Accepted Article

*Friederike Seifert**

Received: 23 March 2021
Accepted: 16 May 2022

The Income-Inequality Relationship within US Metropolitan Areas 1980-2016

Abstract:

Economic growth might both increase and decrease income inequality, also at the city level. This paper examines the income-inequality relationship within US metropolitan areas and finds that it changes over time. A higher average income per capita level was associated with a lower inequality level in earlier years, but this association vanished later. For the 1980-2000 panel, increases in the average income per capita are associated with decreases in inequality. In contrast, increases in the average income per capita are associated with increases in inequality in the 2006-2016 panel. The obtained results hint at polarization resulting from technological change substituting middle-skill routine tasks.

Keywords: Inequality; income; metropolitan areas; United States.

JEL classification: D31; O18; R11.

La Relación Ingresos-Desigualdad en las Áreas Metropolitanas de los EE.UU.

Resumen:

El crecimiento económico puede tanto aumentar como disminuir la desigualdad de ingresos, también al nivel de ciudades. Este artículo examina la relación entre ingresos y desigualdad en las áreas metropolitanas de los EE.UU. y descubre que cambia con el tiempo. Un mayor nivel de ingreso per cápita medio se asoció con un menor nivel de desigualdad en los primeros años, pero esta asociación desapareció posteriormente. Para el panel de 1980—2000, los aumentos del ingreso per cápita medio se asocian con disminuciones de la desigualdad. En cambio, un aumento del ingreso per cápita medio se asocia con un aumento de la desigualdad en el panel 2006—2016. Los resultados obtenidos insinúan a una polarización resultante del cambio tecnológico que sustituye a las tareas rutinarias de cualificación media.

Palabras clave: Desigualdad; ingresos; áreas metropolitanas; Estados Unidos.

Clasificación JEL: D31; O18; R11.

1. Introduction

The income-inequality relationship has been a question of debate since the seminal work of Kuznets, who proposed the Kuznets curve: inequality first increases and then decreases with increasing national income (Kuznets, 1955). However, the income-inequality relationship at the city level does not necessarily follow

* TU Dresden: Faculty of Business and Economics and "Friedrich List" Faculty of Transport and Traffic Sciences, Dresden, Germany.
Leibniz Institute of Ecological Urban and Regional Development, Dresden, Germany.
Dresden Leibniz Graduate School, Dresden, Germany.
Corresponding author: friederike.seifert@tu-dresden.de

that at the national level. Some channels from the national level, such as credit market mechanisms and redistribution policies, do not translate directly to the city level (Glaeser et al., 2009; Royuela et al., 2019). The latter is characterized by more in- and out-migration and the less political maneuver room it has. Other factors level out at the national level, such as segregation. At the same time, income inequality is most visible and prominent in cities due to the spatial proximity of different income levels (Partridge & Weinstein, 2013). Comparatively little is still known about the income-inequality relationship at the city level, mainly due to data limitations. To close this gap, this study assesses this relationship within US metropolitan statistical areas (MSAs) from 1980-2016.

Few studies have analyzed the income-inequality relationship at this scale. For US MSAs, a negative income-inequality relationship has been found for 1980 and 2000: higher income levels are associated with lower inequality levels in MSAs based on cross-section regressions (Glaeser et al., 2009). For European regions, determinants of inequality at the regional level have been analyzed using annual panels over the 1990s and 2000s. These studies find a positive income-inequality relationship: income increases are associated with inequality increases (Rodríguez-Pose & Tselios, 2009; Castells-Quintana et al., 2015).

To assess these opposing results further, the present paper employs both cross-section and fixed effects (FEs) panel regression analyses for one geographic unit (MSAs) over several decades (1980-2016). This procedure provides a consistent background for comparing the results for different techniques and years. The analyses are based on two distinct data sets. The first is an annual panel over 2006-2016 using data from the American Community Surveys (ACSs) (Ruggles et al., 2018; US Census Bureau, n.d.-a). The second is a decennial panel over 1980-2000 using US Census data (Manson et al., 2017; Ruggles et al., 2018). This paper thereby expands the time horizon for local-level studies on the income-inequality relationship up to 2016.

This paper finds that the income-inequality relationship changes over time. A higher average income per capita level is associated with a lower within-MSA inequality level in the earlier years. However, this association stopped being statistically significant in 2000 and remains insignificant for all the following years. For the 1980-2000 panel, average income per capita increases are accordingly associated with decreases in inequality. In contrast, an increase in average income per capita is associated with an increase in inequality in the 2006-2016 panel. The income-inequality relationship changed direction. These results are robust to the use of various inequality measures.

This change in sign might be due to differences in MSA delineations and time dimensions across the two panels. However, it could also originate from qualitative changes in the income-inequality relationship over time, potentially reflecting globalization and specialization. Notably, this study finds hints for polarization in line with the Autor & Dorn (2013) hypothesis of technological change substituting middle-skill routine tasks. However, these explanations cannot be completely distinguished with the data sets at hand. Thus, further research is required.

The following section reviews in greater detail the literature on how income and inequality are linked at the city level. Section 3 describes the data sources used and provides the empirical framework. Section 4 presents the cross-section results on the income-inequality relationship, while section 5 details the panel results. Section 6 presents robustness checks using alternative inequality measures and section 7 discusses potential reasons for the change in sign of the income-inequality relationship. Section 8 concludes.

2. City-Level Links between Income and Inequality

Increases in mean income might both increase and decrease inequality depending on the circumstances. The Kuznets curve theory hypothesizes that the income-inequality relationship follows an inverted-U-shaped curve: inequality first increases and then decreases with increasing income (Kuznets, 1955). The N-shape hypothesis augments this theory, stating that after a certain point, inequality starts increasing again with income for highly-developed economies (Conceição & Galbraith, 2001; Castells-Quintana et al., 2015).

Trade and labor market phenomena such as specialization, technological change substituting middle-skill routine tasks, deunionization, and flexible labor market regulations might lead to a positive income-inequality relationship. They might engender both economic growth and increased inequality (Rigby & Breau, 2008; Autor & Dorn, 2013; Partridge & Weinstein, 2013). On the contrary, theories about residential segregation and disamenities such as crime and sociopolitical unrest predict a negative association: inequality decreases with income. For instance, residential segregation is associated with lower economic growth and higher inequality (Li et al., 2013; Florida & Mellander, 2015). Crime and sociopolitical unrest hinder economic growth while leading to and reinforcing inequality, resulting in vicious circles (Glaeser et al., 2009; Partridge & Weinstein, 2013).

These theories consider implicitly a medium- to long-run perspective where agents can adjust to a new situation. To the best of the author's knowledge, no explicitly short-run theory about the income-inequality relationship exists. However, the relationship between income and inequality might differ between the short, medium, and long run. Transmission channels differ in their manifestation rapidity, with purely economic factors typically realizing faster than sociopolitical ones (Halter et al., 2014).

An MSA's population size, education level, and the sectoral structure of its economy influence within-MSA inequality as well (Glaeser et al., 2009). Studies on the city size-inequality relationship typically identify a positive relationship: larger cities are ceteris paribus more unequal (Glaeser et al., 2009; Baum-Snow & Pavan, 2012; Castells-Quintana et al., 2020). Education proxies for differences in skills and the degree of specialization, which leads to dispersed incomes (Glaeser et al., 2009). Higher education levels are associated with higher levels of inequality (Perugini & Martino, 2008; Glaeser et al., 2009). Shifts in the economy's sectoral structure might influence inequality due to differences in the associated income structure (Bolton & Breau, 2012; Castells-Quintana et al., 2015). Deindustrialization tends to increase inequality (Bolton & Breau, 2012; Partridge & Weinstein, 2013).[1]

Several studies on MSA-level determinants of inequality exist, but they only employ cross-section regression analyses. A higher median income level is related to a lower level of inequality for 1980 and 2000 (Glaeser et al., 2009). Similarly, a higher average income level is associated with lower income inequality for 2010 when wage inequality is controlled for (Florida & Mellander, 2016). Higher income per capita growth appears to lead to lower end-of-period inequality in 1990 (Bhatta, 2001). For 11 OECD countries, including the US, a trend towards a negative income-inequality relationship at the city level emerges in pooled cross-sections over 2000—2014 with year and country FEs (Castells-Quintana et al., 2020). Higher average income per capita is associated with lower inequality. However, this association is not always statistically significant. Furthermore, the results hint at an inverse-U-shape income-inequality relationship (Castells-Quintana et al., 2020).

Cross-sections typically only capture the situation at one point in time and hence incorporate all the past influences leading to differences across MSAs (Forbes, 2000; Partridge, 2005). In this sense, they have rather a long-term perspective. This contrasts with panel studies that assess how changes in income levels result in inequality changes for a given MSA (Partridge, 2005; Atems, 2013). Panel studies have rather a short- to medium-term perspective. Therefore, cross-section and panel results are not directly comparable (Atems, 2013). This study will use both techniques, cross-section and panel analyses, to gain a complete picture of the income-inequality relationship at hand.

