Proceedings of the 2nd International Conference on Financial Technology and Business Analysis
DOI: 10.54254/2754-1169/65/20231586

Analyze the Determinants of the Gini Index in the United States: An Econometrics Approach

Jiahe Yu[1,a,*]

[1]*The Albany Academies, New York, 12208, United States*
a. 24yuj@albanyacademies.org
**corresponding author*

Abstract: Inequality has always been a topic of discussion among scholars. In this essay, one crucial measure of inequality—the Gini Index—is analyzed to find the influencing factors within the labor market. By running four regressions independently, the labor force changed from being positively related to being inversely correlated, and employment in services changed from being negatively associated to being positively related. The empirical results proved that the labor force, employment in agriculture as a percent of total employment, and unemployment rate are negatively related to the Gini index, and the rest of the variables (employment in service as a percent of total employment and population with tertiary degree) have a positive correlation. More labor force, employment in agriculture, and unemployment rate are accompanied by less Gini index. Take education level as an example; those at the top have better access to education and skill development in economies with significant income inequality, making them more likely to engage in higher-paying service or knowledge-based industries. Meanwhile, those with lower incomes may have fewer opportunities for schooling and are more likely to work in low-wage agricultural or manual labor employment.

Keywords: Gini Index, Inequality, Econometrics, United States

1. Introduction

Inequality has received considerable scholarly attention across various fields as a concept and a social phenomenon. Inequality comes in many forms, including health, social, educational, and economic inequality.

Health disparities, frequently connected with socioeconomic position and access to healthcare, have been a significant research focus. Sir Michael Marmot, for example, has investigated the social determinants of health and how they lead to health inequities [1]. This field of study emphasizes the complexities of inequality and its impact on well-being. Social inequality refers to gaps in access to opportunities and resources depending on social identities such as race, gender, ethnicity, and age. Kimberlé Crenshaw pioneered intersectional approaches, emphasizing the compounding impacts of different forms of social inequality [2]. A recent study has examined how these disparities connect and reinforce one another [3,4]. Researchers have examined achievement discrepancies among various demographic groups, differences in access to high-quality education, and educational attainment in order to understand educational inequality [5, 6]. Understanding educational disparity is necessary for developing policies to promote social mobility. Economic inequality, especially income disparity, is the kind of inequality that continues to be one of the most visible and often

DOI: 10.54254/2754-1169/65/20231586

debated. Researchers have focused on income and wealth disparities within and within countries, considering issues such as income distribution, poverty, and resource access. Prominent researchers such as Thomas Piketty have highlighted the growing concentration of wealth in the hands of a few, while others such as Joseph Stiglitz have highlighted the harmful effects of income inequality on economic growth and social stability [7, 8].

The Gini coefficient, which was created by Italian statistician Corrado Gini, quantifies how resources, particularly income, are allocated in an economy when that distribution diverges from perfect equality. [9]. Beginning with the poorest, a Lorenz curve plots the percentages of total income received vs the cumulative recipients. The Gini index, which, according to World Bank estimates, spans from 0 to 100, with 100 representing perfect inequality and 0 indicating ideal equality, measures the distance between the Lorenz curve and a fictitious line of absolute equality by taking the maximum area under the line as a percentage. Research has examined the limitations and nuances of the Gini index. Scholars such as Milanovic have highlighted that while it provides a helpful snapshot of inequality, it may not capture all aspects of distributional dynamics [10]. Nevertheless, the Gini index remains invaluable in academic and policy contexts for tracking and addressing income inequality globally.

Moreover, researchers also tried to decompose the Gini index in different ways. Many studies focused on the contribution of income inequality with different parts of the composition and the population subgroups [11]. Mussini's research, as an example, decomposed the index by income source and population subgroup simultaneously with a multi-decomposition methodology and matrix approach. The study by Heshmatim also included causal factors and other unit characteristics in addition to those factors. Moreover, another existing research proposes three influencing factors of the changes in the Gini index: structural, natural inequality, and interactive effects [12]. However, the influence on income inequality from labor market factors is crucial yet not discussed often with an econometrics approach [13].

In this paper, one difference from the previous studies is that the variables incorporated extensive labor market measures such as unemployment rate, labor force, employment structure (employment in services and agriculture as % of total employment), and population with tertiary education. Using a linear regression model, this paper will discuss the correlation of the Gini index to these factors from the United States labor market over the last thirty-five years.

