The Built Environment, Travel and Income Inequality

Reggie Arkell

Southern New Hampshire University

ECO 700 Applied Economics Capstone

April 8, 2018

Abstract

Research is conducted to determine if either weighted population density (WPD) or per capita vehicle miles traveled (PCVMT) is associated with income inequality as measured by the Gini index (GINI) in U.S. metropolitan statistical areas for the periods of 1990, 2000 and 2010/2015. The findings are that, on average, a 1 percent increase in PCVMT is correlated with a 0.023 percent increase in GINI. The relationship is statistically significant but relatively small compared to other controlling variables. The qualification is that the association is evident using generalized least squares (GLS) regression analysis of panel data in a pooled cross-section. Further, the relationship only occurs in certain mid-range quantiles of the data distribution. The relationship is not apparent using ordinary least squares (OLS) cross-sectional regression analysis for each of the single years of data or changes between the latter two periods. An association was not identified between WPD and GINI. The OLS and GLS analysis did find that a metric highly correlated with WPD, the proportion of population in poor or affluent neighborhoods, does have a statistically significant positive relationship with GINI. The magnitude varies by model. The overall findings lend moderate support to the spatial mismatch hypothesis which states that the friction of distance between residential areas, particularly segregated and low-skilled minority populations, and employment opportunities negatively impacts personal income.

Keywords: weighted population density, per capita vehicle miles traveled, Gini index

The Built Environment, Travel and Income Inequality

As outlined by Autor, Katz, and Kearney (2006), extensive research exists on increasing trends of income inequality in the U.S. over the last several decades. One reason is the rising demand for higher skills due in part to technological advances countered by lower growth in persons with college educations. Other factors are decreasing purchasing power due to stagnant minimum wage rates and reductions in union memberships. Another influence is the lack of middle-income job growth and the rise in both high and low-income employment, otherwise known as polarization of the labor market (Autor, Katz, & Kearney, 2006). Hatch and Rigby (2015) cite studies identifying disproportionate increases in top wage rates and rent-seeking along with shortcomings in state/federal market and redistribution policies. The U.S. Census Bureau documents a gradual rise in the Gini index (GINI) of income inequality over about the past 50 years from a low of 0.351 in 1968 to a high of 0.467 in 2013 (U.S. Census Bureau, 2016). Thus, rising income inequality can be considered an indicator of increasing disparities in quality of life.

Study Purpose and Expected Benefits

The negative relationship between standard population density (SPD) and per capita vehicle miles traveled (PCVMT) is well documented (Cervero & Murakami, 2010). U.S. urban area SPD has been declining for decades due mainly to rising incomes and falling transportation costs (Kim, 2007). Those businesses and households with the financial means take advantage of the evolution in transport and communications by moving outward from urban cores resulting in residential income segregation which is associated with increasing income inequality (Mayer, 2001). Thus, spatial mismatches are created between low-income residential areas and employment locations resulting in larger income disparities (Ewing, Hamidi, Grace, & Wei, 2016). These associations have been measured via the use of a somewhat complex compactness index with metrics of development density, land use mix, activity centering and street accessibility (Ewing & Hamidi, 2014).

More simplified metrics of the built environment that may provide insight to income inequality are weighted population density (WPD) and vehicle miles traveled (VMT). WPD is calculated by determining SPD at census tract levels, weighting each by its proportion of the metropolitan statistical area (MSA) population and adding the results. WPD is a superior metric over SPD in measuring clustering of the population (Eidlin, 2010). Further clarity on the elements associated with income inequality can help to inform decisionmakers on the appropriate policies necessary to address the problem. In turn, insight is provided on the predicted impacts from planning methodologies such as urban growth boundaries limiting land consumption, retention of farmland, natural areas and open space, population/employment densities at neighborhood levels, expressway expansions, roadway pricing, and affordable/mixed-income housing among others. These planning techniques can then be applied as appropriate to reduce income inequality while addressing other social objectives and environmental externalities relative to urban development including, air/water emissions, loss of natural habitat, energy consumption and global warming.

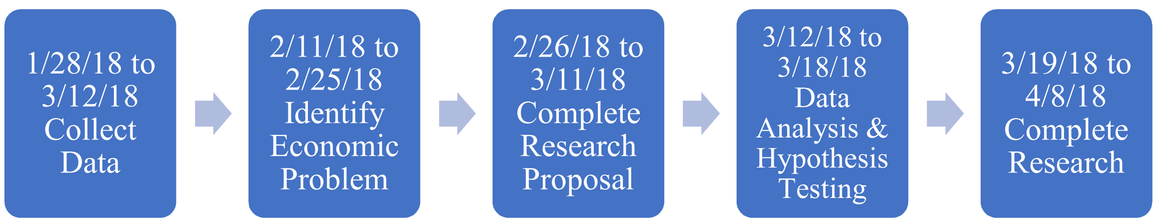
Relevant Audiences

The specific audience targeted is land use planners, stakeholders, and local/regional government officials involved in land use and transportation decision-making. The public is also targeted as they are essentially represented by these individuals and impacted by the policy choices. In particular, those citizens with low incomes are an important audience for the research as the knowledge can provide empowerment to overcome discriminatory practices and policies.

The analysis is presented using the necessary econometric technical terminology to demonstrate the appropriate use of statistical methodologies. At the same time, plain language explanations are provided to maximize comprehension by the reading audience.

Cost of Needed Research and Timeline

The research does not involve any in-person or written survey work due to time constraints. Therefore, monetary costs are negligible. Extensive time is needed to allow for research, compilation, organization and formatting of data. Below is a timeline of the anticipated schedule for the research project.



Economic Theory

The spatial mismatch hypothesis is an economic theory stating that restrictions on black residential choice together with employment disbursement is responsible for high joblessness and low income for those minority populations (Kain, 1994). This hypothesis has roots in groundbreaking econometric analyses of workers in Detroit and Chicago using 1952 and 1956 data of workplace and residential locations. The findings were that racial bias in these housing markets greatly restricted residential choices of black households which impacted the spatial distribution of black employment and caused higher unemployment for this segment of the population. Additionally, further employment dispersal absent reforms in housing discrimination would be expected to exacerbate the problem. Public transportation solutions were devised in the late 1960s to address the dilemma, but it became clear that they were not cost-effective. Various subsequent studies of the spatial mismatch hypothesis either concurred, refuted, or found some validity in the theory with other more critical factors (Kain, 2004).

Holzer (1991) reviewed the empirical literature on the spatial mismatch hypothesis and found that the theory is valid for explaining disparities in black/white employment rates (as opposed to income) but the degree of these impacts is not certain. Most studies have relied upon cross-sectional data which limits predictive ability over time (Holzer, 1991). Andersson, Haltiwanger, Kutzbach, Pollakowski, and Weinberg (2014) allude to the different accessibility measures used in the relative studies such as commute times, distance, and car/transit availability. More recent longitudinal research of nine Great Lakes Region MSAs focusing on low-income displaced workers during 2000-2005 found support for the spatial mismatch hypothesis as improved job accessibility measured by reduced commute time is consistent with lower duration of unemployment. More specifically, an improvement from the 25th to 75th percentile of employment accessibility is related to a 4.2 percent decrease in the amount of time taken to obtain a new job and a 7.0 percent reduction in obtaining a new position with at least 90 percent of the previous job earnings. Further, unemployed blacks are about 71 and 35 percent more susceptible to work accessibility than white job seekers for these respective hiring metrics (Andersson, Haltiwanger, Kutzbach, Pollakowski, & Weinberg, 2014).

Review of Other Literature

President Lyndon Johnson formed the National Advisory Commission on Civil Disorders (Kerner Commission) to investigate the turbulent race riots of the 1960s. Glenn (1968) summarizes the findings. The Kerner Report, published in 1968, concluded that urban problems were largely attributable to white racism which resulted in expanding concentrations of inner city blacks and the ghettos in which they reside. More specifically, discrimination was widespread in the areas of housing and credit, law enforcement and sentencing, employment, and consumer practices. Growing resentment by blacks led further to the explosive nature of their responses.

The Kerner Report recommendations included a national supplemental income program, a welfare program primarily funded by the federal government, massive expansion of low income housing, and educational initiatives. On the one hand, these conclusions and recommendations had been espoused by many for years beforehand. Conversely, polling found that a large proportion of Americans disagreed with several of the findings. As a result, there did not appear to be consensus on developing a roadmap to address the problem (Glenn,1968). Recent analysis by Jones, Schmitt and Wilson (2018) performed on the 50th anniversary of the Kerner Report finds that black Americans have made progress with improved high school/college graduation rates in addition to absolute gains in income and health. However, blacks still trail whites substantially in these areas. Further, virtually no progress has been made regarding black rates for unemployment, homeownership and incarceration (Jones, Schmitt, & Wilson, 2018).

Deaton (2003) cites various studies which generally conclude that income inequality is correlated with reduced social cohesion and higher negative health impacts/mortality. There is disagreement amongst economists as to whether there is causation from income to health or if it is in the opposite direction. Deaton (2003) concludes that health status is more directly associated with poverty than income inequality, and that there is no detrimental effect to escalating incomes for the rich if minimal levels of income are maintained by the poor.

Holland, Peterson, and Gonzalez (2009) studied the correlation of income inequality and biodiversity losses across about 50 countries while controlling for other variables. The finding is that there is a positive relationship that is stronger than between income disparity and per capita gross domestic product (PCGDP). One conclusion is that the proportion of endangered species threatened in the U.S. could be expected to ultimately increase from 2.7 to 3.0 percent based upon the GINI increase of 44 to 49 from 1990 to1997 (Holland, Peterson, & Gonzalez, 2009). Increased income inequality can compromise the ability of institutions to manage natural resources as wealthy populations tend to segregate and distance themselves from problems associated with disadvantaged populations (Dietz, Ostrom, & Stern, 2003; Boyce, 1994).

Sylwester (2003) analyzed a cross-section of up to 90 nations with controlling variables to determine if there is an association between SPD in several historic periods with current income inequality in 1990 as measured by GINI. The finding is a small but statistically significant negative correlation for each of 5 periods/years AD (0, 1000, 1500, 1700, 1900) with income inequality, which diminishes in magnitude as time progresses (Sylwester, 2003). Rothwell and Massey (2010) studied 50 suburban U.S. metropolitan areas with controlling variables to identify a relationship of SPD as determined by restrictions from local zoning codes (density zoning) with income segregation as measured by both GINI and the poor-affluent exposure index. The finding is that there is a strong statistically significant causal relationship between density zoning and both dependent variables (DVs). The associations are evident both statically for the year 2000 and over time from 1990-2000 (Rothwell & Massey, 2010).

Testable Hypothesis

Null hypothesis: $H_0\colon B_j = 0$

Both WPD and PCVMT do not have respective negative and positive statistically significant relationships with GINI, and the slope of the population regression is 0.

Alternative hypothesis: $H_1\colon B_j \neq 0$

WPD and PCVMT do have respective negative and positive relationships with GINI, and the slope of the population regression is not 0.

Statistical Techniques, Strategies and Tactics

The analysis is conducted initially via the application of ordinary least squares (OLS) regression analysis using cross-sectional data with 20 initial controlling independent variables (IVs). The dependent variable (DV) to be predicted is GINI. The primary econometric/statistical software used is STATA. Two separate level-level (raw data) models are developed based upon observations of U.S. MSAs with one based on data for the year 2000 and the other on 2010/2015. The model format follows:

- $\hat{y} = \beta_0 + b_1 x_1 + b_2 x_2 + \cdots b_k x_k + \varepsilon$

The time-invariant element of the error term is addressed through a first-differencing fixed-effects level-level model combining both periods. The methodology counters serial correlation (aka autocorrelation) which can cause biased estimates due to the association of variable observations over separate timeframes. The model format is:

- $\Delta\gamma_{it} = \beta_0 + \beta_1 \Delta x_{1,it} + \beta_2 \Delta x_{2,it} + \cdots \beta_k \Delta x_{k,it} + \Delta u_{it}$

The generalized least squares (GLS) aka weighted least squares (WLS) methodology is also employed for the periods of 1990, 2000, and 2010/2015. This methodology combines the information from all three periods via panel data in a pooled cross-section and involves the techniques of a one-period lag or generalized differences equation and demeaning to create the following respective two additional level-level models while addressing serial correlation:

- $(Y_t - pY_{t-1}) = B_1(1 - p) + B_2(X_t - pX_{t-1}) + \cdots B_k(X_t - pX_{t-1}) + v_t$
- $\tilde{y}_{it} = \beta_0 + \beta_1 \tilde{x}_{1,it} + \beta_2 \tilde{x}_{2,it} + \cdots \beta_k \tilde{x}_{k,it} + \tilde{\varepsilon}_{it}$

The Durbin-Watson *d* Test is used to identify the extent of autocorrelation in the multiple period models:

- $d = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}$

A coefficient of autocorrelation or ρ (rho)(lies between -1 and 1) is also calculated for the multiple period models and used in the aforementioned generalized differences equation:

- $\rho = \left(1 - \frac{d}{2}\right)$

Validation includes testing to ensure the four standard assumptions are confirmed: 1) normal distribution absent skewing or disproportionate outliers; 2) a linear relationship exists between the DVs and IVs; 3) variables have been measured accurately without substantive errors; and 4) using scatterplot analysis to ensure uniformity of the IV errors. A potential issue is that the OLS predicted IV coefficients are determined according to the mean values and therefore only result in average marginal impacts on the DV. Thus, quantile regressions are performed at different percentiles to detect varying marginal effects of the IVs on GINI in the final GLS models.