Some studies of European regions have analyzed the income-inequality relationship in annual panel frameworks with city FEs. Income per capita changes appear to be positively related to inequality changes for European NUTS I and II regions over 1995-2000 based on FEs, random effects, and GMM techniques (Rodríguez-Pose & Tselios, 2009). A U-shaped relationship is found over the 1993-2011 period for NUTS I regions but only when using the GINI as inequality measure (Castells-Quintana et al., 2015). The latter interprets this as inequality having increased more in regions with higher relative increases in income, hence a positive income-inequality relationship as well (Castells-Quintana et al., 2015). However, these results are not directly transferable to US MSAs due to the differing labor market and institutional

[1] The demographic and racial composition of an MSA might influence inequality levels as well. However, the related variables have proved not statistically significant in the regressions. They have been omitted from the presented analysis for clarity.

context, influencing the income-inequality relationship. Furthermore, MSAs provide smaller and more homogeneous regions than the NUTS regions. The present study's sample size is also larger, with up to 399 MSAs available for the analysis.

This paper expands the time horizon for studies on the income-inequality relationship by using data spanning from 1980 to 2016, although with gaps and changes in between as detailed in the next section. This enables assessing whether this relationship changed over time.

3. Data Sources and Empirical Framework

The study unit of this paper is the MSA.[2] MSAs are suitable units for studying regional economic activity and income inequality, as they encompass both the city core and suburbs related through commuting (Madden, 2000). MSAs form a functional economic unit encompassing production and consumption activities (Madden, 2000). Although the concept of MSAs has changed little over time, their county composition does change. A major change in MSA delineations occurred in 2013. Data within the 1990 MSA delineations are available for 1980, 1990, and 2000. Data within the 2013 MSA delineations are available from 2006 onward.

Hence, this study employs two distinct data sets: one with decennial data for 1980-2000 and one with annual data from 2006-2016. For the 2006-2016 data set, the data stems from the 1-year ACSs collected by the US Census Bureau. The data for all the main variables was retrieved from FactFinder (US Census Bureau, n.d.-a). This data includes the pretax household income GINI at the MSA level. Figure 1 shows the MSAs and their respective inequality levels in 2010. All ACS income variables are for the past 12 months prior to the interview moment, which is not publicly disclosed (US Census Bureau, 2009; Peters, 2013; IPUMS-USA, n.d.-b). This paper converts all original income variables into 2010 US-$ using the conversion factors provided by the Integrated Public Use Microdata Series USA (IPUMS) to adjust for inflation (IPUMS-USA, n.d.-b). Table 1 presents descriptive statistics. The resulting panel data set consists of 399 MSAs and 11 years. It is unbalanced due to slight further delineation changes over the time period.

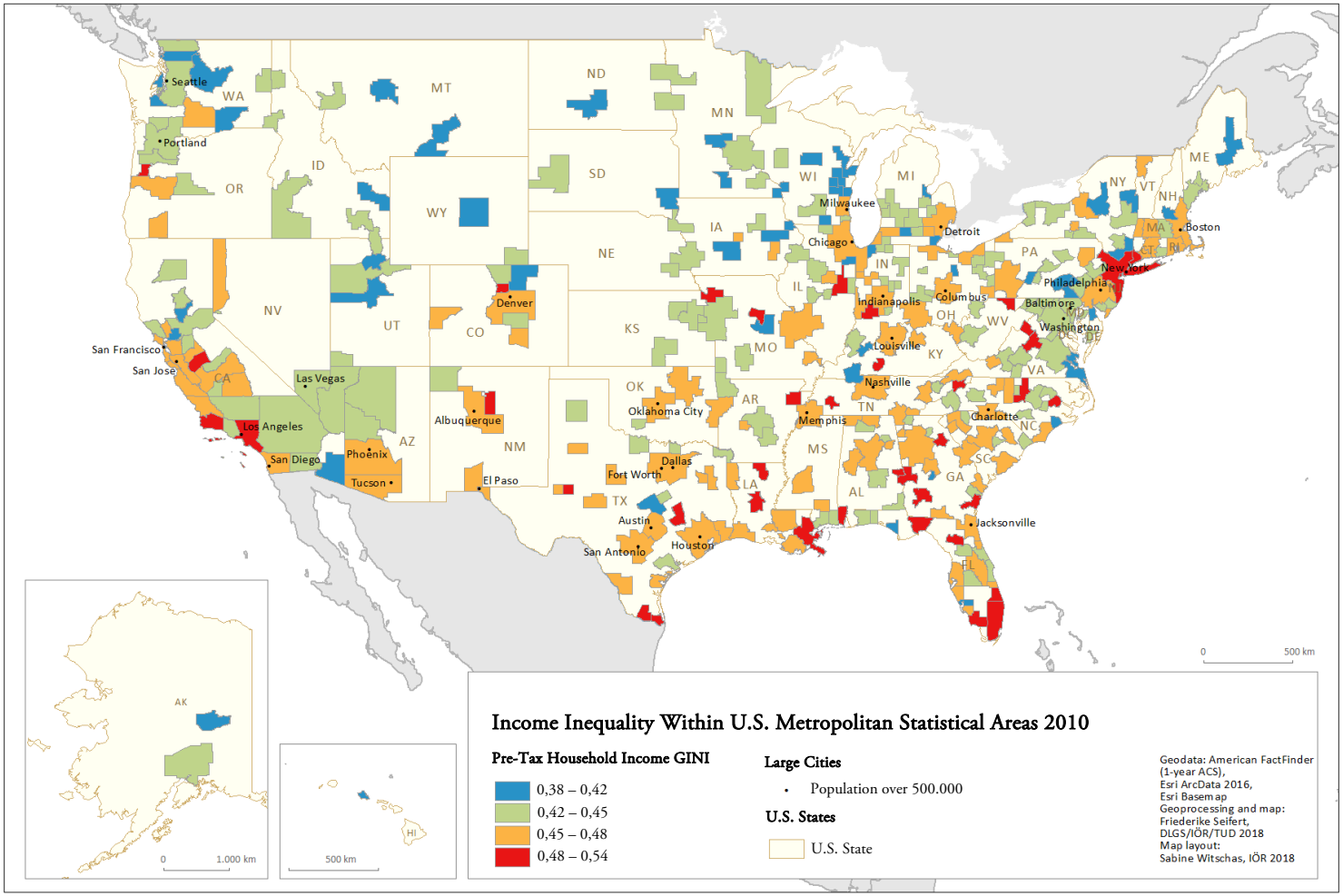
Table 1.
Descriptive Statistics 2006-2016 Data Set

	Obs.	Mean	St. Dev.			Min	Max
			Overall	Between	Within		
Gini	4069	0,450	0,027	0,023	0,015	0,355	0,561
Income per capita	4069	24738	4423	4223	1175	12572	51661
Mean household income	4069	63444	11527	11193	2938	42026	139718

The statistics are for all observations of all MSAs over the entire 2006-2016 period pooled together. The within standard deviation is within MSAs.
Source: FactFinder as well as own calculations.

[2] An MSA is a geographic entity delineated by the Office of Management and Budget for use by US statistical agencies. MSAs consist of the county or counties associated with at least one urbanized area of at least 50,000 inhabitants plus adjacent counties having a high degree of social and economic integration with the core as measured through commuting ties (US Census Bureau, n.d.-b).

FIGURE 1.
MSAs' Inequality Levels 2010



ISSN: 1695-7253 e-ISSN: 2340-2717

TABLE 2.
Descriptive Statistics 1980-2000 Data Set

	Obs.	Mean	St. Dev.			Min	Max
			Overall	Between	Within		
Gini	735	0,416	0,033	0,024	0,024	0,333	0,532
Income per capita	735	23906	4379	3742	2425	11664	42928

The statistics are for all observations of all MSAs over the entire 1980-2000 period pooled together. The within standard deviation is within MSAs.
Source: NHGIS and IPUMS as well as own calculations.

For the 1980-2000 data set, the data stems from the US Census via NHGIS and IPUMS (Manson et al., 2017; Ruggles et al., 2018). NHGIS offers aggregated data at the MSA level for all main variables except the GINI. The latter is calculated from IPUMS, which offers household-level data. There are drawbacks to using IPUMS data to calculate the GINI compared to variables provided by NHGIS or FactFinder directly. First, MSA populations are incompletely identified in the IPUMS data sets (IPUMS-USA, n.d.-a). Second, data confidentiality issues in smaller MSAs reduce the sample size. Third, household income is bottom-coded and the reported incomes are rounded in all years (IPUMS-USA, n.d.-b).[3] The correlation between the 2010 FactFinder and IPUMS-calculated GINIs is nonetheless over 0.9 and statistically significant at the 1% level. Table 2 presents descriptive statistics. The resulting unbalanced panel data set for 1980-2000 consists of 260 MSAs and 3 years.

This paper estimates the income-inequality relationship in cross-sections and panel frameworks using MSA and time FEs. The latter approach controls for time- and MSA-invariant variables. It also allows studying dynamics of change within short time series (Rodríguez-Pose & Tselios, 2009). However, FEs might lead to less variation than in cross-sectional studies as only within variation is considered (Royuela et al., 2019). This effect might be especially relevant for the 2006-2016 panel analysis as inequality is believed to change only slowly over time (Glaeser et al., 2009; Royuela et al., 2019).