2. Experiment Design

2.1. Data Source

Data on the population with tertiary degrees is from the OECD; unemployment, labor force, employment structure, and Gini index are from the World Bank.

2.2. Trend Analysis

The unemployment rate as a percentage of the total labor force can show the stage of an economy in the business cycle since it reflects corporations' ability to hire people. The U.S. started from a growing economy in the late 1980s but soon entered a recession corresponding to the decreasing then increasing unemployment rate (see Figure 1). The recession from 1990-1991 turned out to be the result of a consumption shock caused by lower income, which lasted longer and took more time to recover compared to recessions caused by other factors [14]. Thus, the unemployment rate has fallen since the early 1990s. The United States saw significant economic development in the late 1990s and early 2000s, with unemployment rates falling below 4%. The pace at which people lost their jobs increased dramatically during the Great Recession of 2008. It is most likely due to temporary structural problems brought on by insurance issues and the secular decline in labor mobility that the

DOI: 10.54254/2754-1169/65/20231586

ensuing recovery, which persisted into the early 2010s, was worse in this recession than in prior ones. [15]. From the mid-2010s to early 2020, the U.S. sustained stable economic growth, with the unemployment rate reaching a historically low point, hovering around 3.5%-4%. Then followed the COVID-19 pandemic that slowed the economy by closing borders and trading limitations, yet the recovery phase was much shorter than the previous recessions [16].

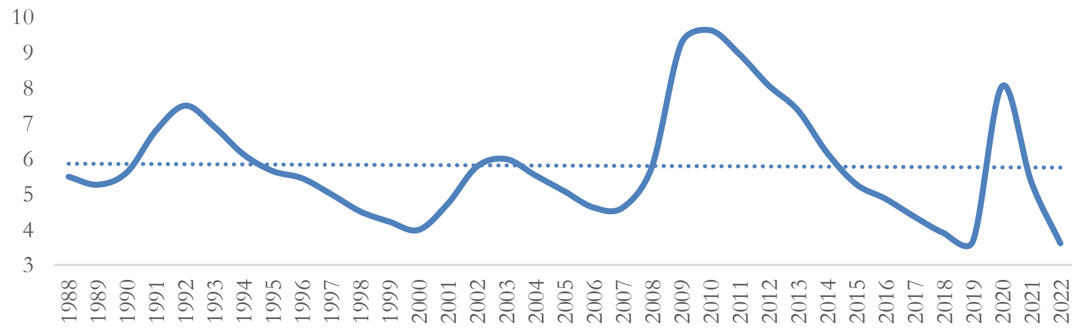

Figure 1: Unemployment rate from 1988 to 2022

Data source: World Bank / Photo credit: Original

The labor force consists of people aged fifteen and over who are eligible to apply for jobs to produce goods and services. It is limited to employed or actively looking for a job. In other words, those who do unpaid jobs, caregivers, and students are not included. Seasonal workers come and go throughout the year, causing changes in the size of the labor force. The United States' growing population and women's higher labor force participation rate are both contributing factors to the labor force's overall stable growth trend (see Figure 2) [17].