Ramsey Regression Error Specification Tests (RESET) are performed in STATA to determine if the models are misspecified in the form of non-linear components which can compromise predictability. Both the Park Test and Brausch-Pagen Test are used to check for heteroscedasticity or unequal variance in the models which can also cause biased results. The testing is supplemented with analysis of normal probability plots and plots of residuals (difference between observed and predicted values) against predicted values.

Data Needs

Data are collected as available for the variables in *Figure 1* covering up to 381 MSAs for the periods of 2010/2015 and 2000. Data coverage during the initial reverse stepwise regression was restricted to about one-half of the MSAs due to limited availability of one of the potential controlling variables; union membership. However, the number of observations increased to about two-thirds (254) of the MSAs as the union membership IV relationship with GINI lacked statistical significance (i.e. it could not be shown that the association likely did not occur by chance). The missing MSAs tend to be smaller geographically but include some of moderate size. Thus, the sample size should be sufficient. One of the difficulties has been older geographic data is not consistent with more recent years. This is because definitions of some MSAs have changed over time.

Figure 1 – Dependent and Tested Independent Variables and Data Sources

Dependent Variable	**Source/Description**	**2000**	**2010/15**	**Change**	**3-Period**
MSA Gini Coefficient	Census, Census-ACS 15' 5-Yr Est.	GINI00	GINI15	GINICHG	GINI
Independent Variables	**Source/Description**	**2000**	**2010/15**	**Change**	**3-Period**
MSA Weight. Pop. Density	Census	WPD00	WPD10	WPDCHG	
Urban Area Std. Pop.	Census	SPD00	SPD10	SPDCHG	SPD
UA Per Cap. Veh. Mile	FHWA Statistics (HM-72)	PCVMT00	PCVMT15	PCVMTCHG	PCVMT
UA Fwy. Portion PCVMT	FHWA Statistics (HM-72)	FPPCVMT00	FPPCVMT15	FPPCVMTCHG	
UA Freeway PCVMT	FHWA Statistics (HM-72)	FPCVMT00	FPCVMT15	FPCVMTCHG	
MSA Population	Census Est., Census DP-1/5-yr B01003	POP10	POP15	POPCHG	
MSA Per Cap. Pers. Income	Census P082 (15$)/5-yr B19301	PCPI00	PCPI15	PCPICHG	
MSA Dissimilarity Index	Census via Brown Univ. (racial seg.)	DISSIM00	DISSIM00	DISSIMCHG	
MSA % in Poor/Affluent	Census via Brown University	PAFSEG00	PAFSEG10	PAFSEGCHG	PAFSEG
MSA % Foreign/Not Ctzn.	Census-00 DP22/ACS-15 S0501	PFBNC00	PFBNC10	PFBNCCHG	
MSA Single Fem. Head	Census-00 DP1/ACS-15/5yr DP02	SFHHC00	SFHHC15	SFHHCCHG	
MSA Median Age	Census-90, 00 PO-13/ACS-15/5yr S0102	MEDAGE00	MEDAGE15	MEDAGECHG	MEDAGE
MSA % Black Population	Census-00 PO-13/ACS-15/5yr B02001	BKPOP00	BKPOP15	BKPOPCHG	
MSA Labor Partic. (Civ.)	Census-00 DP-3/ACS-15/5yr16 S2301	LFPR00	LFPR00	LFPRCHG	
MSA Unemployment Rate	Census-00 DP-3/ACS-15/5yr16 S2301	UNEMP00	UNEMP15	UNEMP	
MSA Poverty Rate	Census-90, 00 DP-3/ACS-15/5yr S1701	PVRTY00	PVRTY15	PVRTYCHG	PVRTY
MSA percent Pop. 25+ BS	Census-90, 00 QT-P20/ACS-15/5yr S1501	EDUBS00	EDUBS15	EDUBSCHG	EDUBS
MSA Violent Crime Rate	FBI	VCR00	VCR15	VCRCHG	
MSA Per Capita GDP	Bureau of Labor Stats. (chained 09' $)	PCGDP00	PCGDP15	PCGDPCHG	
MSA Union Membership	BLS via Unionstats.com	UNMEM00	UNMEM15	UNMEMCHG	

Additionally, some of the older data can only be extracted at the urbanized area level (UZA) which leaves out a relatively small amount of data in the outer areas of a few MSAs.

It also necessitates combining data from two UZAs when they are part of the same MSA. Ideally, supplemental analysis covering data from decennial periods prior to 2000 would likely provide additional substantive insight on evolution of the relationships between urban form and income inequality. The main constraint is the time necessary to convert county-level GINI data to MSA levels. Therefore, 1990 data was gathered for a smaller sample size for use in the three-period modeling. The appropriate sample size in this expanded data scenario is calculated using Yamane's formula (as cited in Israel, 1992) assuming a 95 percent level of confidence and a margin of error (MOE) or level of precision of ±10 percent. Thus, the sample size chosen is 72 of the 254 MSAs with full data sets:

- $n = \frac{N}{1+N(e)^2} \rightarrow \frac{254}{1+254(0.10)^2} = 72$

General Analytical Methodology

Data are analyzed for 20 IVs to identify relations with variations in the DV of GINI in U.S. MSAs. The DVs and IVs are first analyzed both in terms of descriptive statistics and Pearson Correlation Matrices for 254 observations. Descriptive statistics provide measures of central tendency and dispersion while providing indications of data collection errors. Pearson Correlation matrices are in the appendix for each of the 2000 and 2010/15 data in addition to the differences between them (*Figures A-C*). The matrices identify one-on-one relationship strength between each of the variables which serves as an initial screening for multicollinearity. Multicollinearity creates redundancy which can lead to unreliable results. As expected, several IVs are highly associated with each other which is indicative of possible multicollinearity if included in the same model. The correlation threshold used in the analysis is 0.40 for excluding an IV when associated with another IV having a stronger relationship with a DV.

Development of all models occurs via reverse stepwise regression performed in STATA. Successive runs of the regressions are performed after eliminating those IVs that are not close to approaching statistical significance with the DV. Manipulation of the models throughout the reverse stepwise regressions occur based on strength of the standard statistical measures. Specifically, model robustness is measured by R^2 (proportion of explained variability) and statistical significance (relationship does not occur by chance) as measured by F >~2.10 to 2.34 (dependent upon number of IVs) and *significance* $F \leq 0.05$. Model validity is also based on IV p values ≤0.05 and t scores ≥1.96 in relation to a 0.95 level of confidence (repeated sampling mimics the actual population 95 percent of the time), and the aforementioned statistical tests and plots.

Cross-Sectional and First-Differencing Approaches

The reverse stepwise sequencing began by running three separate regressions on the DVs of GINI00 (2000 data), GINI15 (2015 data) and GINICHG (2000-2015 change) against 20 IVs for the respective periods. Of note is the single year models are much more robust than the first-differenced two-period model in terms of the R^2 and higher F values. The resulting level-level equations are as follows:

- $GINI00 = 0.2330 + 0.0985(PAFSEG00) + 0.0034(MEDAGE00) + 0.0971(EDUBS00) + 0.4352(PVRTY00) - 0.0000021(FPCVMT00)$*; R^2=0.6489; F=91.31; Sig. F=0.0000 (note: 253 observations)*
- $GINI15 = 0.2700 + 0.0818(PAFSEG10) + 0.0021(MEDAGE15) + 0.1302(EDUBS15) + 0.3962(PVRTY00) + 0.0248(BKPOP15) - 0.1645(UNEMPR15)$*; R^2=0.6570; F=78.86; Sig. F=0.0000*
- $GINICHG = 0.0087 + 0.333(PAFSEG) + 0.0015(MEDAGECHG) + 0.487(EDUBSCHG) - 0.1069\ (EMPRCHG) + 0.2058(PVRTYCHG) - 0.00000219(FPCVMTCHG)$*; R^2=0.1767; F=8.83; Sig. F=0.000*

Analysis of all three models reveal histograms of the DVs showing the data are relatively normal distributions, i.e. unbiased or not skewed. Normal probability plots show some tailing at both ends indicating a possible concern that the distributions may not be normal in all three cases. The standard residuals (difference between observed GINI and estimated GINI) plotted against the predicted GINIs show moderately uniform scattering and a slope of 0. Yet, in all three models there is some concern about non-normality due to either small amounts of outliers or skewing. Ramsey RESET tests were performed on all three models to check for omission of variables and inappropriate functional form. Results show that the GINI15 and GINICHG models are not misspecified as the respective *F* values are below the *F* table threshold of 2.10 and above the *significance F* threshold of 0.05. However, the Ramsey RESET test for GINI00 shows the model is misspecified. This was rectified by removing one observation (Prescott, AZ) which was an outlier in the standard residual/predicted GINI00 plot.

A Park Test was performed to check for heteroscedasticity (i.e. biased standard errors of the estimates and unreliable confidence intervals) in the models by squaring the residuals or errors, converting them to log format, and regressing the predicted *Y* values against the results. A Breusch-Pagen Test is also used for the same purpose by regressing the squared residuals against the IV data. In all three models, the Park Tests and Breusch-Pagen Tests revealed that the *t stat* values are not above the t table two-tailed threshold of ±1.960 and the *P-value/Significance F* values are above the threshold of 0.05. Thus, there are not statistically significant associations and there is no indication of heteroskedasticity in the models. Printouts of the STATA regression results, Ramsey RESET Tests, histograms of the GINI data, normality plots, and scatter diagrams of the predicted GINI values regressed against the residuals are in the appendix (*Figures D-P*).

The population density and travel-related metrics lack statistical significance in all three models. However, FPCVMTCHG has a negative association with GINICHG. Thus, a one-unit increase in FPCVMTCHG equates to a GINICHG reduction of 0.00000219, ceteris paribus. In terms of magnitude, the association is relatively inconsequential. Of note is that PAFSEG, which is highly correlated with WPD, has a statistically significant positive relationship with the DVs in all three models. Thus, all other circumstances remaining the same, a 1 percent increase in PAFSEG between the two periods is correlated with a considerable increase in GINI of 0.0033 or 0.33 percent. Nevertheless, based on the three models, the null hypotheses are accepted as GINI does not have statistically significant relationships with either WPD or PCVMT.

Within-Transformation Approach

To increase the number of periods, a random sample of 72 MSAs was taken of 1990 data for reappearing IVs in the three regression models. SPD was included in the initial run but lacked statistical significance. Time constraints did not allow for developing 1990 WPD data. The data for the three periods were arranged into a pooled fixed effects panel set which effectively created 216 observations. A Durbin-Watson Statistic was used to identify expected autocorrelation. In response, the data was transformed using a one-period lag. This involves calculating and removing a portion (*p* or rho) of the value of each variable in a previous period from a current period. This does remove the 1990 data, however, it was retained using *p* in a Prais-Winsten transformation. OLS regression was applied to the transformed model to obtain GLS estimators. The resulting model passed the Park Test and the Brausch-Pagen Test. However, it failed the Ramsey RESET Test and the Durbin-Watson statistic showed only slight improvement to 1.485 and remained just outside the range of 1.5 to 2.5 which is considered relatively normal for no autocorrelation. A scatterplot of the predicted GINI against the standard residuals showed a somewhat non-random pattern with coning narrowing to the right.

Alternatively, the three-period pooled cross-sectional panel data was demeaned (removing the means to address bias/autocorrelation between periods) to reveal the following equation and fourth level-level model:

- $GINI = 0.2892 + 0.0310(PAFSEG) + 0.0021(MEDAGE) + 0.1403(EDUBS) + 0.1921(PVRTY) + 0.00000122(PCVMT)$; R^2 *(overall)= 0.5621; F=109.85; Sig. F=0.000*

The Durbin-Watson Statistic showed slightly better improvement to 2.495. Model validity was confirmed by the Park, Brausch-Pagen, and Ramsey RESET Tests. The GINI scatterplot against the residuals showed perhaps a minor improvement. All the IVs have positive statistically significant relationships with GINI at the 95 percent level of confidence and $p<0.05$ except for PAFSEG which is at the 90 percent confidence level and $p<0.10$. Therefore, predicted increases in GINI will be within 10 percentage points of the actual population value 95 percent of the time (except in the case of PAFSEG it will be 90 percent of the time). Thus, a 1-unit increase in PCVMT is associated with a 0.00000122 increase in GINI which is negligible. However, an MSA near the high end of the PCVMT range at 15,000 can be expected to have a GINI of 0.014 (MOE ±0.0014) higher than an MSA at a low end of 3,500 PCVMT 95 percent of the time. Thus, using the within-transformation demeaning approach, the null hypothesis that PCVMT is not associated with GINI can be rejected. Poverty rate and education by far have the highest degree of associations as respective 1 percent increases in each can be expected to increase GINI by 0.001921 and 0.001403. A 1 percent increase in PAFSEG and a 1-year increase in MEDAGE is associated with respective increases in GINI of 0.00031 and 0.0021. Printouts of the STATA regression results, Ramsey RESET Test, histogram of the GINI data, normality plots and scatter diagram of the predicted GINI values regressed against the residuals are in the appendix (*Figures S-U*).