This paper regresses inequality on mean income in the same year. The empirical model is as follows:

$$g_{it} = \alpha + \beta y_{it} + \gamma \boldsymbol{X}_{it} + \mu_i + \tau_t + \varepsilon_{it} \quad (1)$$

where g_{it} is a measure of inequality for MSA *i* at time *t*, y_{it} is an income measure (in logs), $\boldsymbol{X}_{it}$ is a vector of control variables, μ_i and τ_t are respectively MSA and time FEs, and ε_{it} is the error term. Standard errors are clustered at the MSA level.[4] The cross-sections exclude the MSA and time FEs and are only estimated for a given *t*.

Controls for population, education (population share with a bachelor's degree or higher, respectively with a high school diploma or higher, in percent), and sector employment shares (share of persons 16 years and over employed in agriculture, respectively in the manufacturing sector, in percent) are included to avoid confounding factors. They have been shown to influence within-city inequality, as previously discussed. Furthermore, some regressions include quadratic income terms to test for a nonlinear income-inequality relationship. Still, unobservable factors might influence the income-inequality relationship, potentially leading to omitted variable bias. The panel regressions account for time-invariant MSA characteristics, reducing this issue compared to the cross-sections.

[3] A negative income is possible because both the Census and the ACSs include self-employment income from own businesses, that is, net income after business expenses. Furthermore, they include income from an estate or trust, interest, and dividends, which can be negative as well (IPUMS-USA, n.d.-b).
[4] Neither state FEs nor standard errors clustered at the state level can be included. Several MSAs cross state borders and belong to more than one state (not necessarily in equal parts). Furthermore, the number of MSAs per state is limited with several states only having one or two MSAs.

Reverse causality between income and inequality constitutes an issue in these regressions, leading to endogeneity. Income influences inequality, but inequality, in turn, affects income and income growth. Convincing instruments for income have not yet been proposed in this context. Therefore, the obtained coefficients have to be interpreted as associations rather than causal effects of income on inequality. However, this income-inequality association is interesting in its own right and relevant for policy debate.

4. Cross-Section Results

This section presents cross-section results using both data sets. These results constitute a starting point to assess the income-inequality relationship across time. To gain a first impression, Appendix Figures I—III present scatterplots of the GINI and logarithmized income. They all show a slight upward trend, that is, a positive income-inequality relationship: inequality increases with income. Similarly, the income per capita coefficient is positive and statistically significant at the 1% level in pooled cross-section regressions without any control variables included (Appendix Table I). However, these simple regressions do not account for year specificities or the influence of several relevant control variables.

Table 3 presents cross-section results by year, including controls. The first three columns report regression results for 2016, 2010, and 2006. These regressions use the 2013 MSA delineations. The data stems from the ACSs via FactFinder. The last three columns report regression results for 2000, 1990, and 1980. These regressions use the 1990 MSA delineations. The data stems from the Census via NHGIS and IPUMS.

Table 3.
Cross-Section Results Regressing Inequality on Income

	(1)	(2)	(3)	(4)	(5)	(6)
	2016	**2010**	**2006**	**2000**	**1990**	**1980**
	Gini	**Gini**	**Gini**	**Gini**	**Gini**	**Gini**
ln (income per capita)	-0,003 (0,016)	0,001 (0,014)	0,016 (0,015)	-0,032 (0,022)	-0,048*** (0,018)	-0,038*** (0,014)
ln (population)	0,000 (0,002)	-0,001 (0,001)	-0,001 (0,001)	0,001 (0,002)	0,002 (0,002)	0,001 (0,001)
Share bachelor's degree	0,001*** (0,000)	0,002*** (0,000)	0,002*** (0,000)	0,077** (0,036)	0,101*** (0,038)	0,050** (0,023)
Share high school diploma	-0,003*** (0,000)	-0,003*** (0,000)	-0,003*** (0,000)	-0,223*** (0,063)	-0,236*** (0,049)	-0,170*** (0,028)
Share agriculture	-0,002*** (0,001)	-0,002** (0,001)	-0,002*** (0,001)	-0,249*** (0,088)	-0,140** (0,069)	-0,097*** (0,035)
Share manufacturing sector	-0,001*** (0,000)	-0,001*** (0,000)	-0,001*** (0,000)	-0,126*** (0,025)	-0,070*** (0,023)	-0,129*** (0,013)
Constant	0,711*** (0,124)	0,676*** (0,125)	0,545*** (0,126)	0,969*** (0,178)	1,082*** (0,140)	0,916*** (0,114)
MSAs	382	366	359	251	245	239
R^2	0,232	0,311	0,305	0,310	0,338	0,496

The first three columns report results for 2016, 2010, and 2006 respectively. They use 2013 MSA delineations and ACS data from FactFinder. The last three columns report results for 2000, 1990, and 1980 respectively. They use 1990 MSA delineations and Census data from NHGIS and IPUMS. Robust standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.
Source: FactFinder, NHGIS and IPUMS as well as own calculations.

For the years 2000, 2006, 2010, and 2016, the income coefficient is not statistically significant even at the 10% level. Income per capita levels appear not to influence inequality levels in these years: either positively or negatively. The income coefficient is statistically significant at the 1% level and negative in 1980 and 1990. Higher income per capita levels appear to be associated with reduced inequality levels in these years. A 1% increase in income per capita involves ceteris paribus a decrease in the GINI by 0.0004 (1980) respectively 0.0005 points (1990) for a given MSA. This decrement is equivalent to a decrease by about 0.1% at the mean of the GINI. These negative coefficients correspond to the previous findings in the literature for MSAs. Section 7 discusses in detail the reasons for the divergent results across years.

The control variables' coefficients are typically of the expected signs. Only population surprises with a statistically insignificant coefficient. Thus, the MSA size does not seem to influence the inequality level in the considered context, contrary to the existing literature (Glaeser et al., 2009; Baum-Snow & Pavan, 2012; Castells-Quintana et al., 2020). The divergent results in the present study might stem from including other control variables than in previous studies. Still, the coefficient remains in four out of six cases of positive sign as expected.

The control variables do not drive the results as similar results are obtained when excluding them from the regression (Appendix Table II). The statistically significant negative income-inequality relationship persists for 1980 and 1990. The same now applies to 2000. For 2006, 2010, and 2016, the income per capita coefficient remains not statistically significant, as previously. The switch in significance is hence also observed without controls included. The positive coefficient of pooling all 1980-2000 observations together without controls vanishes in the by-year regressions and turns negative. It is probably due to a time trend. The switch in significance is also observed in pooled cross-sections with controls, albeit the other way round. The income coefficient is always negative but only statistically significant for the 2006-2016 panel (Appendix Table I).

When adding quadratic income terms, a similar pattern to that of the main regressions appears (Appendix Table III). In the cross-sections for 2006, 2010, and 2016, neither the linear nor the quadratic income coefficient is statistically significant at the 10% level. For 1980, 1990, and 2000, both are statistically significant at the 1% level. The linear one is positive and the quadratic one negative, indicating an inverted-U-shaped relationship as expected. Most MSAs are situated in the downward sloping part of the curve according to their observed incomes per capita. This pattern indicates that higher-income MSAs have over-proportionally low inequality levels in these years, while a negative income-inequality relationship holds for most MSAs.

5. Panel Results

This section presents panel results using both data sets. They permit evaluating the impact of changes in income per capita on inequality and provide a comparison point to the cross-section results. In addition, they reduce the issue of unobserved heterogeneity in time-invariant MSA characteristics compared to cross-sections.

Table 4 presents the results. The first two columns show the annual 2006-2016 panel results. Column one uses income *per capita* while column two employs *mean household* income. The third column shows the decennial 1980-2000 panel results employing income *per capita*.[5]

For the 2006-2016 panel, the income coefficient is statistically significant at the 1% level and positive in both regressions. Income increases appear to be associated with inequality increases. A 1% increase in per capita (mean household) income involves, ceteris paribus, an increase in the GINI by 0.0015 (0.0014) points for a given MSA. This increment is equivalent to an increase by about 0.3% at the mean

[5] Mean household income is not available for 1980 and 1990. Its cross-section results for the remaining years are very similar to the income per capita ones (available upon request).

of the GINI. These results correspond to the ones obtained for European regions in annual panels over the 1990s and 2000s (Rodríguez-Pose & Tselios, 2009; Castells-Quintana et al., 2015).

For the 1980-2000 panel, the income coefficient is statistically significantly negative. An increase in income seems to have decreased inequality. The absolute size of the income coefficient is smaller than previously. A 1% increase in income per capita involves, ceteris paribus, a decrease in the GINI by 0.0007 points for a given MSA. This decrement is equivalent to a decrease by about 0.2% at the mean of the GINI. However, the within-R^2 increases considerably from 0.29 to 0.85. Section 7 discusses the reasons for these divergent results.

