Whitepaper No. 16006

Measuring Income Inequality in Farm States: Weaknesses of The Gini Coefficient

April 28, 2016

Madelyn McGlynn, *Gail Werner-Robertson Fellow*
Faculty Mentor: Dr. Ernest Goss

Creighton University | Heider College of Business
INSTITUTE FOR ECONOMIC INQUIRY

Executive summary

As the gap between the highest and lowest incomes increases rapidly within the United States, many people have become more convinced of its importance. General interest has sparked academic studies of the issue in many fields of expertise, such as psychology and economics. These studies conclude income inequality results from factors such as education gaps, an aging labor force and the geographic distribution of populations. Much attention has been directed toward determining why income inequality exists and the effects of it; however, little research has been done on the methods used to measure income inequality.

This paper focuses on the primary method used to measure income inequality, the Gini index, and questions its validity because of the inherently inconsistent data used as its main input. The focus of my research is farm income, which is difficult to compare accurately with other income. Using the five US states with the largest farm sectors, I statistically analyze the relationship between the Gini index and such commonly cited factors behind inequality as the prevalence of farm income. The existence of a large farm income altered the estimated Gini coefficient for four of the five states. This result should reduce confidence in the usefulness of the Gini index in measuring income inequality for states with large farm sectors. Further, it calls into question the reliability of the index itself.

Introduction

In the last few years, income inequality has been widely debated. It has worked its way into almost all political discussions, and will undoubtedly be discussed during the presidential debates. Income inequality is most commonly measured using the Gini index, which was developed in 1912 by Italian social scientist Corrado Gini. It measures discrepancies in income as a percentage of a whole, where one equals complete income inequality (a score of one indicates that a single individual has all of the income and everyone else has none; a score of zero indicates income is distributed equally among all people).

The Gini index relies heavily on measurements and definitions of income. The equation's main input is, of course, income. However, measurements of income are subject to many discrepancies including tax credits, depreciation of capital assets, and subsidies. It is almost impossible to determine how best to measure income in the United States. Certain basic principles apply to everyone filing tax returns, but researchers make many exceptions in accordance with their accounting methods and their differential treatment of different types of income-earning activities. Such differences in how countries or states measure income can massively skew the data.

Income for various industries is often even incomparable. For this reason, the focus of my paper is on income inequality, as reflected in the Gini index, in states with the highest farm income as a percentage of total income. Farm income is particularly hard to standardize because of the extent of tax credits and subsidies in the industry. Farmers also often invest in large capital assets that amortize over time, creating an illusion of lost revenue. Often farmers have to reinvest their revenue into their land and equipment to keep their businesses going. Accounting methods consider reinvestment a business expense and so do not include it in the measurement of income. This method might be good for measuring farm income; however, it is not very helpful for comparing farm income with income in other industries.

I chose the five states with the highest farm income as a percentage of total income (South Dakota, Nebraska, Iowa, North Dakota, and Idaho) to determine whether the differences in income-data collection methods significantly impacted the Gini coefficients for these states. I compared farm income to total income to determine whether the Gini coefficient is meaningful when calculated using farming income. In my measurement of farm income, I included crop farming and ranching.

In *Earnings from Inequality and Mobility in the United States: Evidence from Social Security Data since 1937,* Kopcuk claims that since 1953, income inequality, as measured by the Gini index, has risen sharply in the United States. In the past five years, income inequality has risen more gradually across the country. Figure 1 illustrates that the United States' Gini coefficient has risen from 0.378 in 1986 to 0.411 in 2013. While many studies have attempted to discover the factors that have caused US inequality and consequences of the inequality, they have begun with different assumptions and ended with different conclusions.

