United States. Specifically, we intend to investigate how local government expenditures on education and per capita personal income affect county-level high school dropout rates in the United States. At the same time, we look at the effect on high school dropout rates of the male to female ratio, different racial groups, and geographic location of a county. Everything else the same, the more money one spends on a good or a service, the higher the quality of the product or service one expects to receive. Thus, we expect that high school dropout rates should be lower in counties with higher local government expenditures on education. We also expect counties with higher share of per capita personal income to have lower dropout rates. We anticipate this because higher per capita income means lower poverty rates, and poverty is a motivating factor in dropping out of school. As far as the social factors are concerned, their effects on dropout rates could go either way or may have no effect because it is very difficult to predict human behavior.

Our analysis of dropout rates is as timely as a 2010 report by the OECD illustrating how the United States is facing huge challenges in terms of its educational system as compared to the rest of the listed economies. Moreover, in the face of the increasing and fierce international competitiveness of the labor force as a result of globalization, the topic related to education cannot be investigated enough. The aim of this paper is to explain the high school dropout rates using racial disparities across the United States at the county level. Other control variables, as

used by previous studies, are also included in our paper.

Methodological Framework

We use multiple regression models to quantify and analyze the effects that the economic, social, and geographic variables have on high school dropout rates using county-level data collected from the U.S. Census. Our methodological framework includes three models which attempt to explain the high dropout rates across the United States. The first two models are comprised of cross-sectional regressions of data for 1990 and 2000 and the third being a panel regression of the two cross-sectional data sets with the added variable of time.¹ We feel that having the three separate models helps add to the validity of the study and allows a controlled comparison of each regression. Shown below are our starting models; the only difference being that Model 2, the panel model, contains a year variable used as a dummy variable to account for any factors that could affect the high school dropout rates not included in Model 1. The year dummy variable is set to take the value of one for year 2000 and zero otherwise.

Model 1: Cross-sectional Model for 1990 and 2000

$$\begin{aligned} \text{Dropout rate} = & a_0 + a_1 \text{PPI} + a_2 \text{LgSpnd} + a_3 \text{ChgSpnd} + a_4 \text{M/F} + \\ & a_5 \text{White} + a_6 \text{Asian} + a_7 \text{Black} + a_8 \text{AM Ind} + a_9 \text{Other} + a_{10} \text{Midwest} + \\ & a_{11} \text{South} + a_{12} \text{West} + \varepsilon_i \end{aligned}$$

Note: (PPI)= Per capita Personal Income, (LgSpnd)= Local Government Expenditures for Education per Child, (ChgSpnd)= Change in Local Government Expenditure for Education per Child, (M/F)= Ratio of Male population to Female Population. Whites, Asians Blacks and (AM Ind)= American Indians are the major races we will be following. Whereas other category combines race such as Hispanics and others; this is due to the difficulties in separation on Census data. Finally, ε_i is a random error that is assumed to have mean of zero and variance σ^2 . That is to say that there is an unknown variance which is relatively unimportant due to the assumed mean of zero.

¹ Cross-sectional data sets are data sets which consist of variables in one point in time whereas panel data sets allow for the combination of multiple cross-sectional data sets from different points in time.

Model 2: Panel Data Model

$$\begin{aligned} \text{Dropout rate} = & a_0 + a_1 PPI + a_2 LgSpnd + a_3 ChgSpnd + a_4 M/F + a_5 White + \\ & a_6 Asian + a_7 Black + a_8 AM Ind + a_9 Other + a_{10} Midwest + a_{11} South + \\ & a_{12} West + a_{13} Y2000 + \varepsilon_{it} \end{aligned}$$

All the variables have the same meaning as defined in the above cross-sectional model. And Y2000 stands for year 2000 dummy variable. It equals 1 if the year is 2000, and zero if 1990. The introduction of the year dummy variable allows us to track for changes in the variables that have occurred between 1990 and 2000.