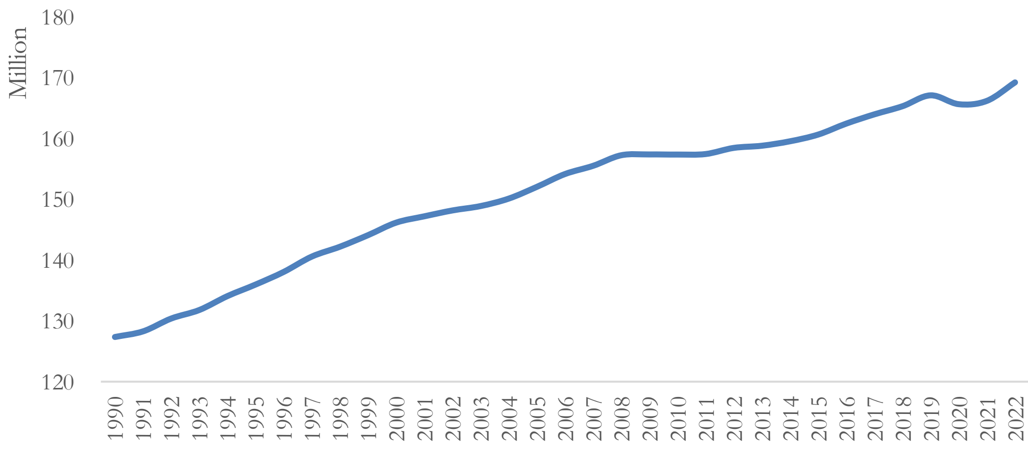

Figure 2: Labor force from 1990 to 2022

Data source: World Bank / Photo credit: Original

Agriculture played a relatively limited influence in the US economy regarding employment at the start of this period. Agricultural work had already fallen due to mechanization and technical developments in earlier decades. Agriculture employment continued to fall progressively in the 2000s as productivity rose, lessening the demand for physical labor. By 2021, the agriculture sector employed a relatively tiny proportion of the US labor force, with most agricultural occupations becoming specialized and mechanized. Agriculture remained an important sector in terms of food production, but it no longer employed a considerable number of people (see Figure 3).

DOI: 10.54254/2754-1169/65/20231586

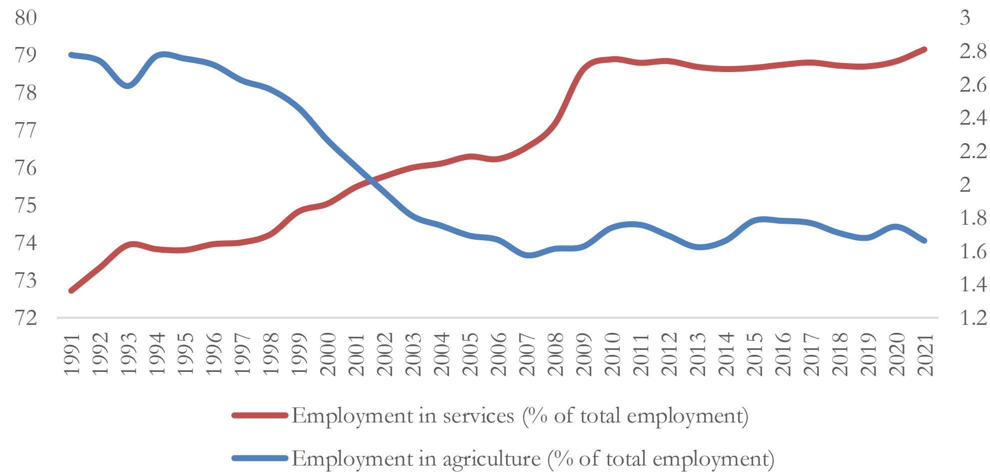

Figure 3: Employment structure from 1991 to 2021

Data source: World Bank / Photo credit: Original

Over time there has been a trend in the proportion of people attaining tertiary education with a slight deviation in 2020 and 2021 due to the impact of the pandemic. In today's job market, the increase in university students can be attributed to population growth and the growing significance of higher education. Another reason for the steady growth trend is that undergraduate students may want to pursue a higher degree to find a better job (see Figure 4).

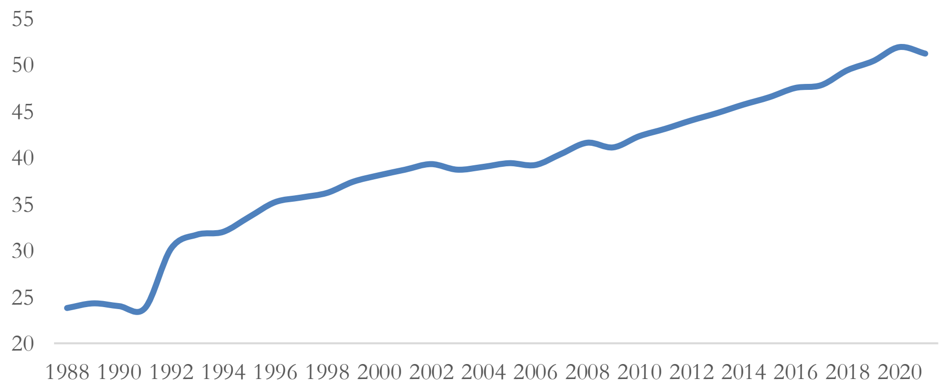

Figure 4: Population with tertiary education from 1988 to 2021

Data source: OECD / Photo credit: Original

In the late '80s through the turn of the millennium, the Gini index consistently climbed, signifying a marked increase in income inequality. This era, notably the 1990s, witnessed substantial economic expansion. This trend continued into the early 21st century, underscoring a deepening income divide. Complex factors, including tax policies, globalization, and technological shifts, played their part in exacerbating this situation. The financial crisis of 2008 led to a temporary blip in the Gini index from 2008 to 2010 due to the reduction in wealth among high-income individuals and the rollout of government stimulus efforts. In the subsequent decade, income inequality persisted upward. Factors such as stagnant wages for many workers and the aggregation of wealth among a select few underscored this trend (see Figure 5).