Figure 1[1]: US Gini coefficient, 1986–2013

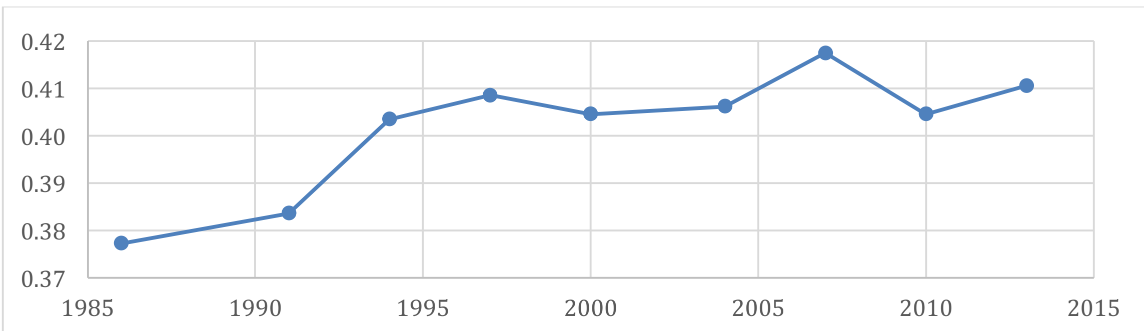
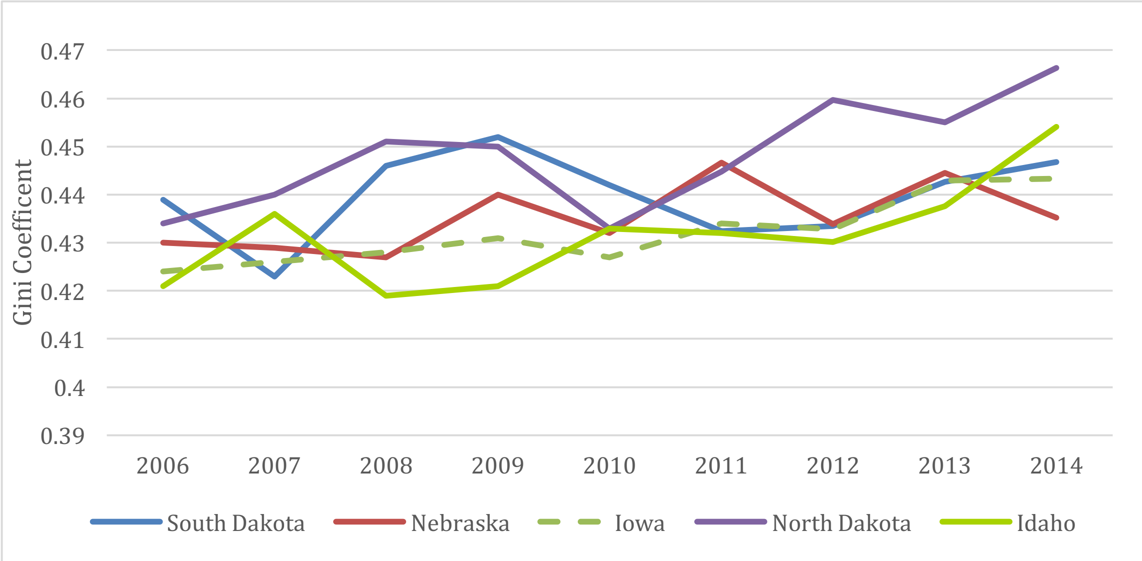


Figure 2 demonstrates the rise in each of the five high-farm-income states' Gini index in recent years. Notice the large increases and decreases present on the state level and how different states are affected by economic changes throughout the years to varying degrees. The model includes South Dakota because it has the highest farm-income percentage and ends with Idaho as it has the lowest share of farm income out of the five states.

[1] Raw data obtained from World Bank estimates.

Figure 2: Increase in Gini index in farm states, 2006–14



Past Research

Academics have written on income inequality since the mid-twentieth century, focusing on varying aspects and reaching varying conclusions. Some have chosen to focus on psychological and social effects of income inequality, and others have looked for the underlying causes with economic models.