TABLE 4.
Panel Results Regressing Inequality on Income

	2006-2016		1980-2000
	(1)	(2)	(3)
	Gini	Gini	Gini
ln (income per capita)	0,149*** (0,009)		-0,072*** (0,017)
ln (mean household income)		0,135*** (0,010)	
ln (population)	-0,020*** (0,004)	-0,023*** (0,005)	-0,013** (0,006)
Share bachelor's degree	-0,001** (0,000)	-0,000* (0,000)	0,150** (0,059)
Share high school diploma	-0,001*** (0,000)	-0,001*** (0,000)	-0,043 (0,035)
Share agriculture	-0,001 (0,000)	-0,000 (0,000)	-0,000 (0,072)
Share manufacturing sector	-0,001*** (0,000)	-0,000** (0,000)	-0,084*** (0,027)
Constant	-0,674*** (0,107)	-0,628*** (0,126)	1,302*** (0,161)
MSA & Time FE	yes	yes	yes
N	4069	4069	735
MSAs	399	399	260
T	11	11	3
Within-R^2	0,288	0,267	0,849

The first two columns report the 2006-2016 annual panel results, while the third column reports the 1980-2000 decennial panel results. Standard errors clustered at the MSA level in parentheses; * p<0.1, ** p<0.05, *** p<0.01.
Source: FactFinder resp. NHGIS and IPUMS as well as own calculations.

From the control variables, population sticks out again. It now exhibits a statistically significant negative coefficient. Thus, increases in MSA size seem to decrease inequality for a given inequality and city size level, whereas the population level per se does not affect an MSA's inequality level. The observed difference in results probably stems from these differing interpretations and foci of cross-section and MSA FEs panel regressions. The existing literature estimates cross-sections or panels without city FEs, rendering the obtained results not directly comparable with those of the MSA FE.

The obtained results are robust to excluding all control variables from the regression while keeping the MSA and sometimes time FEs (Appendix Table IV). The statistically significant positive income-inequality relationship in the 2006-2016 panel persists. For the 1980-2000 panel, the income coefficient remains negative and significant when both MSA and time FEs are included. When only MSA FEs are included, the coefficient turns positive while remaining statistically significant, showing again the importance of controlling for time specificities in this long-term panel.

When adding quadratic income terms, the switch in signs between the two panels is again observed (Appendix Table V). Only the linear income term is statistically significant in the 2006-2016 panel with income per capita. It is positive as previously. With mean household income, both income terms are statistically significant. The linear one is positive, and the quadratic one is negative, indicating an inverse-U-shaped income-inequality relationship. However, all MSAs are located within the upward-sloping part of the curve according to their observed incomes, hence exhibiting a positive income-inequality relationship. For the 1980-2000 panel, both income terms are statistically significant but now of the opposite sign. The linear one is negative while the quadratic one is positive. Thus, the income-inequality relationship is U-shaped. Most MSAs are located within the downward-sloping part of the curve, hence exhibiting a negative income-inequality relationship. Thus, the same pattern as in the linear-only regressions reappears.

6. Employing Alternative Inequality Measures

The obtained opposing results for the two data sets might stem from a peculiarity of the GINI. Therefore, the previous regressions were repeated with several other inequality measures to test the results' robustness. The robustness check sections only present results for the panel regressions as they exhibit most clearly the pattern of switching signs. Furthermore, they can be considered the more reliable results as they abstract from MSA-specific unobservable characteristics, which might bias the cross-section results.[6]

The calculated alternative inequality measures for within-MSA inequality are as follows:

- the GE(0) (Generalized Entropy index with a=0, that is, the mean log deviation),[7]
- the 90/10, 90/50, and 50/10 percentile ratios, and
- the s1, the income share of the top 1% incomes in an MSA.

The GE(0) is an overall inequality measure as the GINI, providing a direct comparison point. The 90/10 percentile ratio is also an overall measure, but it excludes the extreme values at the top and bottom of the income distribution. The 90/50 percentile ratio measures the inequality within top incomes, while the 50/10 percentile ratio measures the inequality within bottom incomes. The s1 indicates the evolution of the very top incomes compared to the rest.

The alternative inequality measures are calculated for both data sets from IPUMS as it offers household-level data. This procedure reduces the number of observations in the 2006-2016 data set to 2,856 (from 4,069 before) and in the 1980-2000 data set to 700 (735 before). The alternative inequality measures replace the GINI as the dependent variable in the regressions. Table 5 presents the 2006-2016 panel results, and Table 6 the 1980-2000 panel results.

For the 2006-2016 panel, GE(0) shows a very similar result to the GINI one: a statistically significant and positive income coefficient. The income coefficient is also statistically significantly positive for s1, while it is not statistically significant in the regressions with the percentile ratios. For the 1980-2000

[6] Robustness checks have also been run for the cross-sections with similar results, indicating that their results are overall robust as well (Appendix Tables VI—X).

[7] Regressions have also been run for the GE(2) (Generalized Entropy index with a=2, that is, half the squared coefficient of variation). The obtained results are very similar to the GE(0) ones. The results have been omitted due to space considerations but are available upon request.

panel, all income coefficients are statistically significant and negative as with the GINI except for the 50/10 percentile ratio and s1. In the latter cases, the coefficient is not statistically significant.

Overall, the regressions with alternative inequality measures confirm the results obtained with the GINI. The oppositional signs of the two panels' income coefficients appear again for the GE(0). The other measures exhibit mixed results. This corresponds to expectations as they only consider parts of the income distribution.

TABLE 5.
Alternative Inequality Measures in the 2006-2016 Panel

	(1)	(2)	(3)	(4)	(5)	(6)
	Gini	**Ge0**	**p90p10**	**p90p50**	**p50p10**	**s1**
ln (income per capita)	0,126*** (0,011)	0,037** (0,016)	-1,097 (0,961)	0,027 (0,100)	-0,480 (0,317)	0,052*** (0,007)
Controls	yes	yes	yes	yes	yes	yes
MSA & Time FEs	yes	yes	yes	yes	yes	yes
Constant	yes	yes	yes	yes	yes	yes
N	2856	2856	2856	2856	2856	2856
MSAs	293	293	293	293	293	293
Within-R^2	0,289	0,248	0,062	0,141	0,028	0,090

Standard errors clustered at the MSA level in parentheses; * p<0.1, ** p<0.05, *** p<0.01.
Source: FactFinder and IPUMS as well as own calculations.

TABLE 6.
Alternative Inequality Measures in the 1980-2000 Panel

	(1)	(2)	(3)	(4)	(5)	(6)
	Gini	**ge0**	**p90p10**	**p90p50**	**p50p10**	**s1**
ln (income per capita)	-0,073*** (0,018)	-0,102*** (0,023)	-3,811*** (0,927)	-0,744*** (0,113)	-0,277 (0,284)	0,012 (0,014)
Controls	yes	yes	Yes	yes	yes	yes
MSA & Time FEs	yes	yes	Yes	yes	yes	yes
Constant	yes	yes	Yes	yes	yes	yes
N	700	700	700	700	700	700
MSAs	254	254	254	254	254	254
within-R^2	0,857	0,852	0,349	0,770	0,140	0,756

Standard errors clustered at the MSA level in parentheses; * p<0.1, ** p<0.05, *** p<0.01.
Source: FactFinder and IPUMS as well as own calculations.

The use of these alternative inequality measures also allows distinguishing between two hypotheses, which have been discussed for the rising inequality in the US: a rise in the top income share and polarization (Piketty & Saez, 2003; Autor et al., 2006; Essletzbichler, 2015). Income per capita has, on average, increased over the study period. Thus, both channels would result in a positive income coefficient for s1 and the 90/50 percentile ratio. Polarization would additionally lead to a negative coefficient for the 50/10 percentile ratio, while the 90/10 ratio should remain relatively unchanged. Notably, the 2006-2016 panel

should exhibit this pattern as it captures the time of technological change substituting middle-skill routine tasks, leading to polarization.

The obtained results hint towards both rises in top incomes and polarization but cannot substantiate these hypotheses unambiguously. The income coefficient for s1 is positive and significant in the new panel compared to being insignificant, albeit already positive, in the old panel. This indicates that the income per capita increases disproportionally benefited the very top incomes. Concurrently, the 90/50 percentile ratio turns insignificantly positive from being significantly negative before. Thus, increasing top incomes played a role in the increasing inequality and switching signs of the income-inequality relationship across the panels. In addition, the 90/10 exhibits an insignificant coefficient in the newer panel, while being significantly negative before, consistent with polarization. However, the coefficient of the 50/10 percentile ratio is not significant but negative in both panels, which questions an income redistribution from the middle to bottom incomes, as suggested by the polarization hypothesis.

7. Reasons for the Change in the Income-Inequality Relationship

There are four possible reasons why the income-inequality relationship changes its sign across panels: differences in the database, changes in the MSA delineations, the different time gaps in the panels, and qualitative changes in the relationship.

First, changes in the underlying data and its aggregation between FactFinder and IPUMS might lead to differing results. The 1980-2000 panel is based on Census data, while the 2006-2016 one uses the ACS. However, both data products are produced by the US Census Bureau according to similar standards. Furthermore, the 2006-2016 results persist when using IPUMS-calculated inequality measures, as shown in the alternative inequality measures regressions. Thus, the differences in the databases cannot account for the changing sign of the income-inequality relationship.