Each equation of the preceding models was assessed using the best subset function of the statistical software package *Minitab 16*. The best subsets output was then analyzed primarily on the basis of the Mallows' Cp value, which helps to choose a model that is highly explanatory with a minimal number of variables. The best model is then chosen and estimated using the Ordinary Least Squares (OLS) regression method². The following are the end models we have chosen (our preferred models) based on the Mallows' Cp value test.

Model 3: Cross-sectional for 1990 and 2000

$$\begin{aligned} \text{Dropout rate} = & a_0 + a_1 PPI + a_2 LgSpnd + a_3 M/F + a_4 White + a_5 Asian + \\ & a_6 Black + a_7 AM Ind + a_8 Other + a_9 Midwest + a_{10} South + \varepsilon_{it} \end{aligned}$$

Model 4: Panel

$$\begin{aligned} \text{Dropout rate} \\ = & a_0 + a_1 PPI + a_2 LgSpnd + a_3 M/F + a_4 White + a_5 Asian + a_6 Black \\ & + a_7 AM Ind + a_8 Other + a_9 Midwest + a_{10} South + a_{11} Y2000 + \varepsilon_{it} \end{aligned}$$

² Interested reader can refer to any introductory statistics textbook for a detailed discussion about OLS regression method.

Data and Descriptive Statistics

Dependent Variable

In economics, the dependent variable is the variable that the model seeks to explain using a set of other variables called explanatory variables.³ Our dependent variable is the dropout rate calculated using U.S. Census Bureau data. This rate was calculated by finding the percentage of persons ages 16 to 19 who are not in school and have not graduated from high school. This is a status dropout rate, meaning that it is a measure of the percentage of a given age group that has already dropped out of high school. As such, this rate is an understatement of the true dropout rate⁴, but it gives a good understanding of the characteristics of the dropout population (Laird, DeBell, & Chapman, 2007).

Independent Variables

The independent variables, also called the right hand-side variables or the explanatory variables, are the variables used to explain changes in the dependent variables. As far as modeling an economic behavior is concerned, researchers must choose between a simple regression model and a multiple regression model. A simple regression model is a model with one dependent variable and one explanatory variable. A multiple regression model has one dependent variable and two or more right hand-side variables. For the purpose of this paper, we use a multiple regression model. As mentioned above, our explanatory variables include a set of social, economic, and geographic factors⁵.

The social independent variables include the ratio of male to female, the percentage of total population, that is, White, Black, Asian, Hispanic, and Other. It is evident in reading

³ We discuss the meaning of the explanatory variables in the section that follows.

⁴ See Grunewald and Chung (2011) for the discussion about how the high school dropout rate as it is currently measured raises some questions.

⁵ In this paper, we interchangeably use the terms “variable” and “factor”.

through papers on the topic related to education that race and gender are correlated to high school dropout rates in the United States; more specifically, we are looking to build upon the work of Allensworth and Easton (2001). In their paper, the authors show graphs of the dropout rates of cohorts of students by either race, gender, or both, resulting in quite interesting findings. We are interested in the profound differences in the performances of the races, especially in the fact that Asians are shown to be far apart and below in dropout rates relative to other races (Allensworth & Easton, 2001). Therefore, we would like to expand upon Allensworth and Easton's (2001) analysis, which was limited to Chicago public school data, by using a country-wide data set with similar race categories. The race categories which we incorporated are all "population of one race" categories for consistency. That is to say that each individual is considered to be of one race and one race only within the data set we used. The racial categories are Asians and Pacific Islanders, Whites, Blacks, Native Americans, and Other. This collection of races was chosen by looking at the work set forth in multiple papers on the subject (e.g., Allensworth & Easton, 2001; Randall, 1997). It is also important to note that the Hispanic race category is included in the Other category due to the difficulties of differentiating between the two (i.e., Hispanic and Other) in the Census data. It seems that a substantial portion of the Other category is made up of Hispanics not reported elsewhere.