A possible concern is that the OLS estimated IV coefficients are determined pursuant to the mean values and only indicate average marginal effects on the DV. Therefore, quantile regressions were run at selected percentiles to identify different marginal effects of PCVMT on GINI. The results are in *Figure 2* below and show that PCVMT increases only have a statistically significant relationship with GINI at the 95 percent confidence level from about the 40th through the 65th quantiles. Thus, the association is only valid in these ranges of the GINI data. At the 40 percent quantile an increase in PCVMT of 1 is associated with an increase in GINI of 0.00001130 while at the 65th quantile the increase in GINI expected is 0.000001230. There is no statistical significance between PCVMT and GINI within the other quantiles. Of note is that at the 90th percentile the PCVMT coefficient turns negative and a reduction in GINI anticipated is -0.000000076 but it lacks statistical significance. A printout of the STATA quantile regression results is in the appendix (*Figure V*).

Figure 2 - Estimated Quantile Regression Slope Coefficients (level-level)

	OLS	Quantile Regression							
	Mean	10th	25th	30th	40th	Median	65th	75th	90th
Intercept	0.28921	0.24792	0.26590	0.27056	0.28343	0.28145	0.29190	0.27824	0.28414
	(0.01162)	(0.01350)	(0.01453)	(0.01468)	(0.01376)	(0.01712)	(0.02011)	(0.02339)	(0.03145)
PAFSEG	0.03103	0.03976	0.04711	0.05178	0.04955	0.04693	0.05224	0.04996	0.06537
	(0.01879)	(0.01982)	(0.01968)	(0.01918)	(0.01462)	(0.01454)	(0.01584)	(0.01727)	(0.01220)
PVRTY	0.19210	0.43337	0.44409	0.43986	0.42182	0.42981	0.44244	0.46240	0.41694
	(0.03854)	(0.04705)	(0.05886)	(0.05130)	(0.03915)	(0.04145)	(0.02624)	(0.04516)	(0.04560)
PCVMT	0.000001220	0.000000359	0.000000689	0.000000885	0.000001130	0.000001510	0.000001230	0.000000374	-0.000000076
	(0.000000632)	(0.000000598)	(0.000000569)	(0.000000532)	(0.000000605)	(0.000000723)	(0.000000636)	(0.000001090)	(0.000000877)
EDUBS	0.14028	0.11234	0.09758	0.09414	0.08657	0.09381	0.08386	0.09188	0.07790
	(0.02248)	(0.02434)	(0.02297)	(0.01996)	(0.01308)	(0.01483)	(0.01001)	(0.00969)	(0.01233)
MEDAGE	0.00210	0.00221	0.00185	0.00172	0.00157	0.00155	0.00140	0.00202	0.00242
	(0.00049)	(0.00030)	(0.00030)	(0.00039)	(0.00041)	(0.00051)	(0.00056)	(0.00082)	(0.00082)
Observations	216	216	216	216	216	216	216	216	216
Note: Dependent variable is gini. Standard errors in parentheses. Dark blue highlighted values are significant at the 5 percent level (majority). Light blue highlighted values are significant at the 10 percent level (two instances).									

In an effort to improve the three-year panel model, the data was transformed to log-log and reverse stepwise regression was performed to reveal the best configuration as follows:

- $\log(GINI) = 1.0463 + 0.0566\log(EDUBS) + 0.2284\log(MEDAGE) + 0.0699\log(PVRTY) + 0.0232\log(PCVMT)$*; R^2 (overall)= 0.5263; F=117.77; Sig. F=0.000*

Of note is that SPD and PAFSEG are not statistically significant. All the IVs in the model have positive statistically significant relationships with GINI at the 95 percent level of confidence and $p<0.05$ except for PCVMT which is at the 90 percent confidence level and $p<0.10$ (respective t and p values of 1.79 and 0.076). Therefore, predicted increases in GINI will be within 10 percentage points of the actual population value 95 percent of the time (except in the case of PCVMT it will be 90 percent of the time). A Ramsey RESET test showed the model just on the cusp of misspecification with $F = 2.52$ and $prob. > F = 0.0587$. Both the Park Test and Brausch-Pagen test indicated there is no heteroscedasticity in the model. Plots for normality and predicted LOGGINI against standard residuals plots are generally consistent with these tests. The Durbin-Watson statistic was satisfactory as it showed no indication of serial correlation. The log-log version of the model indicates that a 1 percent increase in PCVMT is associated with a 0.023 percent (MOE ±0.0023) increase in GINI. Again, as with the level-level regression, the magnitude of the increase in the log-log model is relatively small. The log-log model shows that a 1 percent increase in MEDAGE, the IV most associated to the DV, equates to about a 0.23 percent increase in GINI. Printouts of the STATA regression results, RAMSEY Reset Test, and aforementioned plots are in the appendix (*Figure W-Z*).

Again, a possible concern is that the OLS estimated IV coefficients are determined pursuant to the mean values and only indicate average marginal effects on the DV. Therefore, quantile regressions were run at selected percentiles to identify different marginal effects of LOGPCVMT on LOGGINI. The results are in *Figure 3* below and show that PCVMT increases only have a statistically significant relationship with GINI at the 95 percent confidence level from about the 25th through the 60th quantiles. Thus, the association is only valid in these ranges of the LOG data. At the 95 percent confidence level, in the 35 percent and median quantiles a 1 percent increase in PCVMT equates to a 0.035 percent increase in GINI while at the 25 and 60 quantiles the increase in GINI expected is 0.026 percent. Statistical significance is lacking between PCVMT and GINI in all other quantiles. A printout of the STATA quantile regression results is in the appendix (*Figure AA*).

Figure 3 - Estimated Quantile Regression Slope Coefficients (log-log)

	OLS	Quantile Regression							
	Mean	10th	25th	35th	Median	60th	65th	75th	90th
Intercept	1.04630	1.06023	1.04230	1.02760	1.06890	1.12114	1.11230	1.13723	1.15840
	(0.05790)	(0.07073)	(0.03904)	(0.05546)	(0.06007)	(0.07314)	(0.07382)	0.07552	(0.10013)
PVRTY	0.06988	0.17727	0.15752	0.15870	0.15501	0.14860	0.15535	0.15015	0.15589
	(0.01104)	(0.21677)	(0.01062)	(0.00933)	(0.00698)	(0.00998)	(0.00714)	(0.00957)	(0.01981)
PCVMT	0.02315	-0.00371	0.02602	0.03543	0.03531	0.02641	0.02044	0.01297	-0.00270
	-0.01296	(0.01524)	(0.00698)	(0.01024)	(0.01360)	(0.01177)	(0.01365)	(0.01623)	(0.01690)
EDUBS	0.05658	0.08394	0.06384	0.06352	0.06133	0.05755	0.05937	0.05630	0.06189
	(0.01614)	(0.01543)	(0.00821)	(0.00787)	(0.00679)	(0.00553)	(0.00642)	(0.00644)	(0.00856)
MEDAGE	0.22835	0.17253	0.14690	0.13429	0.11514	0.11446	0.13116	0.14113	0.16462
	(0.04608)	(0.01674)	(0.02191)	(0.02490)	(0.02624)	(0.02584)	(0.02566)	(0.03866)	(0.07066)
Observations	216	216	216	216	216	216	216	216	216
Note: Dependent variable is gini. Standard errors in parentheses. Dark blue highlighted values are significant at the 5 percent level. No shading indicates not significant at either level.									

Ethical Considerations

Caution should be used in applying the research models to the practice of urban planning as the findings are conditioned on the range and accuracy of the contributing data. Additionally, predictions for changes in income inequality could be impacted in the future by unforeseen circumstances due to evolving technology in transport or other fields. Possibilities include alternate fuels that limit emissions, development of renewable fuels, or increases in typical travel speeds. In any of these scenarios, distance within the built environment may not be as much of an impediment to achieving lower income inequality or other social objectives. Ultimately, the subject research should not be looked upon as a rigid prescription for solving the problem by altering urban form or transportation patterns. Rather, it should be used by localities to inform decision-making in balance with other priorities and fiscal realities in relation to collective quality of life values.

It should also be acknowledged that there is criticism of the spatial mismatch hypothesis. As discussed by Blumenberg and Manvillle (2004), researchers have made the case that income inequity is based more on racial discrimination and ethnic separation of the labor market as opposed to distance-based seclusion. Evidence has revealed that distance is not always a factor in income disparities, but it is often intertwined with discriminatory practices to the point where race and space is synonymous with each other. Metropolitan areas such as those in the west and south that matured in later decades tend not to have the same spatial mismatches as in other U.S. urban areas. Other factors are number of jobs available versus number of applicants in addition to skill levels of the local workforce which is consistent with the strong inverse association between education and income inequality (Blumenberg & Manvillle, 2004).

Qualitative Considerations

In identifying and selecting land use and transportation projects, there are qualitative factors that should be taken into account apart from the quantitative elements relating to urban form, travel and income inequality. The quantitative aspects include not only the subject study IVs and models but also conventional benefit-cost analyses typically used to guide decision-making. As discussed by Perugini and Martino (2008), qualitative factors includes the development of intensified highly skilled labor demands in association with evolving trade and specialization. The distributive impacts relative to income inequality can vary based upon the available labor pool, willingness to adapt, country, and the position or strength of a given nation regarding particular industries or services. Anti-distributive policies can also vary by country or within countries based upon local politics. In another vein, survey quality standards based on households, inclusion of all income sources and appropriate population representation are critical components of accurate enumerations. These can be compromised by poor oversight and human error (Perugini & Martino, 2008). Therefore, qualitative features can impact the findings relative to predictability of income inequality.

In a different sense, qualitative factors arguably related to income inequality that can be difficult to measure consist of "imageability" or "likeability" of the built environment as espoused by Lynch (1960). Surveys have found that people tend to identify desirable urban form in terms of paths (roads, sidewalks) edges (continuity breaks), districts (subareas with unique character), nodes (junctions and concentrated gathering points), and landmarks. Such urban traits laid out in an orderly fashion tend to give one a sense of emotional well-being and increased depth and concentration within the human experience (Lynch, 1960). These tend to be consistent with Traditional Neighborhood Development (TND) more prevalent in the pre-World War II period typified by concepts of gridded streets with compact and walkable or more human-scale urban form and community character. Further, studies have shown that older homes in such prewar TND residential areas tend to command price premiums compared to most newer subdivisions, controlling for other factors (Bitter, 2013). Without affordable housing, these qualitative elements can potentially be a factor in segregation, income inequality and gentrification (Koschinsky & Talen, 2015).

According to Dinzey-Flores (2017), addressing imageability deficiencies are both a symbolic and substantive way to reduce the perceptions and experiences of inequality that likely inhibit both well-being and initiative by disadvantaged populations. Yet, close proximity of new and prestigious development with urban ghettos may not provide a pathway to reduced income inequality if there is not true integration of physical and functional form of the two areas. It is the perceptions by disadvantaged populations of prejudice and lack of opportunities that can be more meaningful than quantitative measures. Alternative quality of life metrics have emerged based on surveys of the populace, such as the Gross National Happiness in Bhutan, and arguably provide more insight than traditional numeric measures (Dinzey-Flores, 2017).

Conclusion

Overall, the research revealed that there is not a statistically significant negative relationship between WPD and GINI, therefore, the null hypothesis is accepted and the alternate hypothesis is rejected. Conversely, on average, there is a statistically significant positive relationship between PCVMT and GINI at the 95 percent confidence level (level-level, 90 percent for log-log), thus, the null hypothesis is rejected, and the alternate hypothesis is accepted. The caveat to the latter conclusion is that the association held true in panel data pooled cross-sectional analysis over three periods together at certain mid-range quantiles while the finding could not be verified in single period or more short-term cross-sectional examination. Additionally, the magnitude of the association is relatively small. The log-log transformation of the model revealed that a 0.023 percent increase in GINI can be expected with a 1 percent increase in PCVMT (MOE: 0.0023). The one instance where the null hypothesis is rejected is at least indirectly consistent with the spatial mismatch hypothesis which states that restrictions on black residential choice together with employment disbursement is responsible for high joblessness and low income for those minority populations (Kain, 1994). This is inherent due to the demand of increased travel needs created in part by more scattered urban form.

In a similar vein, the level-level OLS and demeaned GLS analysis did find that PAFSEG, which is highly correlated with WPD, does have a statistically significant positive relationship with GINI. A 1 percent increase in PAFSEG is associated with an increase in GINI ranging from

0.33 percent in the OLS first-differenced two-period model, to 0.031 percent in the GLS demeaned three-period model, and 0.10 in the OLS cross-sectional model for the year 2000 (all level-level). The magnitude of these changes is substantive at the high end but not at the low end.

In U.S. metropolitan areas, only about 30 percent of all jobs and less than one-fourth of low and middle-skilled jobs are located within a 90-minute commute by public transportation (Tomer, Kneebone, Puentes, & Berube, 2011). Therefore, consistent with other research, it is ostensible that there are long-term ramifications from population disbursement, segregation and transport limitations. Consequently, there is value in planning and subdivision/zoning regulations that consider clustered development patterns, limit new land consumption, provide incentives for affordable housing, and support for public transportation. The research is useful for decisionmakers contemplating the potential inequitable ramifications of continued low-density development lacking affordable housing in relation to other regional goals of reducing traffic congestion, energy usage, noise and transport emissions associated with climate change.

The main impediment of the study was the inability to expeditiously obtain data from earlier periods before 1990. In relation, the data used was challenging as some of it was unavailable for various MSAs and some required time-intensive conversions. Therefore, further research is recommended to acquire and analyze the necessary data as far back as the early 20th century for use in a follow-up longitudinal study. This could use other measures of income inequality besides GINI such as the Atkinson index or Generalized Entropy index. Results could be more revealing from earlier periods during construction and maturation of the national highway system, mass exodus from major cities and rampant suburbanization, and the related escalation in PCVMT with proliferation of the automobile.