Second, MSA delineation changes result in different MSAs being considered across the two data sets. These changes lead to a clear difference in the number of MSAs available: 260 in the 1980-2000 data set versus 399 in the 2006-2016 one. The increase in sample size due to the number of MSAs alone is considerable. However, 260 MSAs are a large enough number of observations for regression analyses. Furthermore, the panel and cross-section results remain unchanged when restricting the 2006-2016 sample to only those MSAs already existing in 2000 (Appendix Table XI). Besides, one can calculate both the GINI and mean household income from IPUMS for 2000 and 2010 for both MSA delineations. If one then regresses the GINI on the income, the obtained results are qualitatively the same regarding significance levels and signs (Appendix Table XII). Thus, delineation and sample size changes might play a role in the diverging results, but they appear unlikely to be the sole cause of the opposing results.

Third, the time gaps and time dimensions of the panels differ. The 2006-2016 panel is an annual one with observations for 11 different years. The 1980-2000 panel is a decennial one with observations for only three years. Both might result in statistical issues. There might not be enough within-variation in the former for proper estimation, while the number of observations per MSA might be too small in the latter. The 10-year gap between observations in the latter results in a more medium-run perspective than the short-run one of the annual panel. Transmission channels differ in their manifestation rapidity, as discussed in section 2. Purely economic factors typically realize faster than sociopolitical ones (Halter et al., 2014). The former include trade and labor market phenomena, which result in a positive income-inequality relationship. The latter comprise segregation, crime, and sociopolitical contrast and hence exactly those factors leading to a negative income-inequality relationship. Annual panel studies for European regions found likewise positive income-inequality relationships for 1994-2001 (Rodríguez-Pose & Tselios, 2009), respectively 1993-2011 (Castells-Quintana et al., 2015).

The 2006-2016 panel can be transformed into one with 5-year gaps and observations for three years (2006, 2011, and 2016). This approaches the time gap between observations to the one of the 1980-2000 panel and results in the same number of observation years (three). When regressing the GINI on income and the usual controls in this panel, the income coefficient remains statistically significant and positive for both per capita and mean household income. However, its size diminishes by about one-third. A similar

reduction is observed when basing the 5-year panel on the MSAs already existing over 1980-2000 (Appendix Table XIII). Thus, there appears to be something special about the 2006-2016 time period other than the time gap between observations and the number of observed years resulting in the positive income-inequality relationship. However, the 10-year gaps cannot be simulated due to the 2006-2016 panel's limited time dimension.

Forth, the income-inequality relationship might have changed qualitatively over the years, especially between 2000 and 2006, according to the panel results.[8] The cross-section results also reflect this change. The negative income-inequality association stops in 1990 and no longer exists for 2000 and further years. This timing corresponds to the sharp rise in inequality generally observed in the US in the 1980s and beyond (Piketty & Saez, 2003). This increase in inequality is also observed in the MSA-level data employed in the present study. Apparently, it was not only inequality that increased but also its relationship with income changed. The changed sign of the income-inequality relationship also hints at economic growth having become less inclusive over the years.[9]

The influence of factors resulting in a negative income-inequality relationship might have decreased over time while the influence of those leading to a positive relationship increased. Factors resulting in a negative income-inequality relationship include residential segregation, crime, and sociopolitical unrest, as detailed in section 2. Crime rates have indeed declined for several offenses since the 1980s (Asher, 2017), but residential segregation increased during the considered period (Bischoff & Reardon, 2014). Thus, the evidence for a decline in the ''negative'' factors is mixed.

Factors leading to a positive income-inequality relationship include specialization, technological change substituting middle-skill routine tasks, trade, deunionization, and flexible labor market regulations. Trade and specialization have increased since the 1980s due to globalization and technological change substituting middle-skill routine tasks (Autor et al., 2006; Rigby & Breau, 2008; Autor & Dorn, 2013). Unionization rates declined over the last decades (Hu & Hanink, 2018). All these developments would strengthen a positive income-inequality relationship. Combined, they might have led to the observed change in the sign of the income-inequality relationship if the importance of these positive factors were stronger relative to the negative factors, especially residential segregation.

Given the available data, it is impossible to distinguish data-related issues neatly from qualitative changes in the income-inequality relationship. Thus, one cannot exclude the possibility that the differences in the data and the analysis setup are responsible for the observed change in sign of the relationship. This would require a longer annual panel over at least 20 years to evaluate results for panels of different lengths based on a single, consistent data set. Consequently, further research is required on this topic.

8. Conclusion

This paper analyzed the income-inequality relationship within MSAs using two data sets: a decennial one over 1980-2000 based on the Census and an annual one over 2006-2016 based on the ACS. These data sets enable study of the income-inequality relationship within MSAs over a more extended period than was previously possible, as well as employing both cross-section and panel regression techniques.

A higher income per capita level is still associated with a lower within-MSA inequality level in the earlier years. However, this association stops being statistically significant in 2000 and remained so until 2016. For the 1980-2000 panel, income per capita increases are accordingly associated with inequality decreases. In the 2006-2016 panel, income per capita increases are associated with inequality increases. The income-inequality relationship changes direction over time.

---

[8] The European panel studies finding a positive income-inequality relationship analyzed the 1990s and 2000s (Rodríguez-Pose & Tselios, 2009; Castells-Quintana et al., 2015).

[9] The economic crisis of 2008 might also have influenced the income-inequality relationship. However, the change is already visible in the 2000 cross-section, where the income coefficient is insignificant for the first time. Furthermore, the positive income-inequality association also appears in the 2012-2016 panel, starting after the crisis years.

The main explanations for this change in sign are MSA delineation changes and different time dimensions in the panels, as well as qualitative changes in the income-inequality relationship. The latter are probably due to polarization resulting from technological change substituting middle-skill routine tasks in line with Autor & Dorn (2013). However, these explanations cannot be completely distinguished with the data sets at hand.

Therefore, further research is required to solve this puzzle. On the one hand, studies using a more extended annual panel are needed to evaluate the income-inequality relationship in panels with different time dimensions and time gaps. On the other hand, more research on the transmission channels of the income-inequality relationship at the MSA levels might enlighten the influence of specific factors on this relationship in different periods.

References

Asher, J. (2017). The US murder rate is up but still far below its 1980 peak. FiveThirtyEight. https://fivethirtyeight.com/features/the-u-s-murder-rate-is-up-but-still-far-below-its-1980-peak/

Atems, B. (2013). A note on the differential regional effects of income inequality: Empirical evidence using US county-level data. *Journal of Regional Science, 53*(4), 656-671. https://doi.org/10.1111/jors.12053

Autor, D. H., & Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review, 103*(5), 1553-1597. https://doi.org/10.1257/aer.103.5.1553

Autor, D. H., Katz, L. F., & Kearney, M. S. (2006). The polarization of the US labor market. *American Economic Review, 96*(2), 189-194. https://doi.org/10.1257/000282806777212620

Baum-Snow, N., & Pavan, R. (2012). Understanding the city size wage gap. *The Review of Economic Studies, 79*(1), 88-127. https://doi.org/10.1093/restud/rdr022

Bhatta, S. D. (2001). Are inequality and poverty harmful for economic growth: Evidence from the metropolitan areas of the United States. *Journal of Urban Affairs, 23*(3-4), 335-359. https://doi.org/10.1111/0735-2166.00093

Bischoff, K., & Reardon, S. F. (2014). Residential segregation by income, 1970-2009. In John Logan (Ed.), Diversity and disparities America enters a new century. Russell Sage Foundation.

Bolton, K., & Breau, S. (2012). Growing unequal? Changes in the distribution of earnings across Canadian cities. *Urban Studies, 49*(6), 1377-1396. https://doi.org/10.1177/0042098011410335

Castells-Quintana, D., Ramos, R., & Royuela, V. (2015). Income inequality in European regions: Recent trends and determinants. *Review of Regional Research, 35*(2), 123-146. https://doi.org/10.1007/s10037-015-0098-4

Castells-Quintana, D., Royuela, V., & Veneri, P. (2020). Inequality and city size: An analysis for OECD functional urban areas. *Papers in Regional Science, 99*(4), 1045-1064. https://doi.org/10.1111/pirs.12520

Conceição, P., & Galbraith, J. K. (2001). Toward a new Kuznets hypothesis: Theory and evidence on growth and inequality. In J. K. Galbraith & M. Berner (Eds.), Inequality and industrial change: A global view. Cambridge University Press. https://doi.org/10.1017/CBO9781139175210.008

Essletzbichler, J. (2015). The top 1% in US metropolitan areas. *Applied Geography, 61*, 35-46. https://doi.org/10.1016/j.apgeog.2015.01.019

Florida, R., & Mellander, C. (2015). Segregated city: The geography of economic segregation in America's metros. Martin Prosperity Institute.