The economic variables included in this paper are the Per Capita Personal Income (PPI) and the Local Government Expenditures for Education per Child (LgSpnd). We feel that these variables are likely to influence the educational outcomes of high school students, and as Mayer (2001) states, "Much less research has been done on the consequences of inequality than its cause" (p. 1). It is important to note, however, that the Local Government Direct Expenditures for Education variable has been used as both a three-year lagged variable (LgSpnd) and a percent-change variable (ChgSpnd). The time lag was necessary due to the limited number of

years for which there was available data. It is furthermore important to note that in order to limit the influence of changes in expenditure during our period of study, these dollar-denominated variables have been adjusted for inflation using the Consumer Price Index (CPI) year average value for the respective year, with 1990 being the base year (U.S. Department of Labor, 2010).

To account for possible effects of geography in explaining the high school dropout rates, we included four regional variables using the regions defined by the U.S. Census Bureau. In effect, the U.S. Census Bureau divided the United States into four regions: Northeast, South, West, and Midwest. Those four regions are included in our models as dummy variables. For example, a county located in the Midwest takes a value of 1 for the variable⁶ Midwest, and a value of zero for the remaining three regions. Including all the four regions in our models would cause a multicollinearity⁷ problem when estimating the models. According to Gujarati (1988), “there are various ways to solve this problem, the simplest one is to assign dummy variables and use only m-1 variables if there are m levels or classes of the qualitative data” (p. 436). To address the multicollinearity issue, we dropped the Northeast region from our models.

A final variable we chose to incorporate is a time variable, represented as a dummy variable in our panel regression. To form the panel data models, we stocked our data by county for both 1990 and 2000. And then, we incorporate a time dummy variable which takes a value of one for the year 2000 and zero for the year 1990.

The descriptive statistics for each of the variables may be found in the tables below. Table 1 shows the statistics for the all inclusive model; Table 2 shows the statistics for the 1990 model, and Table 3 shows the statistics for the 2000 model.

⁶ Please note that we are using the regions as variables in our study.

⁷ Multicollinearity means the existence of a linear relationship among some or all explanatory variables of a regression model (Gujrati, 1988).

Table 1: 1990 Descriptive Statistics

Variable	N	Mean	StDev	Minimum	Maximum
Drop	3040	11.058	5.402	0.602	51.064
PPI	3040	15326	3567	5479	50230
LgSpnd	3040	4.2647	1.4709	0	26.0506
ChgSpnd	3040	0.1161	0.7406	-1	39.5501
M/F	3040	96.456	7.415	81.055	211.806
White	3040	84.418	18.263	2.584	99.845
Asian	3040	0.7102	2.5387	0	62.9562
Black	3040	8.622	14.303	0	86.236
Am Ind	3040	1.722	7.153	0	94.668
Other	3040	4.527	11.148	0	97.216

Note. Data calculated from U.S. Census Bureau

Table 2: 2000 Discriptive Statistics

Variable	N	Mean	StDev	Minimum	Maximum
Drop	3040	9.7201	5.129	0.3145	57.9785
PPI	3040	17554	4383	5685	65100
LgSpnd	3040	5.1474	2.2402	0	91.4449
ChgSpnd	3040	0.20347	0.25865	-1	4.3085
M/F	3040	98.591	8.807	74.1	205.4
White	3040	81.234	19.05	2	99.5
Asian	3040	0.8917	2.3884	0	54.9

Black	3040	8.804	14.472	0	86.5
Am Ind	3040	1.868	7.44	0	94.2
Other	3040	6.238	12.094	0.1	97.5

Note. Data calculated from U.S. Census Bureau

Table 3: Panel Descriptive Statistics

Variable	N	Mean	StDev	Minimum	Maximum
Drop	6080	10.389	5.309	0.314	57.979
PPI	6080	16440	4148	5479	65100
LgSpnd	6080	2.1324	2.3726	0	26.0506
ChgSpnd	6080	0.0581	0.5269	-1	39.5501
M/F	6080	97.523	8.21	74.1	211.806
White	6080	82.826	18.727	2	99.845
Asian	6080	0.801	2.4662	0	62.9562
Black	6080	8.713	14.387	0	86.5
Am Ind	6080	1.7952	7.2975	0	94.6677
Other	6080	5.383	11.661	0	97.5

Note. Data calculated from U.S. Census Bureau

It becomes apparent from Table 1, Table 2, or Table 3 that racial composition varies greatly from one county to another. As with the racial groups, there is also a large differentiation in high school dropout rates from one county to another across the United States, reinforcing our hypothesis that difference in high school dropout rates across the United States may be well explained using racial composition within a locality. This hypothesis is better tested using our

estimation results presented in the section that follows.