References

Andersson, F., Haltiwanger, J., Kutzbach, M., Pollakowski, H., Weinberg, D. (2014). Job Displacement and the Duration of Joblessness: The Role of Spatial Mismatch. National Bureau of Economic Research. Retrieved via: http://www.nber.org/papers/w20066.pdf.

Autor, D., Katz, L., & Kearney, M. (2006). Measuring and interpreting trends in economic inequality. *The American Economic Review, 96*(2), 189-194.

Bitter, C. (2013). Subdivision vintage and housing prices: Do home buyers value traditional development? *Urban Studies, 51*(5), 1038-1056.

Blumenberg, E. & Manvillle, M. (2004). Beyond the spatial mismatch: welfare recipients and transportation policy. *Journal of Planning Literature, 19*(2), 182-205.

Boyce, J. (1994). Inequality as a cause of environmental degradation. *Ecological Economics, 11,* 169-178.

Cervero, R. & Murakami, J. (2010). Effects of built environments on vehicle miles traveled: evidence from 370 urbanized areas. *Environment and Planning, 42*(2), 400-418.

Deaton, A. (2003). Health, inequality, and economic development. *Journal of Economic Literature, XLI,* 113-158.

Dietz, T., Ostrom, E. & Stern, P. (2003). The struggle to govern the commons. *Science, 302,*1907-1912.

Dinzey-Flores, Z. (2017). Spatially polarized landscapes and a new approach to urban inequality. *Latin American Research Review, 52*(2), 241-252.

Eidlin, E. (2010). What density doesn't tell us about sprawl. *Access, 37*, 2-9. Available via https://www.accessmagazine.org/fall-2010/density-doesnt-tell-us-sprawl/.

Ewing, R. & Hamidi, S. (2014). Measuring Sprawl 2014. Smart Growth America. Available via https://www.smartgrowthamerica.org/app/legacy/documents/measuring-sprawl-2014.pdf.

Ewing, R., Hamidi, S., Grace, J. & Wei, Y. (2016). Does urban sprawl hold down upward mobility? *Landscape and Urban Planning, 148*, 80-88.

Glenn, N. (1968). The Kerner Report, social scientists, and the American Public: Introduction to a Symposium. *Social Science Quarterly, 49,* (3), 433-437.

Hatch, M. & Rigby, E. (2015). Laboratories of (in)equality? Redistributive policy and income inequality in the American states. *The Policy Studies Journal 43,*(2), 163-187.

Holland, T., Peterson, G., & Gonzalez, A. (2009). A cross-national analysis of how economic inequality predicts biodiversity loss. *Conservation Biology, 23*(5), 1304-1313.

Holzer, H. (1991). The spatial mismatch hypotheses: what has the evidence shown? *Urban Studies, 28*(1), 105-122.

Israel, G. (1992). Determining Sample Size. University of Florida, IFAS Extension. Available Via https://www.tarleton.edu/academicassessment/documents/Samplesize.pdf.

Jones, J., Schmitt, J. & Wilson, V. (2018). 50 years after the Kerner Commission. Economic Policy Institute. Retrieved via https://www.epi.org/publication/50-years-after-the-kerner-commission/.

Kain, J. (1994). The spatial mismatch hypothesis: three decades later. *Housing Policy Debate, 3*(2), 371-392. Retrieved via https://inequality.stanford.edu/sites/default/files/media/_media/pdf/Classic_Media/Kain_1992_Transportation.pdf.

Kain, J. (2004). A pioneer's perspective on the spatial mismatch literature. *Urban Studies, 41*(1), 7-32.

Kim, S. (2007). Changes in the nature of urban spatial structure in the United States, 1890-2000.

Journal of Regional Science, 47(2), 273-287.

Koschinsky, J. & Talen, E. (2015). Affordable housing and walkable neighborhoods: A national urban analysis. *Cityscape, 17*(2), 13-56.

Lynch, K. (1960). *The Image of the City.* Cambridge, MA: MIT Press.

Mayer, S. (2001). How Growth in Income Equality Increased Economic Segregation. Working Paper 230. Chicago: Joint Center for Poverty Research.

Perugini, C. & Martino, G. (2008). Income inequality within European regions: determinants and effects on growth. *Review of Income and Wealth, 54*(3), 373-406.

Rothwell, J. & Massey, D. (2010). Density zoning and class segregation in U.S. metropolitan areas. *Social Science Quarterly, 91*(5), 1123-1143.

Sylwester, K. (2003). Income inequality and population density 1500 AD: a connection. *Journal Of Economic Development, 28*(2), 61-82.

Tomer, A., Kneebone, E., Puentes, R., & Berube, A. (2011). Missed Opportunity: Transit and Jobs in Metropolitan America. Metropolitan Policy Program at Brookings. Retrieved via: https://www.brookings.edu/wp-content/uploads/2016/06/0512_jobs_transit.pdf.

U.S. Census Bureau. (2016). Historical Income Tables: Income Inequality. Retrieved via: https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-income-inequality.html.

Appendix

Figure A – Pearson Correlation Matrix (2000 Data)

	GINI00	*POP00*	*WPD00s*	*SPD00*	*PCPI00*	*PCGDP00*	*DISSIM00*	*PAFSEG00*	*PFBNC00*	*SFHHC00*	*MEDAGE00*	*BKPOP00*	*EDUBS00*	*LFPR00*	*UNEMPR00*	*PVRTY00*	*VCR00*	*PCVMT00*	*FPPCVMT00*
GINI00	1																		
POP00	0.16206593	1																	
WPD00s	0.16126069	0.82181321	1																
SPD00	0.13221556	0.5254067	0.72556726	1															
PCPI00	-0.08922545	0.31488137	0.33200159	0.18760494	1														
PCGDP00	-0.06501621	0.02619572	0.0484107	0.00735748	0.34455367	1													
DISSIM00	0.11797331	0.38628963	0.26470437	0.0731111	0.3261661	-0.02681505	1												
PAFSEG00	0.47034459	0.47931106	0.4998257	0.43009769	0.20568494	-0.01967329	0.44018849	1											
PFBNC00	0.28515241	0.35353029	0.50887862	0.58223633	0.0205307	0.09624764	-0.0666583	0.33718454	1										
SFHHC00	0.17728969	0.02709899	-0.03964962	-0.08141642	-0.36611639	-0.10123409	0.17173717	0.39462634	0.05631706	1									
MEDAGE00	-0.02100523	0.05419345	-0.03454938	-0.23577943	0.41230808	-0.05471373	0.36544627	-0.07742403	-0.23075454	-0.08204385	1								
BKPOP00	0.20758768	0.11713213	-0.04480167	-0.1966199	0.03174606	0.03749968	0.35046372	0.36658543	-0.22578665	0.59892479	0.0279726	1							
EDUBS00	0.05955053	0.13563546	0.20601953	0.20283423	0.55992518	0.36676	-0.09767465	0.11355724	0.01338621	-0.40982885	-0.18727516	-0.01471415	1						
LFPR00	-0.37689126	0.05010655	0.04474662	0.09200509	0.42933655	0.23882294	0.06123898	-0.04331097	-0.16747203	-0.2771504	-0.1145503	-0.12612842	0.42651202	1					
UNEMPR00	0.28240364	-0.00816236	0.06739941	0.18596694	-0.52277338	-0.19112377	-0.14554908	0.21133958	0.41017027	0.50838311	-0.38475153	0.01447384	-0.33781843	-0.30249611	1				
PVRTY00	0.54441786	-0.09844578	-0.06983172	0.03358363	-0.67920856	-0.15323388	-0.2614819	0.12406332	0.31285029	0.3176605	-0.50286727	0.05459691	-0.18807835	-0.47971441	0.63608415	1			
VCR00	0.36080777	0.26958719	0.17392888	0.09903912	-0.01061116	-0.07663494	0.30832767	0.50000383	0.15546055	0.4581335	0.11710006	0.4917753	-0.11735741	-0.20248002	0.17473259	0.16814905	1		
PCVMT00	-0.01015865	-0.00571706	-0.21512134	-0.33595554	0.07994077	0.10225258	0.17735998	0.02048761	-0.15532801	0.02225736	0.20488719	0.18013856	0.00254567	0.0138388	-0.24027878	-0.08576242	0.09187489	1	
FPPCVMT00	-0.07306594	0.09021002	0.08714795	0.08041159	0.13719504	0.1749497	0.10826279	0.05387239	0.07003871	0.06122895	0.08200578	-0.00288436	0.0769473	0.09198617	-0.01771178	-0.10446257	-0.01518797	-0.05534417	1
FPCVMT00	-0.1247676	0.1556574	0.0537299	0.00161542	0.25037593	0.22678558	0.28343602	0.16671433	0.02204129	0.05065945	0.10032331	0.10219371	0.15917932	0.20468372	-0.13997576	-0.19836557	0.01661846	0.52780703	0.64875047

Figure B – Pearson Correlation Matrix (2010/15 Data)

	GINI15 (1yr)	*GINI15 (5yr)*	*POP15*	*WPD10*	*SPD10*	*PCPI15*	*PCGDP15*	*DISSIM10*	*PAFSEG10*	*PFBNC15*	*SFHHC15*	*MEDAGE15*	*BKPOP15*	*EDUBS15*	*LFPR15*	*UNEMPR15*	*PVRTY15*	*VCR15*	*PCVMT15*	*FPPCVMT15*	*FPCVMT15*
GINI15 (1yr)	1																				
GINI15 (5yr)	0.90469442	1																			
POP15	0.24463405	0.26017423	1																		
WPD10	0.21304906	0.22684397	0.78910954	1																	
SPD10	0.1432201	0.14576768	0.50633655	0.73077415	1																
PCPI15	0.08853473	0.12890763	0.3252826	0.38670444	0.25224506	1															
PCGDP15	0.13067601	0.16849974	0.3214382	0.35309144	0.30783966	0.68136934	1														
DISSIM10	0.24778415	0.26355907	0.42819796	0.27179357	0.03021769	0.26660004	0.28352595	1													
PAFSEG10	0.46297601	0.51109992	0.47163834	0.42358465	0.39711952	0.1427936	0.2468285	0.46750719	1												
PFBNC15	0.28930887	0.30179198	0.32058537	0.46357952	0.62324641	0.00764151	0.1466637	-0.06586251	0.40147884	1											
SFHHC15	0.10285883	0.06081855	0.01761135	-0.04837977	-0.03990088	-0.49009216	-0.19427257	0.19824891	0.40523198	0.21275107	1										
MEDAGE15	-0.06076183	-0.03198974	0.0194956	-0.0879883	-0.27062405	0.32697522	-0.00441372	0.2468477	-0.17246837	-0.28766446	-0.26437673	1									
BKPOP15	0.25477417	0.26919245	0.14182096	-0.06720118	-0.25756866	-0.04284228	0.0281222	0.40786097	0.33162796	-0.25017092	0.46069354	-0.00359552	1								
EDUBS15	0.24187857	0.28323702	0.21308822	0.28005988	0.24053383	0.71836784	0.45581697	0.00816415	0.10745998	0.00059765	-0.5890904	-0.08981292	-0.0529529	1							
LFPR15	-0.1907404	-0.22031641	0.17918065	0.20840658	0.19563455	0.51153421	0.56414722	0.10373277	0.1226297	-0.06094098	-0.12076981	-0.33303662	-0.02674977	0.48214701	1						
UNEMPR15	0.13172917	0.13381514	0.05586735	0.01500086	0.03808019	-0.3226565	-0.40389641	0.01755475	0.2241707	0.22442486	0.45487291	0.14686577	0.24170535	-0.4187637	-0.48106772	1					
PVRTY15	0.42501972	0.43471076	-0.14298566	-0.14467324	-0.02011608	-0.73903386	-0.49609097	-0.18040786	0.17549717	0.30763672	0.48457356	-0.4093207	0.12408739	-0.34255965	-0.54194428	0.41870264	1				
VCR15	0.19984811	0.20611481	0.1012045	0.06220919	0.08679316	-0.12161917	-0.01450391	0.28900456	0.3903999	0.05946721	0.44346821	-0.03662294	0.36188908	-0.24644687	-0.09709769	0.32809471	0.19013263	1			
PCVMT15	0.01419844	0.04231028	-0.05916582	-0.33540748	-0.5178942	-0.01665791	-0.0339141	0.18054168	-0.05887722	-0.29188871	0.00863131	0.33221753	0.37002934	-0.08875273	-0.11306647	0.01317446	-0.08335174	0.06759288	1		
FPPCVMT15	-0.10195314	-0.07024979	0.26481881	0.27698928	0.24721791	0.32760592	0.27521002	0.33294011	0.26186326	0.15338845	0.01952125	0.04474344	-0.00948842	0.12508849	0.25817904	-0.11398556	-0.29093998	0.08159382	0.15236652	1	
FPCVMT15	-0.06000193	-0.02713214	0.14512126	0.00392599	-0.09919541	0.21756857	0.18099809	0.32509384	0.14226881	-0.0559299	0.0087192	0.16672034	0.17109944	0.0638284	0.1537427	-0.09565313	-0.24210204	0.06643099	0.64899206	0.82379014	1

Figure C - Pearson Correlation Matrix (2000-2010/15 First Differenced Data)