Florida, R., & Mellander, C. (2016). The geography of inequality: Difference and determinants of wage and income inequality across US metros. *Regional Studies, 50*(1), 79-92. https://doi.org/10.1080/00343404.2014.884275

Forbes, K. J. (2000). A reassessment of the relationship between inequality and growth. *American Economic Review, 90*(4), 869-887. https://doi.org/10.1257/aer.90.4.869

Glaeser, E. L., Resseger, M., & Tobio, K. (2009). Inequality in cities. *Journal of Regional Science, 49*(4), 617-646. https://doi.org/10.1111/j.1467-9787.2009.00627.x

Halter, D., Oechslin, M., & Zweimüller, J. (2014). Inequality and growth: The neglected time dimension. *Journal of Economic Growth, 19*(1), 81-104. https://doi.org/10.1007/s10887-013-9099-8

Hu, Q., & Hanink, D. M. (2018). Declining union contract coverage and increasing income inequality in US metropolitan areas. *The Professional Geographer, 70*(3), 453-462. https://doi.org/10.1080/00330124.2017.1416298

IPUMS-USA. (n.d.-a). IPUMS USA: Incompletely identified metropolitan areas. Retrieved December 22, 2021, from https://usa.ipums.org/usa/volii/incompmetareas.shtml

IPUMS-USA. (n.d.-b). IPUMS-USA: INCTOT. Retrieved December 22, 2021, from https://usa.ipums.org/usa-action/variables/INCTOT

Kuznets, S. (1955). Economic growth and income inequality. *The American Economic Review, 45*(1), 1-28.

Li, H., Campbell, H., & Fernandez, S. (2013). Residential segregation, spatial mismatch and economic growth across US metropolitan areas. *Urban Studies, 50*(13), 2642-2660. https://doi.org/10.1177/0042098013477697

Madden, J. F. (2000). Changes in income inequality within US metropolitan areas. W. E. Upjohn Institute for Employment Research.

Manson, S., Schroeder, J., Van Riper, D., & Ruggles, S. (2017). IPUMS National Historical Geographic Information System: Version 12.0 [database]. University of Minnesota. https://doi.org/10.18128/D050.V14.0

Partridge, M. D. (2005). Does income distribution affect US state economic growth? *Journal of Regional Science, 45*(2), 363-394. https://doi.org/10.1111/j.0022-4146.2005.00375.x

Partridge, M. D., & Weinstein, A. L. (2013). Rising inequality in an era of austerity: The case of the US. *European Planning Studies, 21*(3), 388-410. https://doi.org/10.1080/09654313.2012.716247

Perugini, C., & Martino, G. (2008). Income inequality within European regions: Determinants and effects on growth. *Review of Income and Wealth, 54*(3), 373-406. https://doi.org/10.1111/j.1475-4991.2008.00280.x

Peters, D. J. (2013). American income inequality across economic and geographic space, 1970-2010. *Social Science Research, 42*(6), 1490-1504. https://doi.org/10.1016/j.ssresearch.2013.06.009

Piketty, T., & Saez, E. (2003). Income Inequality in the United States, 1913-1998. *The Quarterly Journal of Economics, 118*(1), 1-41. https://doi.org/10.1162/00335530360535135

Rigby, D., & Breau, S. (2008). Impacts of trade on wage inequality in Los Angeles: Analysis using matched employer-employee data. *Annals of the Association of American Geographers, 98*(4), 920-940. https://doi.org/10.1080/00045600802219786

Rodríguez-Pose, A., & Tselios, V. (2009). Education and income inequality in the regions of the European Union. *Journal of Regional Science, 49*(3), 411-437. https://doi.org/10.1111/j.1467-9787.2008.00602.x

Royuela, V., Veneri, P., & Ramos, R. (2019). The short-run relationship between inequality and growth: Evidence from OECD regions during the Great Recession. *Regional Studies, 53*(4), 574-586. https://doi.org/10.1080/00343404.2018.1476752

Ruggles, S., Flood, S., Goeken, R., Grover, J., Meyer, E., Pacas, J., & Sobek, M. (2018). IPUMS USA: Version 8.0 [dataset]. IPUMS. https://doi.org/10.18128/D010.V10.0

US Census Bureau. (n.d.-a). American FactFinder. Retrieved December 22, 2021, from http://fact-finder2.census.gov

US Census Bureau. (n.d.-b). Glossary. Retrieved December 22, 2021, from https://www.census.gov/programs-surveys/metro-micro/about/glossary.html

US Census Bureau. (2009). A compass for understanding and using American Community Survey data: What researchers need to know. US Government Printing Office. https://www.census.gov/library/publications/2009/acs/researchers.html

ORCID

Friederike Seifert https://orcid.org/0000-0003-3117-8410

Appendix

Figures

Figure I.
Scatterplot of the GINI against ln (income per capita) 2006-2016

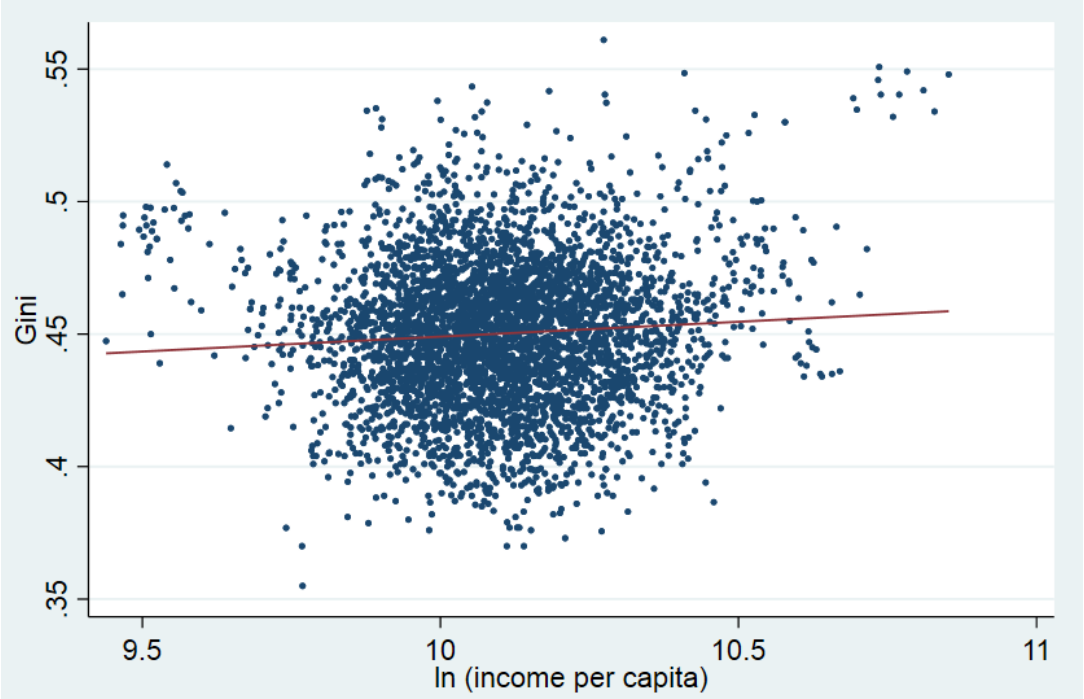


FIGURE II.
Scatterplot of the GINI against ln (mean household income) 2006-2016

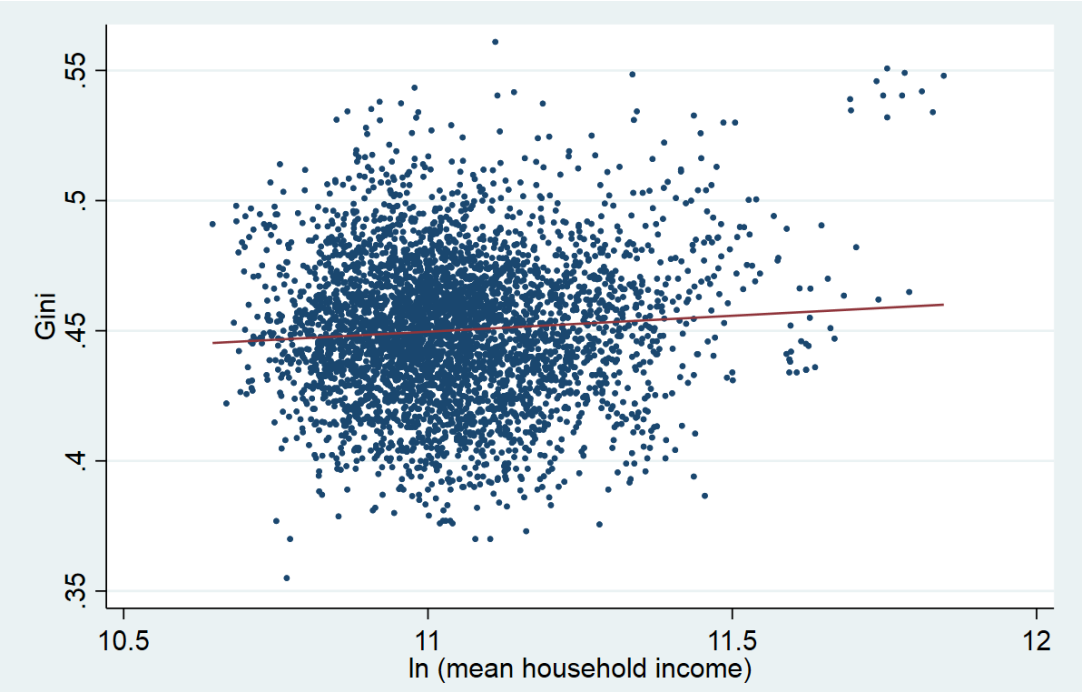
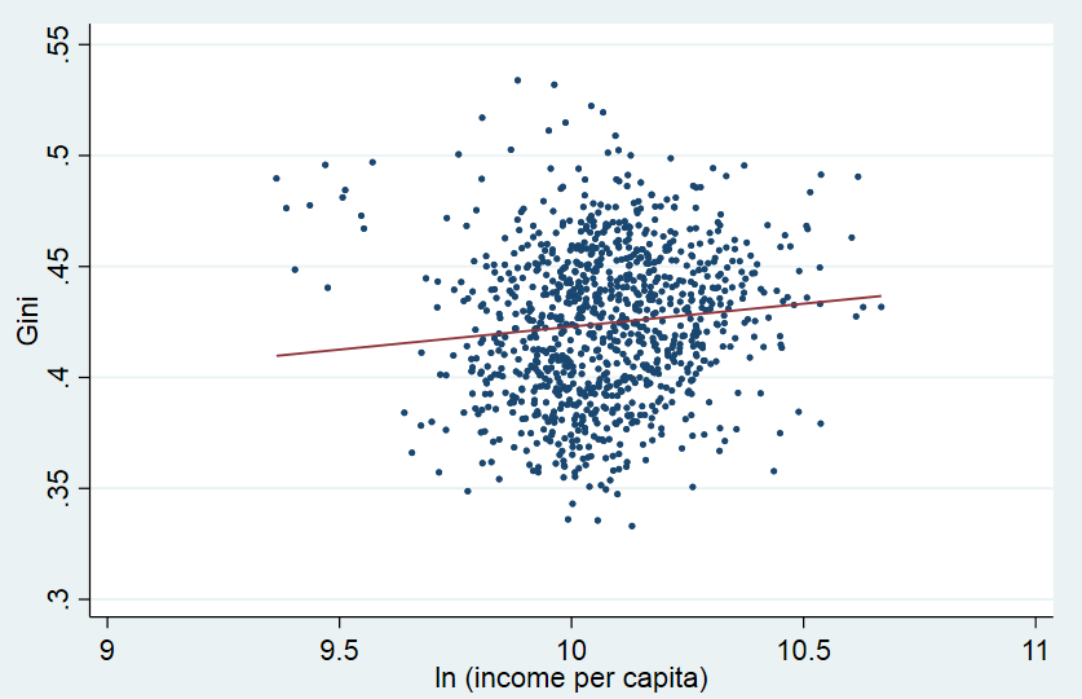