Regression Analyses

Estimation Results using 1990 Data

Unfortunately, problems with the U.S. Census data caused us to drop the Hispanic data and instead use the racial category Other as a proxy. Even in doing so, the racial category Other was highly correlated with the rest of the racial variables used in the 1990 data set. Therefore, Other was dropped from the 1990 equation altogether. The estimated regression from Model 1, using 1990 data, is expressed as follows:

Equation 1: 1990 Regression

Dropout Rate

$$\begin{aligned} = & 15.307 - 0.00018014 PPI - 0.42925 LgSpnd + 0.0439 ChgSpnd \\ & + 0.03249 M/F - 0.037947 White - 0.08853 Asian - 0.02859 Black \\ & + 0.01446 Am Ind - 1.7027 Midwest + 2.6818 South + 0.3386 West \end{aligned}$$

Estimation Results using 2000 Data

The regression output of Model 1 using data from 2000 yields the following:

Equation 2: 2000 Regression

Dropout Rate

$$\begin{aligned} = & -17.09 - 0.0001455 PPI - 0.15669 LgSpnd + 0.2318 ChgSpnd \\ & + 0.047978 M/F + 0.2384 White + 0.223 Asian + 0.3014 Black \\ & + 0.3014 Am Ind + 0.2856 Other - 0.5499 Midwest + 2.0526 South \\ & + 0.3426 West \end{aligned}$$

All Inclusive Regression

As mentioned earlier in the paper, the main difference between Model 1 (using either data from 1990 or data from 2000) and Model 2, which we call the all inclusive model or the paned data model, is the inclusion of the time dummy variable. The estimated equation is as follows:

Equation 3: Panel Regression

Dropout Rate

$$\begin{aligned}
 = & -13.358 - 0.00016309 PPI - 0.44858 LgSpnd + 0.0456 ChgSpnd \\
 & + 0.040742 M/F + 0.23458 White + 0.2038 Asian + 0.27198 Black \\
 & + 0.30306 Am Ind + 0.27671 Other - 1.0407 Midwest + 2.4709 South \\
 & + 0.3963 West - 2.8263 Y2000
 \end{aligned}$$

For ease of comparison, we have summarized the results from all the above three estimated models into a table which includes the estimated coefficients and their T-statistics, and the R-Squared for each model.

Table 4: Regression Analysis (using Models 1 and 2)

Predictor	1990	2000	1990-2000
Constant	15.307 (9.73)*	-17.09 (-1.68)***	-13.358 (-1.51)
PPI	-0.00018014 (-6.31)*	-0.0001455 (-6.63)*	-0.00016309 (-9.4)*
LgSpnd	-0.42925 (-6.32)*	-0.15669 (-3.96)*	-0.44858 (-7.27)*
ChgSpnd	0.0439 (0.38)	0.2318 (-0.69)	0.0456 (0.4)
M/F	0.03249 (2.59)*	0.047978 (4.88)*	0.040742 (5.24)*
White	-0.037947 (-4.58)*	0.2384 (2.32)**	0.23458 (2.66)*
Asian	-0.08853	0.223	0.2038

	(-2.31)**	(1.64)***	(1.99)**	
Black	-0.02859	0.3014	0.27198	
	(-2.85)*	(2.98)*	(3.1)*	
Am Ind	0.01446	0.3207	0.30306	
	(1)	(3.03)*	(3.37)*	
Other	N/A	0.2856	0.27671	
	N/A	(2.84)*	(3.16)*	
Midwest	-1.7027	-0.5499	-1.0407	
	(-4.63)*	(-1.57)	(-4.09)*	
South	2.6818	2.0526	2.4709	
	(6.97)*	(5.62)*	(9.37)*	
West	0.3386	0.3426	0.3963	
	(0.8)	(-0.84)	(1.35)	
Y2000	N/A	N/A	-2.8263	
	N/A	N/A	(-9.01)*	