Many scholars, such as Lisa Berkman (2014), take a psychological approach, searching for the consequences of the massive amount of inequality on personal well-being. Research conducted by Michael Norton (2011) showed that Americans tend to underestimate income inequality. Citizens think the income of the richest members of society and the poorest members of society are closer than they actually are. These scholars have proven that income inequality affects psychology, but this paper focuses on economic concerns.

The Gini index is derived from the estimated difference between the slopes along the Lorenz curve. Economists have tried to correct for these errors by using statistical models that isolate aspects of the equation. Scholars such as Joseph Gastwirth (1972) are willing to admit the results are skewed: “the method used by the Census Bureau often leads to estimates which are outside of mathematically possible bounds.” But very few call into question the data-collection methods as problems in and of themselves.

Perhaps the most important agreement of these works is that the Gini index measures "relative inequality." This makes the debate on data collection more important than before because it means the debate relies on consistent data collection even more.

Popular modern theories focus less on the actual equation Gini created and more on the reasons why inequality is increasing. These studies start with different assumptions, which makes comparing them difficult. Despite their varied theories about what factors influence income inequality the most, it is clear income inequality is rising in the United States. According to Janet Yellen (chair of the Board of Governors of the Federal Reserve), since 1973 the top 10 percent of American incomes increased by about 30 percent. The bottom 50 percent of workers' real income only rose by about 5 percent. This difference is significant and changes the dynamics of the American economy, and it is not going away. Researchers tend to conclude that income inequality is exacerbated by differences in education, an aging labor force, and the concentration of populations.

My research argues that other factors can be added to the list of factors that affect the Gini coefficient. When an industry deviates from the normal income tax structure, it may have an effect on the index coefficient. I seek to show that high concentration of farm income decreases the Gini coefficient.

Theory and Data Description

To find the Gini coefficient one must rely on tax data from federal and state income tax reports. I collected the Gini coefficients for the counties of each of the five states from the Census Bureau's 2013 information. Farm data measuring 2013 incomes comes from the Bureau of Economic Analysis. Farm income here includes both the cultivation of crops and cattle ranching. However, a difficulty arises because most farmers' wealth is not measured using the traditional income tax model.

Compared to other producers, farmers have more of their wealth invested in capital assets such as expensive mechanical implements and land. In addition, a lot of farm expenses and tax deductions offset profits and thus taxable income. This poses a problem for the Gini coefficient's reliability for farms because income is the main input for deriving it. In fact, the Department of Agriculture, in its statistics on farm income and wealth, does not measure the success of farms in terms of their income, but rather their wealth. To measure wealth, they use cash receipts and ownership of capital. This has such a large effect on the data that it actually changes the ranking of which states profit the most from farming, meaning that the department's method implies that just because certain states have the highest reported farm income does not mean they benefit the most from farming.

I compared the income inequality of individual counties of each of the five farming states considered here (South Dakota, Nebraska, Iowa, North Dakota, and Idaho) to each of the other counties within their state first, and collected information about each county to determine possible factors that could cause Gini index discrepancies. Then I aggregated this information to compose a state profile by arranging the counties into five quintiles, with quintile 1 containing the counties with the lowest Gini coefficient, suggesting the lowest income inequality, and quintile 5 containing the counties with the highest Gini coefficient, suggesting the highest income inequality.

Observations collected on commonly cited factors for income inequality include education (measured by the percentage of the population with a high school degree or higher), median age, and population density. I compared and averaged the data from the counties in each quintile to more easily see the relationship between the Gini index and each factor. Then I examined the impact farm income as a percentage of total income has on the Gini coefficient. After this I compared the states on the same basis.

Data Analysis

I further examined the information described above by using regression correlation analysis. Figure 3 demonstrates how the commonly evaluated factors relate to the Gini coefficient in each state. It shows that all of the factors have a low impact on the Gini coefficient. Notice the wide variations. This demonstrates that the cited factors are not as relevant as researchers tend to believe and that they do not fully carry over across state lines. The results of the regression analysis are displayed in figures 4–6.