FIGURE III.
Scatterplot of the GINI against ln (income per capita) 1980-2000



Tables

Table VII.
Pooled Cross-Sections Without and With Control Variables Included

	2006-2016			1980-2000	
	(1)	(2)	(3)	(4)	(5)
	Gini	Gini	Gini	Gini	Gini
ln (income per capita)	0,011*** (0,003)	0,011*** (0,003)	-0,013*** (0,004)	0,022*** (0,008)	-0,020 (0,013)
Controls	no	no	yes	no	yes
MSA & Time FEs	no	no	no	no	no
Constant	yes	yes	yes	yes	yes
N	4070	4069	4069	735	735
MSAs	399	399	399	260	260
R^2	0,005	0,005	0,244	0,014	0,270

The first three columns report results for cross-sections pooling all observations over 2006-2016 together, that is, without including MSA and time FEs. Column 1 presents the results for an unrestricted sample without control variables included. Column 2 is restricted to those observations with data for the control variables without including them into the regression. Column 3 then includes control variables. These regressions use 2013 MSA delineations and ACS data from FactFinder. The last two columns report results for cross-sections pooling all observations over 1980-2000 together, without including MSA and time FEs. Column 4 presents the results without control variables included and column 5 with them included. (The sample size is, in this case, unaffected by the inclusion of control variables.) These regressions use 1990 MSA delineations and Census data from NHGIS and IPUMS. Robust standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.
Source: FactFinder, NHGIS and IPUMS as well as own calculations.

Table VIII.
Cross-Section Results Without Control Variables Included

	(1)	(2)	(3)	(4)	(5)	(6)
	2016	2010	2006	2000	1990	1980
	Gini	Gini	Gini	Gini	Gini	Gini
ln (income per capita)	0,011 (0,009)	0,008 (0,010)	0,011 (0,011)	-0,031*** (0,010)	-0,057*** (0,010)	-0,068*** (0,010)
Constant	0,348*** (0,094)	0,364*** (0,101)	0,334*** (0,114)	0,755*** (0,107)	0,990*** (0,096)	1,061*** (0,104)
MSAs	382	366	359	251	245	239
R^2	0,005	0,003	0,004	0,039	0,142	0,173

The first three columns report results for 2016, 2010, and 2006 respectively. They use 2013 MSA delineations and ACS data from FactFinder. The last three columns report results for 2000, 1990, and 1980 respectively. They use 1990 MSA delineations and Census data from NHGIS and IPUMS. No control variables are included in these regressions. Robust standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.
Source: FactFinder, NHGIS and IPUMS as well as own calculations.

TABLE III.
Quadratic Cross-Section Results

	(1)	(2)	(3)	(4)	(5)	(6)
	2016	**2010**	**2006**	**2000**	**1990**	**1980**
	Gini	**Gini**	**Gini**	**Gini**	**Gini**	**Gini**
ln (income per capita)	0,847 (0,777)	1,072 (0,885)	0,569 (0,888)	2,262** (0,900)	2,286*** (0,702)	1,636*** (0,434)
squared ln (income per capita)	-0,042 (0,038)	-0,053 (0,044)	-0,027 (0,044)	-0,112** (0,044)	-0,116*** (0,035)	-0,084*** (0,022)
Controls	yes	yes	yes	yes	yes	yes
Constant	yes	yes	yes	yes	yes	yes
MSAs	382	366	359	251	245	239
R^2	0,237	0,317	0,308	0,337	0,376	0,511

The first three columns report results for 2016, 2010, and 2006 respectively. They use 2013 MSA delineations and ACS data from FactFinder. The last three columns report results for 2000, 1990, and 1980 respectively. They use 1990 MSA delineations and Census data from NHGIS and IPUMS. The usual linear control variables are included in these regressions. Robust standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.
Source: FactFinder, NHGIS and IPUMS as well as own calculations.

TABLE IV.
Panel Results Without Control Variables Included

	2006-2016				1980-2000	
	(1)	(2)	(3)	(4)	(5)	(6)
	Gini	**Gini**	**Gini**	**Gini**	**Gini**	**Gini**
ln (income per capita)	0,065*** (0,008)	0,135*** (0,009)			0,184*** (0,006)	-0,056*** (0,012)
ln (mean household income)			0,069*** (0,008)	0,124*** (0,009)		
Controls	no	no	no	no	no	no
MSA FEs	yes	yes	yes	yes	yes	yes
Time FEs	no	yes	no	yes	no	yes
Constant	yes	yes	yes	yes	yes	yes
N	4069	4069	4069	4069	735	735
MSAs	399	399	399	399	260	260
T	11	11	11	11	3	3
Within-R^2	0,044	0,268	0,047	0,250	0,579	0,836

The first four columns report the 2006-2016 annual panel results, while the fourth and fifth column reports the 1980-2000 decennial panel results. All regressions do not include control variables but include MSA FEs. The pair columns (2, 4, and 6) additionally include time FEs. The sample is restricted to those observations that have observations for all control variables. The unrestricted (full sample) results are identical. Standard errors clustered at the MSA level in parentheses; * p<0.1, ** p<0.05, *** p<0.01.
Source: FactFinder resp. NHGIS and IPUMS as well as own calculations.

TABLE V.
Quadratic Panel Results

	(1)	(2)	(3)
	2006-2016		1980-2000
	Gini	Gini	Gini
ln (income per capita)	0,716** (0,345)		-1,248*** (0,464)
Quadratic ln (income per capita)	-0,028 (0,017)		0,059** (0,023)
ln (mean household income)		1,700*** (0,477)	
Quadratic ln (mean household income)		-0,071*** (0,022)	
Controls	yes	yes	yes
MSA and Time FEs	yes	yes	yes
Constant	yes	yes	yes
N	4069	4069	735
MSAs	399	399	250
T	11	11	3
within-R^2	0,289	0,272	0,851

The first two columns report the 2006-2016 annual panel results, while the third column reports the 1980-2000 decennial panel results. Standard errors clustered at the MSA level in parentheses; The usual linear control variables are included in these regressions. * p<0.1, ** p<0.05, *** p<0.01.
Source: FactFinder resp. NHGIS and IPUMS as well as own calculations.

TABLE VI.
Cross-Section Results regressing GE(0) on Income

	(1)	(2)	(3)	(4)	(5)	(6)
	2016	2010	2006	2000	1990	1980
	ge0	ge0	ge0	ge0	ge0	ge0
ln (income per capita)	-0,023 (0,031)	-0,035 (0,036)	-0,052 (0,036)	-0,071* (0,038)	-0,116*** (0,032)	-0,057*** (0,018)
Controls	yes	yes	yes	yes	yes	yes
Constant	yes	yes	yes	yes	yes	yes
MSAs	260	261	259	251	225	229
R^2	0,233	0,352	0,270	0,280	0,385	0,518

The first three columns report results for 2016, 2010, and 2006 respectively. They use 2013 MSA delineations and ACS data from FactFinder. The last three columns report results for 2000, 1990, and 1980 respectively. They use 1990 MSA delineations and Census data from NHGIS and IPUMS. The usual controls are included. Robust standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.
Source: FactFinder, NHGIS and IPUMS as well as own calculations.