DOI: 10.54254/2754-1169/65/20231586

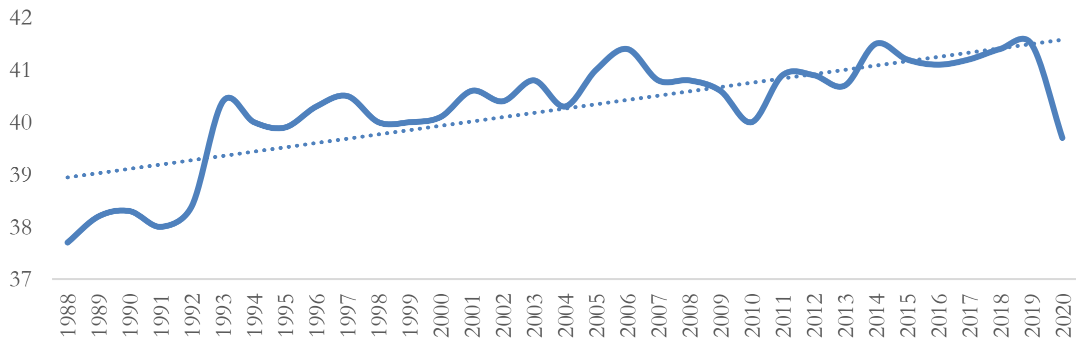

Figure 5: Gini index from 1988 to 2020

Data source: World Bank / Photo credit: Original

2.3. Regression Setting

α_0 is always the constant. For the variables, Lf stands for the total labor force, S means employment in services as a percent of all employment, A is employment in agriculture as a percent of all employment, Ur stands for the unemployment rate as a percent of total labor force, and E stands for education with tertiary degrees. Tertiary degrees encompass bachelor's, master's, and doctoral degrees are defined as the focus. The data utilized here are people in the age group of 25 to 30.

$$Gini = \alpha_0 + \alpha_1 \cdot Lf + \alpha_2 \cdot S + \alpha_3 \cdot A + \alpha_4 \cdot Ur + \alpha_5 \cdot E \tag{1}$$

3. Empirical Results

The first regression $Gini = \alpha_0 + \alpha_1 \cdot Lf$ shows that the labor force alone is positively correlated with the Gini index, which means a higher labor force comes with a higher Gini index and vice versa.

Table 1: Regression results

VARIABLES	(1) OLS Gini	(2) OLS Gini	(3) OLS Gini	(4) OLS Gini
Labor force, in million	0.0565***	0.0681*	-0.0849	-0.1488**
	(0.0089)	(0.0358)	(0.0514)	(0.0714)
Service		-0.1816	0.5749**	0.5526**
		(0.1719)	(0.2516)	(0.2492)
Agriculture		-0.4376	-0.9063*	-1.4384**
		(0.5679)	(0.4856)	(0.6367)
UR			-0.3569***	-0.3494***
			(0.0981)	(0.0971)
Education				0.0894
				(0.0703)
Constant	31.9464***	44.9910***	13.2720	22.0367
	(1.3424)	(10.0933)	(12.0523)	(13.7589)
Observations	31	30	30	30
R-squared	0.5796	0.5471	0.7039	0.7226

The following two variables added are employment structure in service and agriculture as a percent of total employment. The equation is now $Gini = \alpha_0 + \alpha_1 \cdot Lf + \alpha_2 \cdot S + \alpha_3 \cdot A$. Both employment in agriculture and service are negatively correlated with the Gini index, expressing that if the Gini is higher, there would be more income disparity, and employment in agriculture and service would be lower. Also, it is worth noticing that after inputting these two variables, the correlation between the labor force and Gini decreased from 0.0556 to 0.0681. Therefore, employment structure is proven more relevant to Gini than the labor force.