	1990	2000	1990- 2000
R-Sq	22.90%	21.50%	22.60%
R-Sq(Adj)	22.60%	21.20%	22.50%

Note:

* : denotes the variable is statistically significant at 1%

** : denotes the variable is statistically significant at 5%

***: denotes the variable is statistically significant at 10%

We chose not to discuss the results from Table 4 because the models expressed in the form of equations 1, 2, and 3 are not our preferred models. We did not choose these models because of the relative weakness of the variables in the models. In the table, there are many variables which are not significant to the equation. The significance is to some degree denoted by the asterisks in the table. Variables lacking asterisks are deemed to be insignificant at the ten percent level.

As mentioned earlier in the paper, the choice of our preferred models was done by running the subset test which allows us to improve upon our model. Thus, we move on to Model 3 and Model 4.

Our Preferred End Regression Equations

The following equations are the estimated Models 3 and 4, with equations 4 and 5 being the estimated Model 3 using the 1990 and the 2000 data respectively, and equation 6 being the

estimated panel data Model 4.

Equation 4: 1990 End Equation

Dropout Rate

$$\begin{aligned}
 &= 15.5 - 0.00018 PPI - 0.43247 LgSpnd + 0.3425 M/F \\
 &\quad - 0.3888 White - 0.8688 Asian - 0.0295 Black + 0.01461 Am Ind \\
 &\quad - 1.9139 Midwest + 2.466 South
 \end{aligned}$$

Equation 5: 2000 End Equation

Dropout Rate

$$\begin{aligned}
 &= -15.86 - 0.0001472 PPI - 0.16655 LgSpnd + 0.048946 \frac{M}{F} \\
 &\quad + 0.2284 White + 0.2122 Asian + 0.2919 Black + 0.3123 Am Ind \\
 &\quad + 0.277 Other - 0.7574 Midwest + 1.8143 South
 \end{aligned}$$

Equation 6: Panel End Equation

Dropout Rate

$$\begin{aligned}
 &= -12.234 - 0.0001663 PPI - .45195 LgSpnd + 0.4223 \frac{M}{F} \\
 &\quad + 0.22502 White + 0.1965 Asian + 0.26239 Black + 0.2947 Am Ind \\
 &\quad + 0.26883 Other - 1.2861 Midwest + 2.2185 South - 2.8538 Y2000
 \end{aligned}$$

See Table 5 for a summary of the values from Equations 4, 5 and 6.

Table 5: Regression Analysis Adjusted (using Models 3 and 4)

Predictor	1990	2000	1990-2000
Constant	15.5 (9.97)*	-15.86 (-1.57)	-12.234 (-1.39)
PPI	-0.00018 (-6.46)*	-0.0001472 (-6.77)*	-0.00016627 (-9.68)*
LgSpnd	-0.43247 (-6.39)*	-0.16655 (-4.38)*	-0.45195 (-7.35)*

M/F	0.03425 (2.78)*	0.048946 (5.02)*	0.04223 (5.48)*
White	-0.03888 (-4.74)*	0.2284 (2.24)**	0.22502 (2.56)*
Asian	-0.08688 (-2.28)**	0.2122 (1.57)	0.1965 (1.92)**
Black	-0.0295 (-2.96)*	0.2919 (2.91)*	0.26239 (3)*
Am Ind	0.01461 (1.01)	0.3123 (2.97)*	0.2947 (3.28)*
Other	N/A N/A	0.277 (2.77)*	0.26833 (3.07)*
Midwest	-1.9139 (-7.38)*	-0.7574 (-3.07)*	-1.2861 (-7.18)*
South	2.466 (8.95)*	1.8143 (7.15)*	2.2185 (11.88)*
Y2000	N/A N/A	N/A N/A	-2.8538 (-9.16)*

	1990	2000	1990- 2000
R-Sq	22.90%	21.50%	22.60%
R- Sq(Adj)	22.70%	21.20%	22.50%

Note:

* : denotes the variable is statistically significant at 1%

** : denotes the variable is statistically significant at 5%

***: denotes the variable is statistically significant at 10%

Relative to Models 1 and 2, Models 3 and 4 show great improvements in the significance of individual variables with no adjustment to the R-Squared terms. Most of the variables are significant at the one percentage level. This means that we can be much more confident of the impact each individual variable has on dropout rates.