TABLE IX.
Cross-Section Results regressing the 90/10 Percentile Ratio on Income

	(1)	(2)	(3)	(4)	(5)	(6)
	2016	**2010**	**2006**	**2000**	**1990**	**1980**
	p90p10	**p90p10**	**p90p10**	**p90p10**	**p90p10**	**p90p10**
ln (income per capita)	-7,982 (5,427)	-4,235** (2,091)	-4,711*** (1,787)	-3,089* (1,611)	-4,768*** (1,256)	-1,355 (0,824)
Controls	yes	yes	yes	yes	yes	yes
Constant	yes	yes	yes	yes	yes	yes
MSAs	260	261	259	251	225	229
R^2	0,115	0,323	0,286	0,246	0,351	0,400

The first three columns report results for 2016, 2010, and 2006 respectively. They use 2013 MSA delineations and ACS data from FactFinder. The last three columns report results for 2000, 1990, and 1980 respectively. They use 1990 MSA delineations and Census data from NHGIS and IPUMS. The usual controls are included. Robust standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.
Source: FactFinder, NHGIS and IPUMS as well as own calculations.

TABLE XIII.
Cross-Section Results regressing the 90/50 Percentile Ratio on Income

	(1)	(2)	(3)	(4)	(5)	(6)
	2016	**2010**	**2006**	**2000**	**1990**	**1980**
	p90p50	**p90p50**	**p90p50**	**p90p50**	**p90p50**	**p90p50**
ln (income per capita)	0,075 (0,167)	-0,085 (0,187)	-0,144 (0,183)	0,002 (0,214)	-0,568*** (0,121)	-0,358*** (0,080)
Controls	yes	yes	yes	yes	yes	yes
Constant	yes	yes	yes	yes	yes	yes
MSAs	260	261	259	251	225	229
R^2	0,245	0,322	0,330	0,376	0,537	0,574

The first three columns report results for 2016, 2010, and 2006 respectively. They use 2013 MSA delineations and ACS data from FactFinder. The last three columns report results for 2000, 1990, and 1980 respectively. They use 1990 MSA delineations and Census data from NHGIS and IPUMS. The usual controls are included. Robust standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.
Source: FactFinder, NHGIS and IPUMS as well as own calculations.

TABLE IX.
Cross-Section Results regressing the 50/10 Percentile Ratio on Income

	(1)	(2)	(3)	(4)	(5)	(6)
	2016	**2010**	**2006**	**2000**	**1990**	**1980**
	p50p10	**p50p10**	**p50p10**	**p50p10**	**p50p10**	**p50p10**
ln (income per capita)	-2,690 (1,648)	-1,223** (0,543)	-1,424*** (0,462)	-1,000** (0,406)	-0,981*** (0,353)	0,066 (0,315)
Controls	yes	yes	yes	yes	yes	yes
Constant	yes	yes	yes	yes	yes	yes
MSAs	260	261	259	251	225	229
R^2	0,107	0,241	0,228	0,145	0,216	0,195

The first three columns report results for 2016, 2010, and 2006 respectively. They use 2013 MSA delineations and ACS data from FactFinder. The last three columns report results for 2000, 1990, and 1980 respectively. They use 1990 MSA delineations and Census data from NHGIS and IPUMS. The usual controls are included. Robust standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.
Source: FactFinder, NHGIS and IPUMS as well as own calculations.

TABLE X.
Cross-Section Results regressing the Top 1% Income Share on Income

	(1)	(2)	(3)	(4)	(5)	(6)
	2016	**2010**	**2006**	**2000**	**1990**	**1980**
	s1	**s1**	**s1**	**s1**	**s1**	**s1**
ln (income per capita)	-0,013* (0,007)	0,008 (0,006)	-0,000 (0,007)	-0,026*** (0,004)	-0,013*** (0,005)	0,051*** (0,010)
Controls	yes	yes	yes	yes	yes	yes
Constant	yes	yes	yes	yes	yes	yes
MSAs	260	261	259	251	225	229
R^2	0,144	0,100	0,076	0,442	0,274	0,442

The first three columns report results for 2016, 2010, and 2006 respectively. They use 2013 MSA delineations and ACS data from FactFinder. The last three columns report results for 2000, 1990, and 1980 respectively. They use 1990 MSA delineations and Census data from NHGIS and IPUMS. The regressions include the usual controls. Robust standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.
Source: FactFinder, NHGIS and IPUMS as well as own calculations.

TABLE XI.
2006-2016 Panel Restricted to the 1980-2000 MSAs Results

	(1)	(2)	(3)	(4)	(5)
	Annual panel		**2016**	**2010**	**2006**
	Gini	**Gini**	**Gini**	**Gini**	**Gini**
ln (income per capita)	0,119*** (0,011)		0,000 (0,021)	-0,016 (0,019)	-0,004 (0,015)
ln (mean household income)		0,107*** (0,010)			
Controls	yes	yes	yes	yes	yes

TABLE XII. CONT.
2006-2016 Panel Restricted to the 1980-2000 MSAs Results

	(1)	(2)	(3)	(4)	(5)
	Annual panel		2016	2010	2006
	Gini	Gini	Gini	Gini	Gini
MSA & time FEs	yes	yes	-	-	-
Constant	yes	yes	yes	yes	yes
N	2600	2600	-	-	-
MSAs	240	240	232	238	240
T	11	11	-	-	-
(within-) R^2	0,289	0,274	0,270	0,348	0,325

The table reports regression results for the annual panel 2006-2016 (columns 1 and 2) and the cross-sections for 2016, 2010, and 2006 (columns 3-5). In all these regressions, the sample is reduced to those MSAs, in 2013 MSA delineations, that already existed in 2000, resulting in at most 240 MSAs. This number is slightly smaller than the 251 MSAs available for 2000 because some MSAs grew between 2000 and 2006 in such a way that they fused with other MSAs, reducing their number. After 2006, some further slight MSA changes occurred, reducing their number in later years. The regressions include the usual controls. Standard errors clustered at the MSA level in parentheses; * p<0.1, ** p<0.05, *** p<0.01.
Source: FactFinder as well as own calculations.

TABLE XII.
Cross-Section Results for 2000 and 2010 with Different MSA Delineations

	(1)	(2)	(3)	(4)	(5)	(6)
	2010 old IPUMS	2010 new IPUMS	2010 aggregated	2000 aggregated	2000 old IPUMS	2000 new IPUMS
	Gini	Gini	Gini	Gini	Gini	Gini
ln (mean household income)	0,015 (0,012)	0,015 (0,011)	0,014 (0,010)	-0,054*** (0,009)	-0,000 (0,000)	-0,019 (0,013)
Controls	no	no	no	no	no	no
Constant	yes	yes	yes	yes	yes	yes
MSAs	283	261	366	251	283	258
R^2	0,011	0,010	0,007	0,115	0,000	0,015

The table reports results for regressing the GINI on log mean household income without control variables included. The first three columns report results for 2010. The regression of the first column uses the 1990 MSA delineations together with mean household income calculated from IPUMS micro data. The regression of the second column also calculates from IPUMS but uses the 2013 MSA delineations. The regression of the third column then uses the aggregated FactFinder data for the GINI and the mean household income and the 2013 MSA delineations as in the main regressions. The last three columns report results for 2000. The regression of the fourth column uses the aggregated Census NHGIS data for mean household income and the 1990 MSA delineations as in the main regressions. The regression of the fifth column also uses the 1990 MSA delineations but calculates mean household income from IPUMS. The regression of the third column then uses the 2013 MSA delineations while calculating the mean household income from IPIMS. The GINI is in these regressions always calculated from IPUMS. Robust standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.
Source: FactFinder, NHGIS and IPUMS as well as own calculations.

TABLE XIII.
5-Year-Period Panels 2006-2016 Results

	(1)	(2)	(3)	(4)
	All MSAs		Only 2000 MSAs	
	Gini	Gini	Gini	Gini
ln (income per capita)	0,099*** (0,013)		0,071*** (0,017)	
ln (mean household income)		0,094*** (0,014)		0,060*** (0,017)
Controls	yes	yes	yes	yes
MSA & Time FE	yes	yes	yes	yes
Constant	yes	yes	yes	yes
N	1106	1106	710	710
MSAs	397	397	240	240
T	3	3	3	3
within-R^2	0,330	0,322	0,350	0,338

The table reports regression results for a 5-year-period subpanel of the 2006-2016 one. Thus, it includes observations from 2006, 2011, and 2016 only. The first two columns report results for the full sample, while the last two columns present the results for restricting the sample to only those MSAs, in 2013 MSA delineations, that have already existed in 2000 (as in Appendix Table XI). The regressions include the usual controls. Standard errors clustered at the MSA level in parentheses; * p<0.1, ** p<0.05, *** p<0.01.
Source: FactFinder as well as own calculations.