Figure 3[2]: Correlation of factors to the Gini coefficient by state

[2] Raw data obtained from the US Census Bureau.

Factor	Correlation Coefficient				
	South Dakota	Nebraska	Iowa	North Dakota	Idaho
Farm income as percentage of total income	0.10	0.26	-0.07	-0.06	-0.31
Median age	-0.07	0.05	-0.08	-0.11	0.17
High school degree or higher	0.11	0.02	-0.34	0.13	-0.32
Population density	-0.14	-0.18	0.19	-0.08	0.13

Further regression analysis led to a deeper understanding of the relationships between these factors. Figure 6 highlights the results. Working with five states and five sets of data required a thorough, multi-test approach. Using Idaho as the base state, across the states I analyzed farm income as a share of total income in relation to the Gini index. This approach produced initially puzzling results. Farm income as a percentage of total income when compared on a county level showed a significantly positive relationship, meaning as farm income as a percentage of total income increased so did the Gini index for that area, as shown in figure 6.

Then I evaluated the states individually and in comparison with each other. As expected, South Dakota and Nebraska both demonstrated that as farm income as a share of total income increased among their counties, so did the counties' Gini coefficient. The opposite was true of Idaho. It presented a negative relationship between farm income as a share of total income and the Gini index, meaning as farm income as a share of total income increased, the Gini coefficient decreased. North Dakota and Idaho both showed negative relationships between their counties' farm income and Gini coefficients, but the results were not statistically significant. Figure 5 demonstrates the net impact of farm income as a percentage of total income on each state's Gini coefficient.

Figure 4: Variable explanations for regression results

Variable Name	Definition of Variable
Gini	2013 Gini coefficients for each county in the five state area.
Farm income as a share of total income	2013 Farm income as a percentage of the total income for all counties In all 5 states.
Median Age	Median age of the population for all counties
Educational Attainment	Percentage of the population with a high school or higher educational degree
Population Density	County population per square mile
Nebraska	A binary variable equal to 1.0 for all Nebraska counties, and equal to 0 for all non-Nebraska counties
South Dakota	A binary variable equal to 1.0 for all South Dakota counties, and equal to 0 for all non-South Dakota counties
Iowa	A binary variable equal to 1.0 for all Iowa counties, and equal to 0 for all non-Iowa counties
North Dakota	A binary variable equal to 1.0 for all North Dakota counties, and equal to 0 for all non-North Dakota counties
INE	Nebraska's farm income as a percent of total income times the indicator variable, Nebraska.
ISD	South Dakota's farm income as a percent of total income times the indicator variable, South Dakota.
IIA	Iowa's farm income as a percent of total income times the indicator variable, Iowa

Figure 5: Summary of the effects of the Gini index

State	Effect of the Gini Coefficient	Average Farm Income as a Percentage of Whole Income	Net Impact
Nebraska	0.0716	23.4504%	1.6795%
South Dakota	0.0393	21.2396%	0.8345%

Iowa	-0.0100	13.0175%	-0.1297%
North Dakota	-0.0132	15.0718%	-0.1989%
Idaho	-0.0971	9.7126%	-0.9436%

Figure 6: Regression analysis

SUMMARY OUTPUT	
Regression statistics	
Multiple R	0.4278
R-squared	0.1830
Adjusted R-squared	0.1543
Standard error	0.0315
Observations	355

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	12	0.0762	0.0063	6.3828	3.06199E-10
Residual	342	0.3402	0.0010		
Total	354	0.4164			