There is consistency across the models with respect to several of our variables as far as their signs and magnitudes are concerned. PPI and LgSpnd per student are both negatively correlated with dropout rates. That means counties with higher per capita personal income have

seen a decline in the dropout rates of their high school students for the period between 1990 and 2000, all else equal. Also, the increases in the local government educational expenditures are correlated with lower dropout rates; although, the effects of educational expenditures on dropout rates seem to kick in with a time lag (three years in this study). Specifically, a 1% increase in LgSpnd is associated with a 0.45% reduction in high school dropout rates three years later. This negative association between high school dropout rates and LgSpnd appears in the 1990 and the 2000 data as well. These results suggest that economic factors play a role in explaining the high school dropout rates.

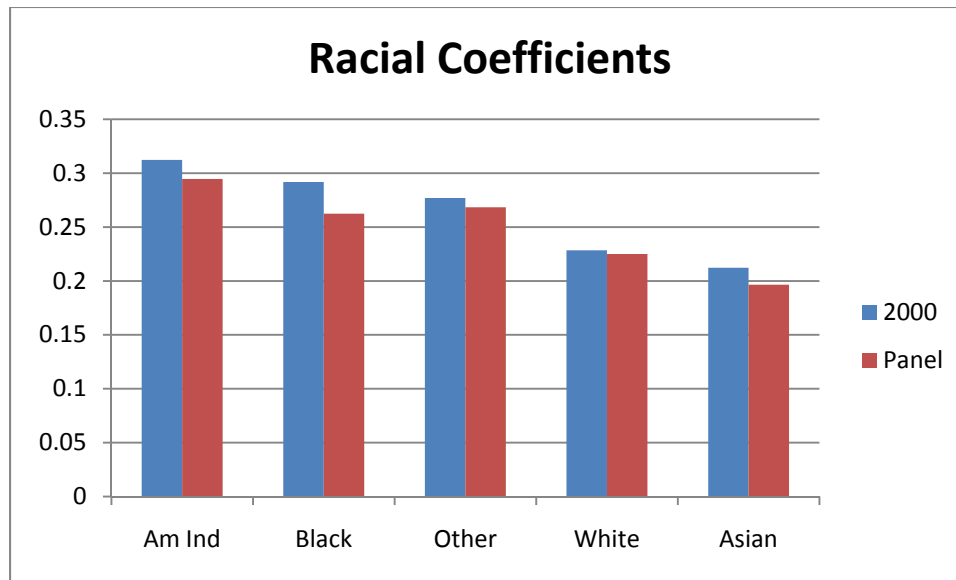
Also, there does seem to be a regional dynamic involved because the Midwest is consistently shown as having an inverse relationship with dropout rates, while the South has a positive relationship. In other words, someone living in the southern parts of the United States is more likely to drop out of high school than someone who lives in the Midwest. This result reinforces the relationship between per capita personal income and dropout rates. Indeed, previous studies have documented that the southern United States have a higher number of people living below the poverty level compared to the Midwest region.

Another important finding is that the male to female coefficient consistently shows males have higher dropout rates compared to females, all else equal. The gender difference in high school dropout seems to widen over time. Namely, the estimated coefficient of M/F has increased from 0.034 in 1990 to 0.049 in 2000.

One startling result is in regard to racial categories. The sign of the coefficients on different race groups, with the exception of the Native Americans, has turned from negative in 1990 to positive in 2000. The results suggest that, all else equal, in the 1990s and before, Native Americans were most likely to drop from high school compared to either Whites, Blacks, or Asians. However, between 1990 and 2000, race alone did not explain the likelihood of dropping

out of high school. Nevertheless, Blacks and Native Americans have shown a greater tendency to drop out of high school. A graphical illustration of the coefficient estimates of the different race groups is shown in Figure 2 below.