	GINICHG	POPCHG	WPDCHG	SPDCHG	PCPICHG	PCGDPCHG	DISSIMCHG	PAFSEGCHG	PFBNCCHG	SFHHCCHG
GINICHG	1									
POPCHG	0.06521036	1								
WPDCHG	-0.08721114	0.10274303	1							
SPDCHG	-0.07543245	0.01793745	0.57668821	1						
PCPICHG	-0.09534865	-0.05698292	0.13243405	0.06675636	1					
PCGDPCHG	-0.04397416	-0.23589588	-0.16242455	-0.1077159	-0.02513607	1				
DISSIMCHG	0.01191821	0.08095479	0.04506071	-0.07928695	0.2014858	-0.03097008	1			
PAFSEGCHG	0.17766055	0.08904929	0.00265663	0.05206295	-0.14746759	-0.14973598	0.04589552	1		
PFBNCCHG	0.08514864	0.10061982	0.01135617	-0.16575462	-0.08758158	0.00308593	-0.05542929	-0.00301337	1	
SFHHCCHG	-0.06642109	-0.01529029	-0.02210706	-0.09134133	-0.18494284	0.30512863	-0.10806991	0.05317084	-0.07698919	1
MEDAGECHG	0.06549619	-0.00214134	0.09668123	0.05534964	0.16852331	-0.51020422	0.04768708	0.03482557	-0.06629168	-0.30894579
BKPOPCHG	0.02899465	-0.03754757	0.03483528	-0.01412886	-0.1266638	0.21639661	-0.06923662	-0.07645541	0.07408167	0.3220701
EDUBSCHG	0.05680763	-0.00080134	-0.06847403	-0.26216071	0.0097509	0.00409784	-0.01419994	-0.076913	0.06670259	-0.05287601
LFPRCHG	-0.07273251	-0.01470747	-0.06758098	-0.00511494	0.29523388	0.13390546	0.06967499	-0.18857989	-0.19652103	0.22831763
UNEMPRCHG	0.04040345	0.05144789	-0.04178168	0.10557857	-0.37434864	-0.09866868	-0.10170497	0.15386943	-0.0426598	0.02112632
PVRTYCHG	0.30517479	-0.01659779	-0.13083087	-0.06905783	-0.54397032	0.17579227	-0.12504103	0.24931621	0.21141495	0.1769407
VCRCHG	0.08601122	-0.10238784	0.0177736	-0.02814951	0.09738448	0.06006151	0.00644426	0.0492296	0.17516942	0.0715915
PCVMTCHG	-0.01700864	-0.02792239	-0.04248052	-0.09547509	0.04967175	-0.07662676	0.0695265	-0.04723664	0.13672132	-0.17929342
FPPCVMTCHG	-0.06434847	-0.01128807	-0.03249171	0.03422148	-0.09677034	0.17728019	-0.02051535	0.0283349	0.00384162	0.18565461
FPCVMTCHG	-0.11657558	-0.05963985	-0.05283181	-0.09836442	-0.03306303	0.02250142	0.01596072	0.01912922	0.16075901	-0.0186163

	MEDAGECHG	BKPOPCHG	EDUBSCHG	LFPRCHG	UNEMPRCHG	PVRTYCHG	VCRCHG	PCVMTCHG	FPPCVMTCHG	FPCVMTCHG
GINICHG										
POPCHG										
WPDCHG										
SPDCHG										
PCPICHG										
PCGDPCHG										
DISSIMCHG										
PAFSEGCHG										
PFBNCCHG										
SFHHCCHG										
MEDAGECHG	1									
BKPOPCHG	-0.19586784	1								
EDUBSCHG	0.08514578	0.00833154	1							
LFPRCHG	-0.23420239	-0.00320114	-0.03002746	1						
UNEMPRCHG	0.00954795	0.03209225	0.03159378	0.09014481	1					
PVRTYCHG	-0.303315	0.18308871	0.02767822	-0.18845555	0.40629883	1				
VCRCHG	-0.13332883	0.11631493	-0.03240623	-0.0142206	-0.20617106	0.02071542	1			
PCVMTCHG	0.01362581	-0.06994333	0.03149542	-0.00080155	-0.01458378	-0.02189131	-0.11348252	1		
FPPCVMTCHG	-0.06101996	0.04715502	-0.00309289	0.01331785	-0.06267391	0.05971533	0.01824238	-0.56331695	1	
FPCVMTCHG	-0.03888422	-0.00745671	0.03750406	0.03238261	-0.06120248	0.01713544	-0.11271493	0.44275944	0.29403835	1

Figure D – Cross-Section Regression of 2000 Data Prior to Correction and Ramsey RESET Test

```
. regress gini00 pafseg00 medage00 edubs00 pvrty00 fpcvmt00

      Source |       SS           df       MS      Number of obs   =       254
-------------+----------------------------------   F(5, 248)       =     79.45
       Model |  .106462669         5  .021292534   Prob > F        =    0.0000
    Residual |  .066466311       248  .000268009   R-squared       =    0.6156
-------------+----------------------------------   Adj R-squared   =    0.6079
       Total |  .172928981       253  .000683514   Root MSE        =    .01637

------------------------------------------------------------------------------
      gini00 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    pafseg00 |   .1067008   .0107226     9.95   0.000     .0855817    .1278198
    medage00 |   .0028752   .0003135     9.17   0.000     .0022576    .0034927
     edubs00 |   .0855232   .0144094     5.94   0.000     .0571428    .1139035
     pvrty00 |   .3873413   .0258335    14.99   0.000     .3364604    .4382223
    fpcvmt00 |  -2.02e-06   6.35e-07    -3.18   0.002    -3.27e-06   -7.68e-07
       _cons |   .2569877   .0145151    17.70   0.000      .228399    .2855763
------------------------------------------------------------------------------

. ovtest

Ramsey RESET test using powers of the fitted values of gini00
       Ho:  model has no omitted variables
                 F(3, 245) =      8.76
                  Prob > F =      0.0000
```

Figure E - Cross-Section Regression of 2000 Data After Correction and Ramsey RESET Test

```
. regress gini00 pafseg00 medage00 edubs00 pvrty00 fpcvmt00

      Source |       SS           df       MS      Number of obs   =       253
-------------+----------------------------------   F(5, 247)       =     91.31
       Model |  .112200151         5  .02244003    Prob > F        =    0.0000
    Residual |  .060700047       247  .000245749   R-squared       =    0.6489
-------------+----------------------------------   Adj R-squared   =    0.6418
       Total |  .172900198       252  .000686112   Root MSE        =    .01568

------------------------------------------------------------------------------
      gini00 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    pafseg00 |   .0985244   .0104055     9.47   0.000     .0780296    .1190193
    medage00 |   .0033699   .0003171    10.63   0.000     .0027453    .0039945
     edubs00 |   .0970774   .0140027     6.93   0.000     .0694975    .1246573
     pvrty00 |   .4352516   .0266414    16.34   0.000     .3827783    .4877249
    fpcvmt00 |  -2.05e-06   6.08e-07    -3.38   0.001    -3.25e-06   -8.57e-07
       _cons |    .232897   .0147622    15.78   0.000     .2038211    .2619729
------------------------------------------------------------------------------

. ovtest

Ramsey RESET test using powers of the fitted values of gini00
       Ho:  model has no omitted variables
                 F(3, 244) =      1.84
                  Prob > F =      0.1401

.
```

Figure F – Histogram of GINI00 Data

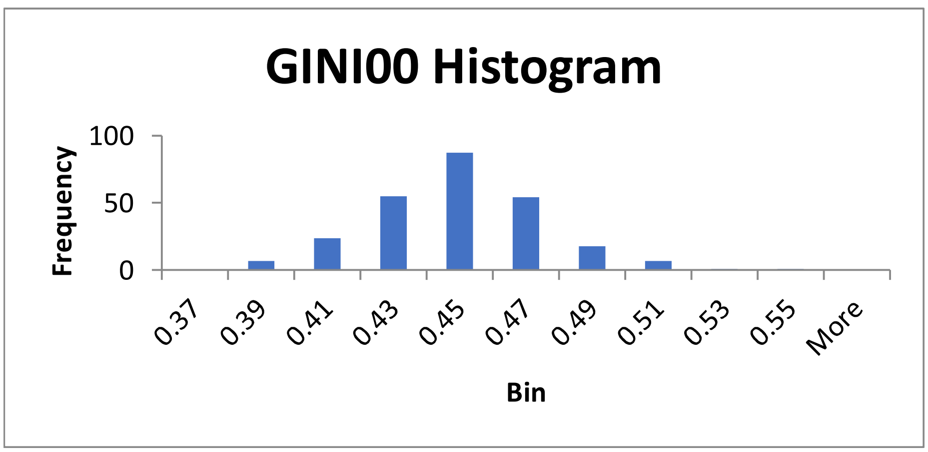


Figure G – Normal Probability Plot of GINI00 Data

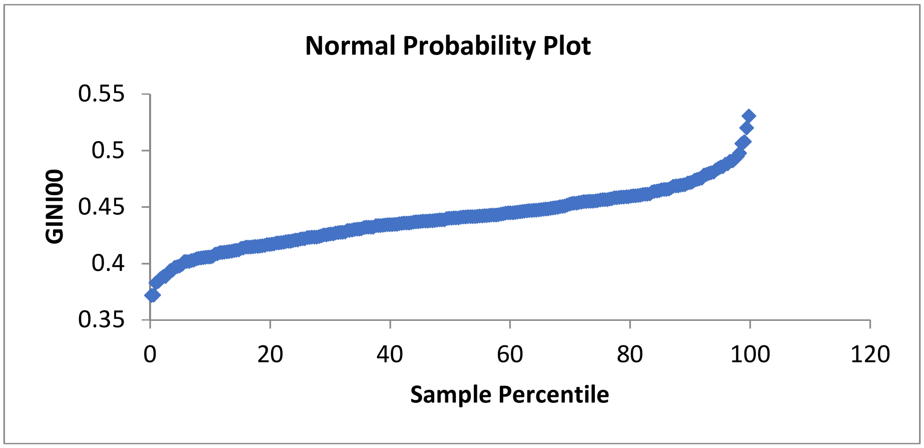


Table H – Standard Residuals and Predicted GINI00 Prior to Removing Outlier (upper left)

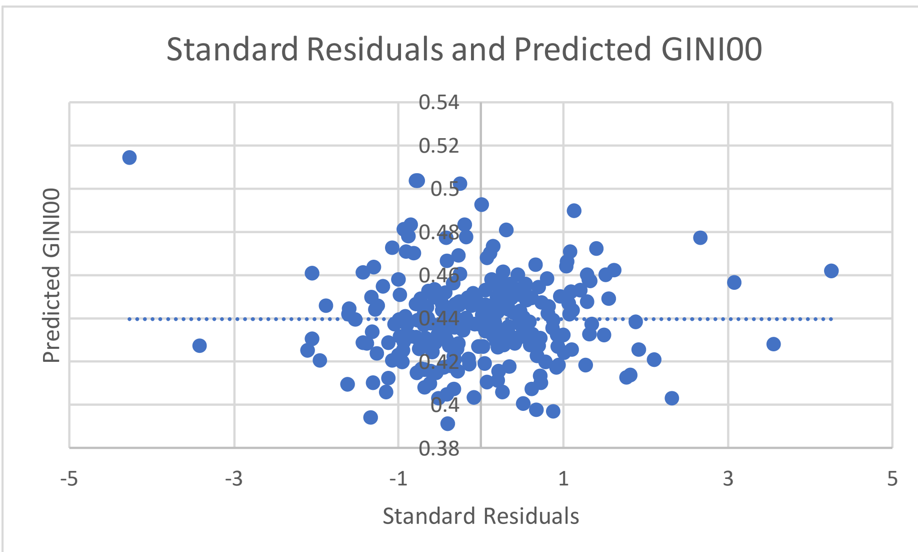


Figure I - Cross-Section Regression of 2010/15 Data and Ramsey RESET Test

```
. regress gini155yr pafseg10 medage15 edubs15 pvrty15 bkpop15 unempr15

      Source |       SS           df       MS      Number of obs   =       254
-------------+----------------------------------   F(6, 247)       =     78.86
       Model |  .085291067         6  .014215178   Prob > F        =    0.0000
    Residual |  .044523259       247  .000180256   R-squared       =    0.6570
-------------+----------------------------------   Adj R-squared   =    0.6487
       Total |  .129814327       253   .0005131    Root MSE        =    .01343

------------------------------------------------------------------------------
   gini155yr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    pafseg10 |   .0818458   .0086163     9.50   0.000      .064875    .0988166
    medage15 |   .0021009   .0002343     8.97   0.000     .0016394    .0025623
     edubs15 |   .1301849   .0120147    10.84   0.000     .1065205    .1538493
     pvrty15 |   .3961513   .0265263    14.93   0.000     .3439046     .448398
     bkpop15 |   .0247684   .0091609     2.70   0.007     .0067249    .0428119
    unempr15 |  -.1644896     .05096    -3.23   0.001    -.2648611    -.064118
       _cons |   .2700343   .0121157    22.29   0.000      .246171    .2938975
------------------------------------------------------------------------------

. ovtest

Ramsey RESET test using powers of the fitted values of gini155yr
       Ho:  model has no omitted variables
                 F(3, 244) =      1.99
                  Prob > F =      0.1165
```