	Coefficients	*Standard Error*	*t stat*	*P-value*
Idaho—used for comparison	0.4465	0.0159	28.0616	0.0000
Farm income as share of total income	-0.0971	0.0445	-2.1838	0.0297
Median age	-0.0001	0.0003	-0.3169	0.7515
Educational attainment	-0.0566	0.0353	-1.6002	0.1105
Population density	0.0000	0.0000	-0.2753	0.7832
Nebraska	-0.0163	0.0093	-1.7510	0.0808
South Dakota	0.0130	0.0097	1.3473	0.1788
Iowa	-0.0020	0.0090	-0.2227	0.8239
North Dakota	0.0277	0.0097	2.8617	0.0045
INE	0.1688	0.0518	3.2550	0.0012
ISD	0.1364	0.0533	2.5583	0.0109
IIA	0.0872	0.0612	1.4234	0.1555
IND	0.0839	0.0610	1.3752	0.1700

Conclusion

These data show that farm income has an impact on the Gini coefficient in four of the five states I tested. If the Gini index was operating properly, the strong presence of an industry should not have swayed the results. Thus this study undermines the usefulness of the Gini index for measuring income inequality for states heavily dependent upon farm income.

This has large implications when considering differences in tracking income inequality. The presence of the farming industry sways the outcome of the most common measure of income inequality. Income inequality relies on consistent income data, which is not usually available. Farm income is not the only distinctively measured form of income. The Gini index will not have reliable results for any such income data. This means comparing county inequality is not very useful when the data are collected differently.

In a time when income inequality is such a highly controversial issue, it is important for policy makers and the public to understand that the measures used to describe income inequality are flawed. Following a flawed system to reach important decisions leads to flawed decisions. To reach more relevant decisions, economists, policy makers, and the public need to consider alternative ways of evaluating income inequality.

The existence of a relationship between farm income and the Gini index should motivate further investigation. Perhaps this difference is caused by the inclusion of both crop cultivation and cattle ranching in farm-income measurements. This inquiry should be further tested in future research.

References

Bee, Adam. 2012. "Household Income Inequality within U.S. Counties: 2006–2010." *American Community Survey Briefs* (February). US Census Bureau.

Berkman, Lisa, S. V. Subramanian, and Ichiro Kawachi. 2014. "Income Inequality." In *Social Epidemiology*, edited by Lisa Berkman, Ichiro Kawachi, and Maria Glymour. Oxford: Oxford University Press. Print.

Gastwirth, Joseph L. 1972. "The Estimation of the Lorenz Curve and Gini Index." *Review of Economics and Statistics* 54, no. 3: 306–16.

Kopczuk, Wojciech, Emmanuel Saez, and Jae Song. 2010. "Earnings Inequality and Mobility in the United States: Evidence from Social Security Data since 1937." *Quarterly Journal of Economics* 125, no. 1: 91–128.

Nebraska Department of Agriculture. 2013. "Nebraska Agriculture Fact Card." *Nebraska Agriculture 5*–30.

Norton, M. I., and D. Ariely. 2011. "Building a Better America—One Wealth Quintile at a Time." *Perspectives on Psychological Science* 6, no. 1: 9–12.

Thompson, Eric, Bruce Johnson, and Anil Giri. 2012. "The 2010 Economic Impact of the Nebraska Agricultural Production Complex." Report No. 192, Department of Agricultural Economics, University of Nebraska-Lincoln.

United States Department of Agriculture. 2016."Farm Income and Wealth Statistics." Accessed April 12. http://www.ers.usda.gov/data.

United States Environmental Protection Agency. 2015. "Agriculture." October 1. Accessed April 12, 2016. https://www.epa.gov/agriculture.

United States Internal Revenue Service. 2007. "Reporting Farm Income and Expenses." Fact Sheets FS-2007-20. Accessed November 24, 2015.

Weinberg, Daniel. 2011. "U.S. Neighborhood Income Inequality in the 2005–2009 Period." *American Community Survey Reports* (October). US Census Bureau.

World Bank. 2015. "GINI Index (World Bank Estimate)." Accessed April 12, 2016. http://data.worldbank.org.

Yellen, Janet. 2006. Speech to the Center for the Study of Democracy, Irvine, CA. November 6.