Figure 2: Racial Coefficients by Model



Our results showing lower dropout rates attributed to Asian populations are in line with the findings in the Chicago public schools by Allensworth and Easton (2001).

Lastly, the time trend was found to be statistically significant and showed an inverse relationship with dropout rates. Specifically, dropout rates have decreased between 1990 and 2000, and there is reason to hope that the U.S. dropout rates will continue to decline, but this hope is contrary to other research, suggesting that dropout rates will soon begin trending upward (Laird, DeBell, & Chapman, 2007).

Conclusion

The findings of this paper have some policy implications. Based on our results, we can conclude Local Educational Spending is a variable that has a large and positive impact in decreasing dropout rates. Assuming causality between the variables, one dollar per school child

at the 1990 price level is associated with a 0.16-0.45% likelihood of reducing high school dropout rates three years forward. The decrease in the dropout rates between 1990 and 2000 can also be partially attributed to an increase in the Per Capita Personal Income. In addition, we can assert that the findings of Allensworth and Easton (2001) in regard to the Chicago public schools also seem to hold true in the United States as a whole. We can also show that the Asian population is sustaining lower dropout rates compared to other races in the United States.

Lastly, our findings point to regional disparities in the United States that do not seem to be attributed to other variables, such as personal income alone. The indication is that something other than income, spending, or race may cause dropout rates in the South to be significantly higher than in the rest of the United States. At the same time, the Midwest seems to be experiencing the opposite phenomenon as its dropout rates are considerably lower.

Future Work

While our findings have successfully extended some of the more localized research we reviewed, further questions remain. Thus, there is an opportunity for future research into the nature of the regional disparity between the Midwest and the South. Income, race, and local spending do play a part in explaining the differences, but there are still unnamed variables accounting for the rest of the difference. These could perhaps be ideological differences, state and federal funding differences, or any number of other factors which we did not explore. In the future, we would like to explore the reasons behind the lower dropout rates of Asian-Americans compared to the rest of the population.

References

- Allensworth, E., & Easton, J.Q. (2001). Calculating a cohort dropout rate for the Chicago public schools. *Consortium on Chicago School Research*, Retrieved from <http://ccsr.uchicago.edu>
- Grunewald, R., & Chung, W. (2011, January). Another look at measuring high school graduation rates. *Fedgazette*, 12.
- Gujrati, D.N. (1988). *Basic Econometrics*. New York: McGraw-Hill.
- Mayer, S.E. (2001). How did the increase in economic inequality between 1970 and 1990 affect children's educational attainment? *American Journal of Sociology*, 107(1). Retrieved from <http://professorreed.com>
- Laird, J., DeBell, M., & Chapman, C. (2007). Dropout rates in the United States: 2004 (nces 2007-024). *U.S. Department of Education. Washington, DC: National Center for Education Statistics*. Retrieved from <http://nces.ed.gov>
- Mallows, C., L. (1973). Some comments on Cp. *Technometrics*, 15(4). Retrieved from www.jstor.org
- Randall, S.C. (1997). The effect of school dropout rates on estimates of adolescent substance use among three racial/ethnic groups. *American Journal of Public Health*, 87(1), Retrieved from <http://ajph.aphapublications.org>
- Randel, B., Moore, L., & Blair, P. (2008). High school dropout and graduation rates in the central region. *Issues & Answers Report*, 40. Retrieved from <http://ies.ed.gov>
- U.S. Census Bureau, (2001). USA counties Retrieved from <http://censtats.census.gov>
- U.S. Department of Education Institute of Education Sciences, (2010). *The dropout rates of high school students*. Retrieved from <http://nces.ed.gov>

U.S. Department Of Labor, Bureau of Labor Statistics. (2010). *Consumer Price Index (CPI)*.

U.S. city average, Retrieved from <http://ftp.bls.gov>