Figure J - Histogram of GINI15 Data

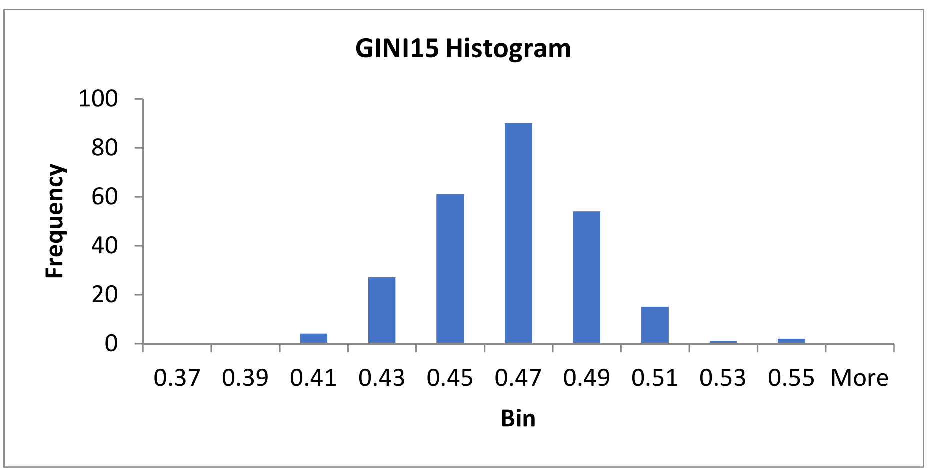


Figure K – Normal Probability Plot of GINI15 Data

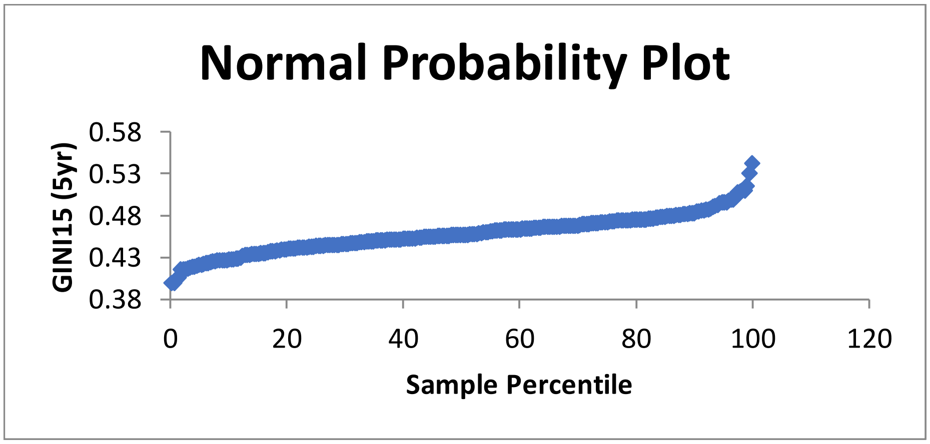


Figure L – Standard Residuals and Predicted GINI15

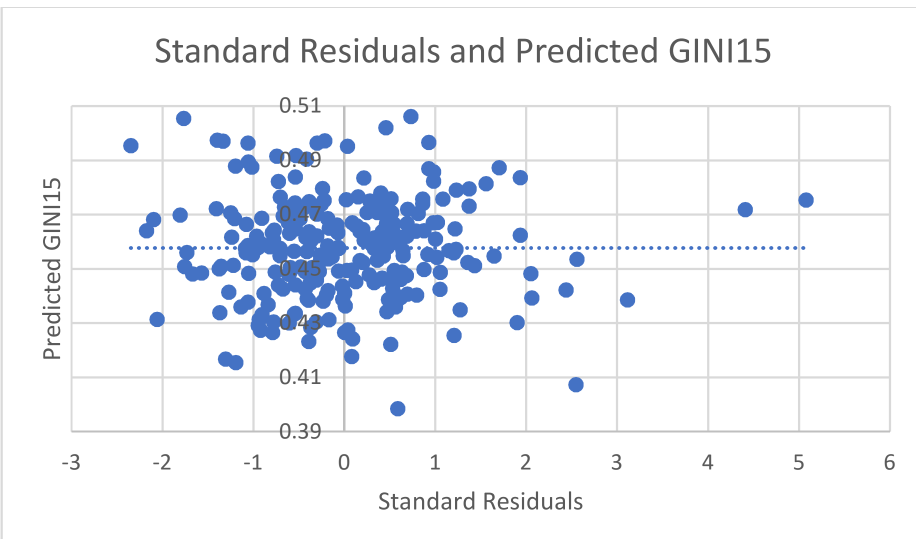


Figure M – Multivariable Regression – GINI Change Between 2000-015

```
. regress ginichg pafsegchg medagechg edubschg unemprchg pvrtychg fpcvmtchg

      Source |       SS           df       MS      Number of obs   =       254
-------------+----------------------------------   F(6, 247)       =      8.83
       Model |  .009156409         6  .001526068   Prob > F        =    0.0000
    Residual |  .042676695       247   .00017278   R-squared       =    0.1767
-------------+----------------------------------   Adj R-squared   =    0.1567
       Total |  .051833104       253  .000204874   Root MSE        =    .01314

------------------------------------------------------------------------------
     ginichg |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   pafsegchg |   .0332829    .017002     1.96   0.051    -.0002046    .0667704
   medagechg |   .0015202   .0005043     3.01   0.003     .0005269    .0025135
    edubschg |    .048705   .0210777     2.31   0.022     .0071901    .0902199
   unemprchg |  -.1069281   .0474541    -2.25   0.025    -.2003945   -.0134617
    pvrtychg |   .2057795   .0365402     5.63   0.000     .1338093    .2777497
   fpcvmtchg |  -2.19e-06   1.03e-06    -2.11   0.036    -4.22e-06   -1.49e-07
       _cons |    .008697   .0028698     3.03   0.003     .0030446    .0143495
------------------------------------------------------------------------------

. ovtest

Ramsey RESET test using powers of the fitted values of ginichg
       Ho:  model has no omitted variables
                 F(3, 244) =      1.23
                  Prob > F =      0.2992
```

Figure N - Histogram of GINI Change Between 2000-2015

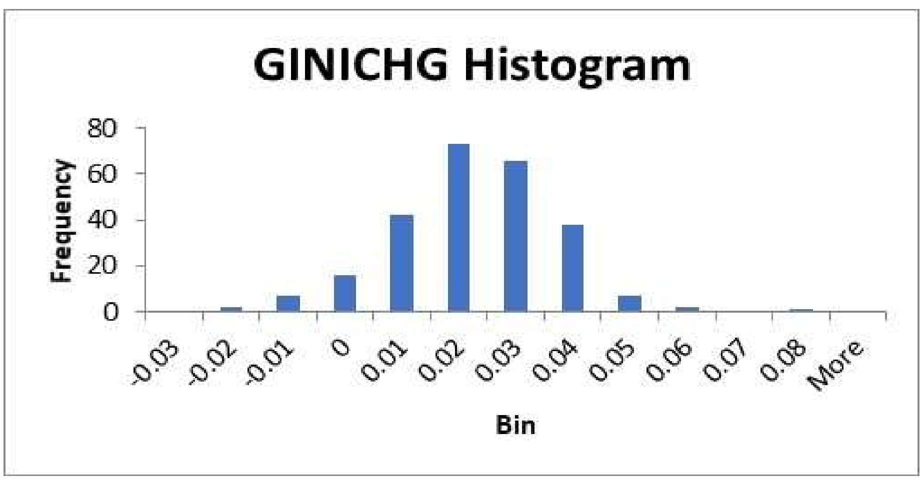


Figure O – Normal Probability Plot of GINI Change Between 2000-2015

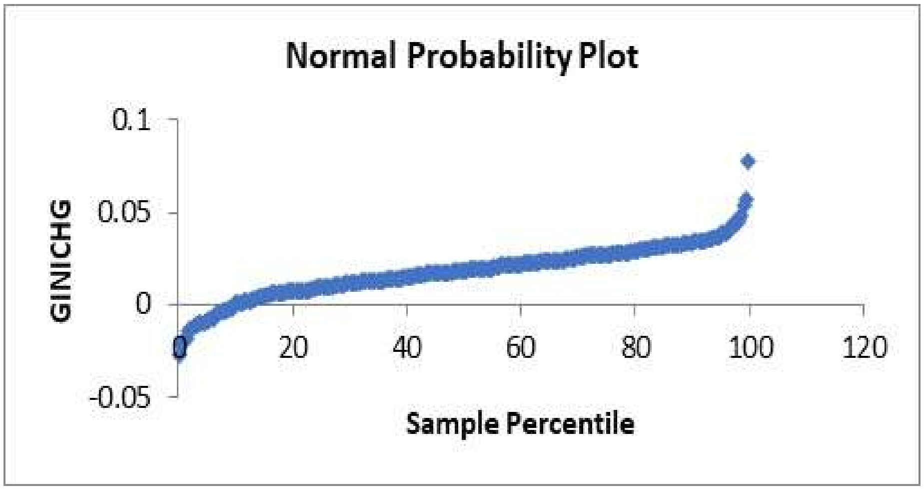


Figure P – Standard Residuals and Predicted GINI Change Between 2000-2015

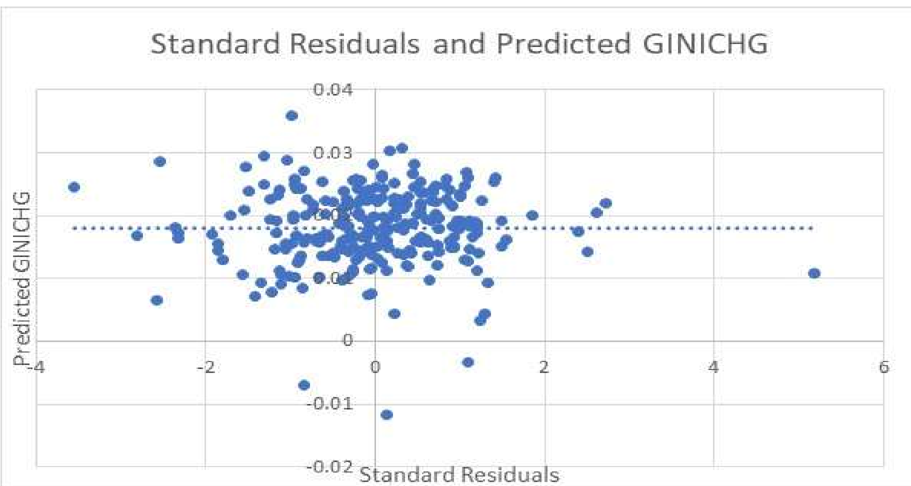


Figure Q – GINI Histogram for 1990, 2000 and 2015 Panel Data

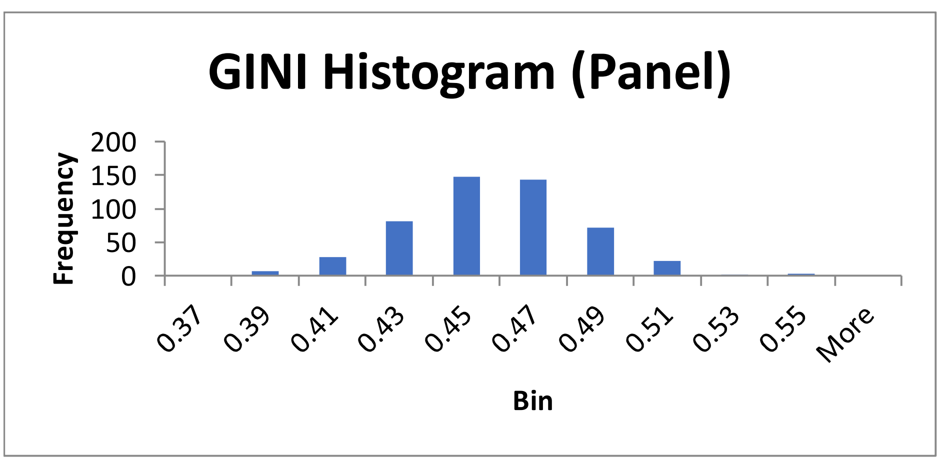


Figure R – STATA Level-Level Demeaned Panel Data Regression (1990-2000-2015)

```
. xtreg gini pafseg pvrty pcvmt edubs medage, fe

Fixed-effects (within) regression               Number of obs     =        216
Group variable: msa                             Number of groups  =         72

R-sq:                                           Obs per group:
     within  = 0.7980                                         min =          3
     between = 0.4659                                         avg =        3.0
     overall = 0.5621                                         max =          3

                                                F(5,139)          =     109.85
corr(u_i, Xb)  = 0.0881                         Prob > F          =     0.0000
```

gini	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
pafseg	.0310325	.0187876	1.65	0.101	-.0061139	.0681788
pvrty	.1920993	.0385382	4.98	0.000	.1159024	.2682963
pcvmt	1.22e-06	6.32e-07	1.93	0.055	-2.70e-08	2.47e-06
edubs	.1402827	.0224832	6.24	0.000	.0958293	.184736
medage	.0020968	.0004899	4.28	0.000	.0011282	.0030653
_cons	.2892053	.0116245	24.88	0.000	.2662215	.312189
sigma_u	.01655827					
sigma_e	.00815245					
rho	.80488855	(fraction of variance due to u_i)				

```
F test that all u_i=0: F(71, 139) = 5.76                     Prob > F = 0.0000

.
```

Figure S – Level-Level Manually Demeaned Panel Data Regression And Ramsey RESET Test (1990-2000-2015)

```
. regress gini pafseg pvrty pcvmt edubs medage
```