The final variable imputed is the unemployment rate in the United States with the equation $Gini = \alpha_0 + \alpha_1 \cdot Lf + \alpha_2 \cdot S + \alpha_3 \cdot A + \alpha_4 \cdot Ur + \alpha_5 \cdot E$. As shown in Table 1, the correlation between the unemployment rate and the Gini is negative. In other words, the Gini index is lower when the unemployment rate is higher. Specifically, there could be two extreme scenarios–shared prosperity or poverty. With higher unemployment, most people would be jobless, which tends to be shared poverty. On the contrary, when unemployment is lower, a more significant portion of the population will work, leading to shared prosperity and a lower Gini index.

Finally, all variables are incorporated into the last regression, adding the population of people with tertiary education. $Gini = \alpha_0 + \alpha_1 \cdot Lf + \alpha_2 \cdot S + \alpha_3 \cdot A + \alpha_4 \cdot Ur + \alpha_5 \cdot E$. The regression result presents that when the number of people who completed college with any degree increased, so did the Gini index. In this case, more educated people getting paid higher causes higher income disparity. In addition, the positive correlation between the Gini index and the labor force slightly decreased as education was a variable. For employment in service, a higher percentage reveals that capitalism is more prevalent. As a result, employees may be more severely exploited, which leads to a higher Gini index.

4. Conclusion

To sum up, as an indicator of income disparity, the Gini index has many determinants. In this paper, variables, including the labor force, unemployment structure, unemployment rate, and education level, are analyzed using linear regression to find the correlation of these variables to the Gini index. By running four regressions independently, the labor force changed from being positively related to being inversely correlated, and employment in services changed from being negatively associated to being positively related. The empirical results proved that the labor force, employment in agriculture as a percent of total employment, and unemployment rate are negatively related to the Gini index, and the rest of the variables (employment in service as a percent of total employment and population with tertiary degree) have a positive correlation. More labor force, employment in agriculture, and unemployment rate are accompanied by less Gini index. Take education level as an example; those at the top have better access to education and skill development in economies with significant income inequality, making them more likely to engage in higher-paying service or knowledge-based industries. Meanwhile, those with lower incomes may have fewer opportunities for schooling and are more likely to work in low-wage agricultural or manual labor employment.

However, these data are not the only impacting factors for sure. There are various ways to interpret and understand the Gini coefficient, and this paper only provides a view focused on the labor market perspective.

References

[1] Marmot, M. (2005). Social Determinants of Health Inequalities. The Lancet, 365(9464), 1099-1104.
[2] Crenshaw, K. (1989). Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. University of Chicago Legal Forum, 139-167.
[3] Collins, P. H. (2015). Intersectionality's definitional dilemmas. Annual Review of Sociology, 41, 1-20.

DOI: 10.54254/2754-1169/65/20231586

[4] *McCall, L. (2005). The Complexity of Intersectionality. Signs: Journal of Women in Culture and Society, 30(3), 1771-1800.*

[5] *Reardon, S. F. (2013). The Widening Academic Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations.*

[6] *G. J. Duncan & R. J. Murnane (Eds.), Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances (pp. 91-116). Russell Sage Foundation.*

[7] *Piketty, T. (2013). Capital in the Twenty-First Century. Harvard University Press.*

[8] *Stiglitz, J. E. (2012). The Price of Inequality: How Today's Divided Society Endangers Our Future.*

[9] *Catalano, M. T., Leise, T. L., & Pfaff, T. J. (2009). Measuring resource inequality: The Gini coefficient. Numeracy, 2(2), 4.*

[10] *Milanovic, B. (2011). Global Inequality: From Class to Location, from Proletarians to Migrants. Global Policy, 2(2), 198-208.*

[11] *Shao, B. (2021). Decomposition of the Gini index by income source for aggregated data and its applications. Computational statistics, 36, 2135-2159.*

[12] *Wan, G. H. (2001). Decomposing changes in the Gini index by factor components. Applied Economics Letters, 8(1), 1-3.*

[13] *Mocan, H. N. (1999). Structural unemployment, cyclical unemployment, and income inequality. Review of Economics and Statistics, 81(1), 122-134.*

[14] *Blanchard, O. (1993). Consumption and the Recession of 1990-1991. The American Economic Review, 83(2), 270-274.*

[15] *Pissarides, C. A. (2013). Unemployment in the great recession. Economica, 80(319), 385-403.*

[16] *Kozicki, B., & Gornikiewicz, M. (2020). Unemployment rate in Poland and USA during COVID-19 pandemic: a case study.*

[17] *Lee, M., & Mather, M. (2008). US labor force trends (Vol. 63, No. 2). Washington, DC: Population Reference Bureau.*