Source	SS	df	MS
Model	.036503114	5	.007300623
Residual	.009238306	210	.000043992
Total	.04574142	215	.000212751

```
Number of obs     =        216
F(5, 210)         =     165.95
Prob > F          =     0.0000
R-squared         =     0.7980
Adj R-squared     =     0.7932
Root MSE          =     .00663
```

gini	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
pafseg	.0310326	.0152851	2.03	0.044	.0009007	.0611645
pvrty	.192099	.0313537	6.13	0.000	.1302906	.2539073
pcvmt	1.22e-06	5.14e-07	2.38	0.018	2.09e-07	2.24e-06
edubs	.1402828	.0182918	7.67	0.000	.1042238	.1763419
medage	.0020968	.0003985	5.26	0.000	.0013111	.0028824
_cons	-1.63e-06	.0004513	-0.00	0.997	-.0008913	.000888

```
. ovtest

Ramsey RESET test using powers of the fitted values of gini
       Ho:  model has no omitted variables
                 F(3, 207) =      1.46
                  Prob > F =      0.2266
```

Figure T – Manually Demeaned Level-Level Panel Data Regression Normal Probability Plot (1990-2000-2015)

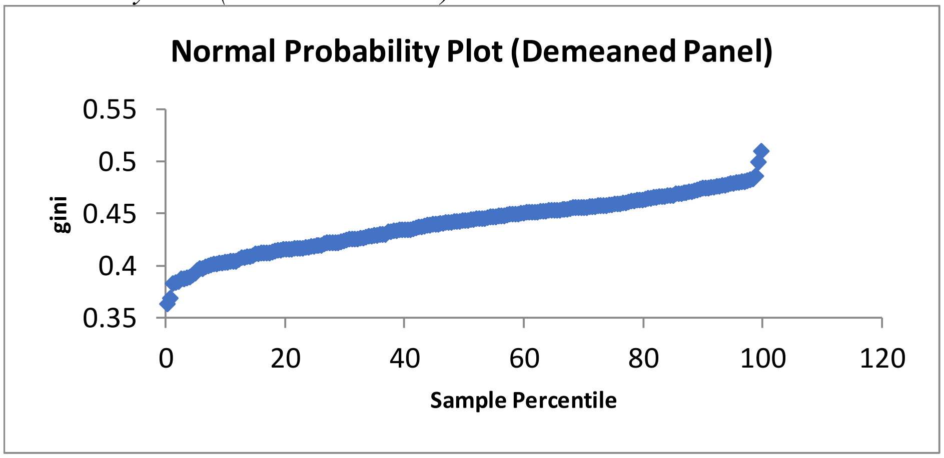


Figure U – Manually Demeaned Level-Level Panel Data Standard Residuals and Predicted GINI (1990-2000-2015)

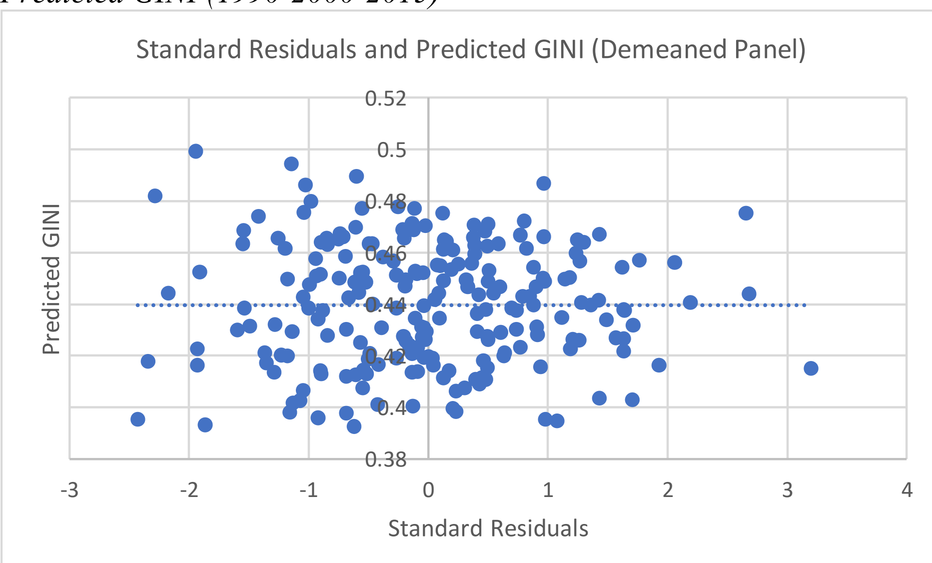


Figure V – Quantile Regression on GINI for 1990-2000-2010 Level-Level Panel Data

```
. sqreg gini pafseg pvrty pcvmt edubs medage, quantiles(10 25 30 35 40 45 50 55 60 65 70 75 90)
(fitting base model)

Bootstrap replications (20)
────┼─── 1 ───┼─── 2 ───┼─── 3 ───┼─── 4 ───┼─── 5
....................

Simultaneous quantile regression                    Number of obs =        216
  bootstrap(20) SEs                                 .10 Pseudo R2 =     0.5487
                                                    .25 Pseudo R2 =     0.5494
                                                    .30 Pseudo R2 =     0.5510
                                                    .35 Pseudo R2 =     0.5477
                                                    .40 Pseudo R2 =     0.5446
                                                    .45 Pseudo R2 =     0.5429
                                                    .50 Pseudo R2 =     0.5388
                                                    .55 Pseudo R2 =     0.5308
                                                    .60 Pseudo R2 =     0.5190
                                                    .65 Pseudo R2 =     0.5063
                                                    .70 Pseudo R2 =     0.4917
                                                    .75 Pseudo R2 =     0.4814
                                                    .90 Pseudo R2 =     0.4459
```

gini	Coef.	Bootstrap Std. Err.	t	P>|t|	[95% Conf.	Interval]
q10						
pafseg	.0397583	.0238503	1.67	0.097	-.0072585	.086775
pvrty	.4333726	.0468237	9.26	0.000	.341068	.5256773
pcvmt	3.59e-07	7.26e-07	0.50	0.621	-1.07e-06	1.79e-06
edubs	.1123394	.027505	4.08	0.000	.0581182	.1665607
medage	.0022136	.0003167	6.99	0.000	.0015893	.0028378
_cons	.2479188	.0127385	19.46	0.000	.2228071	.2730305
q25						
pafseg	.0471072	.0200839	2.35	0.020	.0075154	.0866991
pvrty	.4440869	.0491954	9.03	0.000	.3471068	.541067
pcvmt	6.89e-07	5.15e-07	1.34	0.182	-3.25e-07	1.70e-06
edubs	.0975842	.0210431	4.64	0.000	.0561015	.139067
medage	.0018484	.0002805	6.59	0.000	.0012955	.0024013
_cons	.2659038	.0125641	21.16	0.000	.2411359	.2906717
q30						
pafseg	.051781	.0191781	2.70	0.007	.0139746	.0895873
pvrty	.4398566	.0513043	8.57	0.000	.3387192	.540994
pcvmt	8.85e-07	5.32e-07	1.66	0.098	-1.64e-07	1.93e-06
edubs	.0941413	.0199577	4.72	0.000	.0547982	.1334843
medage	.0017233	.0003877	4.44	0.000	.000959	.0024876
_cons	.2705595	.0146798	18.43	0.000	.2416208	.2994982
q35						
pafseg	.0539666	.017651	3.06	0.003	.0191709	.0887624
pvrty	.4560859	.049846	9.15	0.000	.3578233	.5543485
pcvmt	9.46e-07	6.59e-07	1.43	0.153	-3.54e-07	2.25e-06
edubs	.0956819	.0172209	5.56	0.000	.061734	.1296298
medage	.0015902	.0003974	4.00	0.000	.0008068	.0023735
_cons	.2736002	.0169909	16.10	0.000	.2401056	.3070948
q40						
pafseg	.0495473	.0146156	3.39	0.001	.0207353	.0783594
pvrty	.421823	.0391463	10.78	0.000	.344653	.498993
pcvmt	1.13e-06	6.05e-07	1.87	0.063	-6.23e-08	2.32e-06
edubs	.0865656	.0130784	6.62	0.000	.0607838	.1123473
medage	.0015698	.0004131	3.80	0.000	.0007554	.0023843
_cons	.2834333	.0137631	20.59	0.000	.2563018	.3105649

Figure V – Quantile Regression on GINI for 1990-2000-2010 Level-Level Panel Data (Cont.)

q45						
pafseg	.0420595	.0140896	2.99	0.003	.0142842	.0698347
pvrty	.443442	.0361317	12.27	0.000	.3722146	.5146693
pcvmt	1.59e-06	5.56e-07	2.86	0.005	4.95e-07	2.69e-06
edubs	.09757	.0097535	10.00	0.000	.0783427	.1167974
medage	.0016237	.0004498	3.61	0.000	.0007371	.0025104
_cons	.2751761	.0154339	17.83	0.000	.2447509	.3056013
q50						
pafseg	.0469259	.0143126	3.28	0.001	.0187111	.0751408
pvrty	.4298142	.0341947	12.57	0.000	.3624053	.4972231
pcvmt	1.51e-06	5.37e-07	2.82	0.005	4.56e-07	2.57e-06
edubs	.0938104	.0107024	8.77	0.000	.0727126	.1149082
medage	.0015472	.0004071	3.80	0.000	.0007447	.0023497
_cons	.2814521	.0148651	18.93	0.000	.2521481	.310756
q55						
pafseg	.0447635	.0154811	2.89	0.004	.0142452	.0752818
pvrty	.4295361	.0346712	12.39	0.000	.3611879	.4978843
pcvmt	1.68e-06	4.35e-07	3.86	0.000	8.23e-07	2.54e-06
edubs	.0933778	.0098798	9.45	0.000	.0739015	.1128541
medage	.0014751	.0004557	3.24	0.001	.0005768	.0023734
_cons	.2837554	.0165754	17.12	0.000	.25108	.3164308
q60						
pafseg	.0509954	.0150267	3.39	0.001	.0213729	.0806178
pvrty	.434611	.0284431	15.28	0.000	.3785405	.4906815
pcvmt	1.52e-06	5.39e-07	2.82	0.005	4.59e-07	2.58e-06
edubs	.0877868	.0114221	7.69	0.000	.0652701	.1103035
medage	.0014329	.0005573	2.57	0.011	.0003343	.0025314
_cons	.2875528	.0188338	15.27	0.000	.2504252	.3246804
q65						
pafseg	.0522418	.0158405	3.30	0.001	.0210151	.0834686
pvrty	.4424439	.0262444	16.86	0.000	.3907077	.4941802
pcvmt	1.23e-06	6.36e-07	1.93	0.055	-2.51e-08	2.48e-06
edubs	.0838569	.0100131	8.37	0.000	.0641179	.103596
medage	.0014044	.0005601	2.51	0.013	.0003001	.0025086
_cons	.291903	.0201081	14.52	0.000	.2522634	.3315426
q70						
pafseg	.0468962	.016062	2.92	0.004	.0152328	.0785596
pvrty	.4574047	.0274617	16.66	0.000	.4032687	.5115407
pcvmt	8.47e-07	6.35e-07	1.34	0.183	-4.04e-07	2.10e-06
edubs	.0940059	.0122635	7.67	0.000	.0698305	.1181812
medage	.002043	.0005115	3.99	0.000	.0010347	.0030513
_cons	.2722466	.0182766	14.90	0.000	.2362175	.3082756
q75						
pafseg	.0499587	.0127925	3.91	0.000	.0247404	.0751769
pvrty	.4623984	.0293896	15.73	0.000	.404462	.5203348
pcvmt	3.74e-07	6.52e-07	0.57	0.567	-9.12e-07	1.66e-06
edubs	.0918807	.0112439	8.17	0.000	.0697152	.1140462
medage	.0020238	.0003737	5.42	0.000	.0012871	.0027605
_cons	.2782371	.0153466	18.13	0.000	.247984	.3084902
q90						
pafseg	.0653673	.0096726	6.76	0.000	.0462994	.0844352
pvrty	.4169367	.0461901	9.03	0.000	.325881	.5079923
pcvmt	-7.59e-08	6.68e-07	-0.11	0.910	-1.39e-06	1.24e-06
edubs	.0778994	.0105516	7.38	0.000	.0570987	.0987
medage	.0024246	.0007197	3.37	0.001	.0010059	.0038434
_cons	.2841354	.0316694	8.97	0.000	.2217048	.346566

Figure W - STATA Log-Log Demeaned Panel Data Regression (1990-2000-2015)

```
. xtreg loggini logpvrty logpcvmt logedubs logmedage, fe

Fixed-effects (within) regression               Number of obs     =        216
Group variable: msa                             Number of groups  =         72

R-sq:                                           Obs per group:
     within  = 0.7709                                         min =          3
     between = 0.4392                                         avg =        3.0
     overall = 0.5263                                         max =          3

                                                F(4,140)          =     117.77
corr(u_i, Xb)  = 0.1418                         Prob > F          =     0.0000

------------------------------------------------------------------------------
     loggini |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    logpvrty |   .0698827   .0110426     6.33   0.000     .0480509    .0917145
    logpcvmt |   .0231506   .0129633     1.79   0.076    -.0024784    .0487797
    logedubs |   .0565768   .0161367     3.51   0.001     .0246738    .0884799
   logmedage |   .2283519   .0460779     4.96   0.000     .1372535    .3194503
       _cons |   1.046261   .0578979    18.07   0.000     .9317936    1.160728
-------------+----------------------------------------------------------------
     sigma_u |  .01727774
     sigma_e |  .00861247
         rho |  .80097737   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0: F(71, 140) = 5.87                     Prob > F = 0.0000

.
```

Figure X - Log-Log Manually Demeaned Panel Data Regression and Ramsey RESET Test (1990-2000-2015)

```
. regress loggini logpvrty logpcvmt logedubs logmedage

      Source |       SS           df       MS      Number of obs   =       216
-------------+----------------------------------   F(4, 211)       =    172.78
       Model |  .034723037         4  .008680759   Prob > F        =    0.0000
    Residual |  .010601165       211  .000050242   R-squared       =    0.7661
-------------+----------------------------------   Adj R-squared   =    0.7617
       Total |  .045324203       215   .00021081   Root MSE        =    .00709

------------------------------------------------------------------------------
     loggini |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    logpvrty |   .0720178   .0090316     7.97   0.000     .0542141    .0898215
    logpcvmt |   .0237839   .0106905     2.22   0.027     .0027101    .0448577
    logedubs |   .0658127   .0125892     5.23   0.000      .040996    .0906294
   logmedage |   .2013842   .0356326     5.65   0.000     .1311427    .2716258
       _cons |   .0000748   .0004825     0.15   0.877    -.0008763    .0010258
------------------------------------------------------------------------------

. ovtest

Ramsey RESET test using powers of the fitted values of loggini
       Ho:  model has no omitted variables
                  F(3, 208) =      2.52
                  Prob > F =      0.0587
```

Figure Y - Manually Demeaned Level-Level Panel Data Regression Normal Probability Plot (1990-2000-2015)

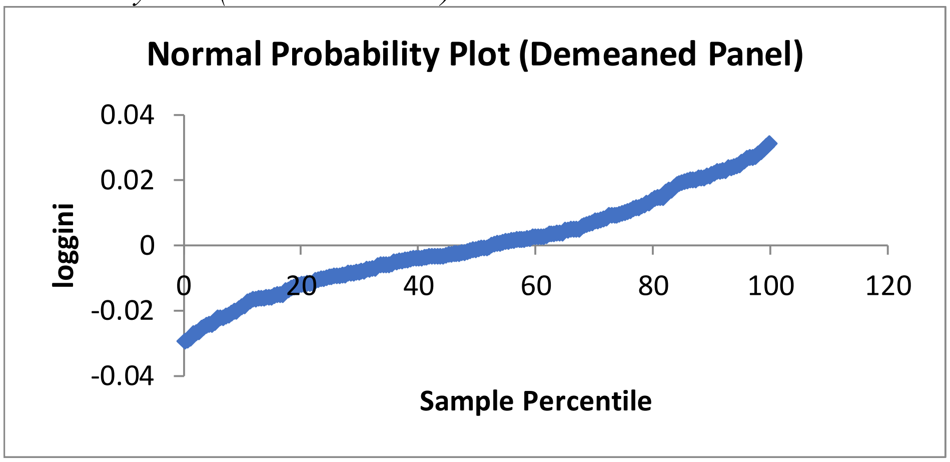


Figure Z - Manually Demeaned Log-Log Panel Data Standard Residuals and Predicted LOGGINI (1990-2000-2015)

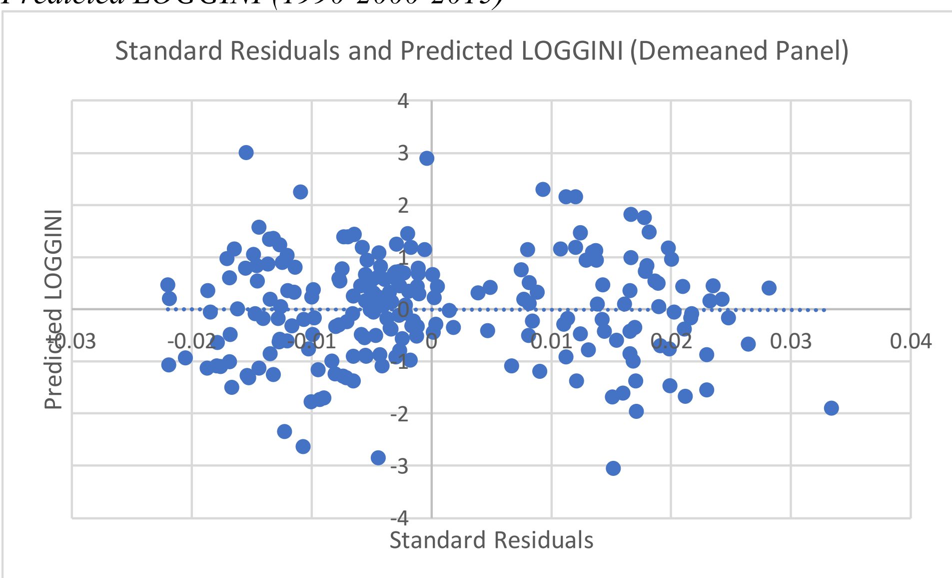


Figure AA - Quantile Regression on GINI for 1990-2000-2010 Log-Log Panel Data

```
. sqreg loggini logpvrty logpcvmt logedubs logmedage, quantiles (10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90)
(fitting base model)

Bootstrap replications (20)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
....................

Simultaneous quantile regression                    Number of obs =        216
  bootstrap(20) SEs                                 .10 Pseudo R2 =     0.5287
                                                    .15 Pseudo R2 =     0.5292
                                                    .20 Pseudo R2 =     0.5270
                                                    .25 Pseudo R2 =     0.5281
                                                    .30 Pseudo R2 =     0.5309
                                                    .35 Pseudo R2 =     0.5318
                                                    .40 Pseudo R2 =     0.5277
                                                    .45 Pseudo R2 =     0.5195
                                                    .50 Pseudo R2 =     0.5082
                                                    .55 Pseudo R2 =     0.4966
                                                    .60 Pseudo R2 =     0.4846
                                                    .65 Pseudo R2 =     0.4721
                                                    .70 Pseudo R2 =     0.4595
                                                    .75 Pseudo R2 =     0.4444
                                                    .80 Pseudo R2 =     0.4253
                                                    .85 Pseudo R2 =     0.4040
                                                    .90 Pseudo R2 =     0.3813
```

loggini	Coef.	Bootstrap Std. Err.	t	P>|t|	[95% Conf.	Interval]
q10						
logpvrty	.1772656	.0216765	8.18	0.000	.1345354	.2199958
logpcvmt	-.0037065	.0152437	-0.24	0.808	-.0337559	.026343
logedubs	.0839442	.0154305	5.44	0.000	.0535266	.1143618
logmedage	.1725382	.0167397	10.31	0.000	.1395397	.2055366
_cons	1.060227	.0707343	14.99	0.000	.9207905	1.199663
q15						
logpvrty	.1701339	.0154127	11.04	0.000	.1397514	.2005165
logpcvmt	.0069795	.0153205	0.46	0.649	-.0232214	.0371803
logedubs	.0765204	.0094918	8.06	0.000	.0578094	.0952314
logmedage	.1570775	.021718	7.23	0.000	.1142654	.1998896
_cons	1.063624	.0457714	23.24	0.000	.9733963	1.153852
q20						
logpvrty	.1685097	.0105504	15.97	0.000	.147712	.1893075
logpcvmt	.0161055	.0110121	1.46	0.145	-.0056024	.0378134
logedubs	.0702686	.0105251	6.68	0.000	.0495208	.0910164
logmedage	.1496065	.0232686	6.43	0.000	.1037377	.1954753
_cons	1.05298	.0274682	38.33	0.000	.9988326	1.107127
q25						
logpvrty	.157516	.0106174	14.84	0.000	.1365862	.1784457
logpcvmt	.026021	.0096819	2.69	0.008	.0069353	.0451067
logedubs	.0638445	.0082061	7.78	0.000	.047668	.0800209
logmedage	.1468988	.0219084	6.71	0.000	.1037114	.1900862
_cons	1.042298	.0390439	26.70	0.000	.9653319	1.119264
q30						
logpvrty	.1565847	.0093333	16.78	0.000	.1381862	.1749832
logpcvmt	.0275059	.0104169	2.64	0.009	.0069714	.0480404
logedubs	.0602046	.0072385	8.32	0.000	.0459355	.0744736
logmedage	.1326524	.0205583	6.45	0.000	.0921263	.1731784
_cons	1.065914	.0437054	24.39	0.000	.9797592	1.152069
q35						
logpvrty	.1586975	.0082711	19.19	0.000	.1423928	.1750022
logpcvmt	.0354337	.010242	3.46	0.001	.015244	.0556234
logedubs	.0635222	.007874	8.07	0.000	.0480005	.0790439
logmedage	.1342871	.0248955	5.39	0.000	.0852113	.183363
_cons	1.027603	.0554587	18.53	0.000	.9182791	1.136927
q40						
logpvrty	.1580068	.0070576	22.39	0.000	.1440943	.1719193
logpcvmt	.0381481	.0116439	3.28	0.001	.0151947	.0611014
logedubs	.0678058	.0069559	9.75	0.000	.0540939	.0815177
logmedage	.13843	.0256497	5.40	0.000	.0878675	.1889924
_cons	1.006942	.0477545	21.09	0.000	.9128051	1.101079
q45						
logpvrty	.1578509	.0060456	26.11	0.000	.1459333	.1697685
logpcvmt	.0347694	.0136349	2.55	0.011	.0078913	.0616475
logedubs	.0598831	.0068934	8.69	0.000	.0462944	.0734718
logmedage	.132849	.0282958	4.70	0.000	.0770703	.1886277
_cons	1.041411	.0497199	20.95	0.000	.9434	1.139423

Figure Y - Quantile Regression on GINI for 1990-2000-2010 Log-Log Panel Data (Cont.)

q50						
logpvrty	.1550121	.0069816	22.20	0.000	.1412496	.1687747
logpcvmt	.0353101	.0136043	2.60	0.010	.0084923	.062128
logedubs	.0613289	.006788	9.03	0.000	.0479478	.0747099
logmedage	.1151421	.0262433	4.39	0.000	.0634094	.1668748
_cons	1.068872	.0600683	17.79	0.000	.9504607	1.187283
q55						
logpvrty	.153788	.0097114	15.84	0.000	.1346442	.1729319
logpcvmt	.031066	.0111646	2.78	0.006	.0090575	.0530744
logedubs	.0636354	.0070259	9.06	0.000	.0497855	.0774854
logmedage	.1224421	.027017	4.53	0.000	.0691842	.1757
_cons	1.074131	.0661188	16.25	0.000	.9437931	1.204469
q60						
logpvrty	.1485947	.0099779	14.89	0.000	.1289257	.1682638
logpcvmt	.0264092	.0117693	2.24	0.026	.0032088	.0496096
logedubs	.0575491	.0055278	10.41	0.000	.0466523	.0684459
logmedage	.1144561	.025844	4.43	0.000	.0635107	.1654016
_cons	1.121139	.0731376	15.33	0.000	.9769654	1.265313
q65						
logpvrty	.1534838	.0071441	21.48	0.000	.1394008	.1675668
logpcvmt	.0204413	.0136487	1.50	0.136	-.006464	.0473465
logedubs	.0593678	.0064185	9.25	0.000	.0467152	.0720204
logmedage	.1311647	.025658	5.11	0.000	.0805859	.1817435
_cons	1.112303	.073819	15.07	0.000	.9667854	1.25782
q70						
logpvrty	.1495969	.0090044	16.61	0.000	.1318467	.1673471
logpcvmt	.0133385	.0127575	1.05	0.297	-.0118099	.0384869
logedubs	.0564645	.0050682	11.14	0.000	.0464737	.0664552
logmedage	.1281444	.0279478	4.59	0.000	.0730517	.1832372
_cons	1.154888	.071353	16.19	0.000	1.014232	1.295544
q75						
logpvrty	.1501547	.0095709	15.69	0.000	.1312878	.1690216
logpcvmt	.0129625	.016295	0.80	0.427	-.0191593	.0450843
logedubs	.0562962	.0064387	8.74	0.000	.0436038	.0689886
logmedage	.141127	.0386592	3.65	0.000	.0649193	.2173348
_cons	1.137226	.0755201	15.06	0.000	.9883554	1.286096
q80						
logpvrty	.144586	.0092285	15.67	0.000	.126394	.1627779
logpcvmt	.0186239	.0167742	1.11	0.268	-.0144425	.0516903
logedubs	.0528018	.0093472	5.65	0.000	.0343761	.0712276
logmedage	.1266436	.0590215	2.15	0.033	.0102963	.242991
_cons	1.151128	.0914709	12.58	0.000	.9708142	1.331442
q85						
logpvrty	.1402751	.0143912	9.75	0.000	.1119061	.1686441
logpcvmt	.0073157	.017161	0.43	0.670	-.0265131	.0411446
logedubs	.0578508	.0089844	6.44	0.000	.0401402	.0755614
logmedage	.1032303	.0606391	1.70	0.090	-.0163057	.2227662
_cons	1.233456	.0896359	13.76	0.000	1.056759	1.410152
q90						
logpvrty	.1558871	.0198127	7.87	0.000	.1168309	.1949432
logpcvmt	-.0027036	.0168995	-0.16	0.873	-.036017	.0306099
logedubs	.0618827	.0085622	7.23	0.000	.0450043	.0787611
logmedage	.1646244	.0706537	2.33	0.021	.0253469	.3039019
_cons	1.158339	.1001278	11.57	0.000	.9609601	1.355718

.