

Seifert, Friederike

Working Paper

The income-inequality relationship within U.S. metropolitan areas 1980-2016

CEPIE Working Paper, No. 01/21

Provided in Cooperation with:

Technische Universität Dresden, Faculty of Business and Economics

Suggested Citation: Seifert, Friederike (2021) : The income-inequality relationship within U.S. metropolitan areas 1980-2016, CEPIE Working Paper, No. 01/21, Technische Universität Dresden, Center of Public and International Economics (CEPIE), Dresden, <https://nbn-resolving.de/urn:nbn:de:bsz:14-qucosa2-740877>

This Version is available at:

<https://hdl.handle.net/10419/232577>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



CEPIE Working Paper No. 01/21

Center of Public and International Economics

THE INCOME-INEQUALITY RELATIONSHIP
WITHIN U.S. METROPOLITAN AREAS
1980—2016

March 2021

Friederike Seifert

Editors: Faculty of Business and Economics, Technische Universität Dresden.

This paper is published on the [Open Access Repository Qucosa](#).

The complete Working Paper Series can be found at the [CEPIE Homepage](#) | [EconStor](#) | [RePEc](#)

The Income-Inequality Relationship within U.S. Metropolitan Areas 1980—2016

Friederike Seifert*

TU Dresden and Leibniz Institute Dresden

March 2021

Abstract

Economic growth might both increase and decrease income inequality, depending on the circumstances. The nature of this relationship matters at the city level as well. This paper examines the income-inequality relationship within U.S. metropolitan areas using cross-section and panel regression techniques over the 1980—2016 period. It finds that this relationship changes over time. A higher per capita income level was associated with a lower within-MSA inequality level in earlier years, but this association vanished later. For the 1980—2000 panel, per capita income increases are accordingly associated with decreases in inequality. In contrast, an increase in per capita income is associated with an increase in inequality in the 2006—2016 panel. The obtained results hint at polarization resulting from technological change substituting middle-skill routine tasks, but further research is still required to solve this puzzle.

JEL classification: D31, O18, R11

Keywords: Inequality, Income, Metropolitan Areas, United States

*Corresponding author: Faculty of Business and Economics and "Friedrich List" Faculty of Transport and Traffic Sciences, Technische Universität Dresden, 01062 Dresden, Germany and Leibniz Institute of Ecological Urban and Regional Development, Weberplatz 1, 01217 Dresden, Germany. E-mail: friederike.seifert@tu-dresden.de

1 Introduction

The income-inequality relationship has been a question of debate since the seminal work by Kuznets proposing 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 the national one. Some channels from the national level, such as credit market mechanisms and redistribution policies, do not translate directly to the city level (Glaeser, Resseger, & Tobio, 2009; Royuela, Veneri, & Ramos, 2019). The latter is characterized by more in-and-out-migration and less political room for maneuver than countries. 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). Still, comparatively little is known about the income-inequality relationship at the city level, mainly due to data limitations. To close this gap, this study is going to assess this relationship within U.S. metropolitan statistical areas (MSAs) from 1980–2016.

Few studies have analyzed the income-inequality relationship at this scale. For U.S. 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: increases in income are associated with increases in inequality (Castells-Quintana, Ramos, & Royuela, 2015; Rodríguez-Pose & Tselios, 2009).

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 to compare against 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; U.S. Census Bureau, n.d.-a). The second is a decennial panel over 1980–2000 using U.S. Census data (Manson, Schroeder, Van Riper, & Ruggles, 2017; Ruggles et al., 2018). Thereby, this paper expands the time horizon for local-level studies on the income-inequality relationship up until 2016.

This paper finds that the income-inequality relationship changes over time. A higher per capita income level was still associated with a lower within-MSA inequality level in the earlier years. However, this association stopped being statistically significant in 2000 and remained insignificant for all the following years. For the 1980–2000 panel, per capita income increases are accordingly associated with decreases in inequality. In the 2006–2016 panel, in contrast, an increase in per capita income is associated with an increase in inequality. The income-inequality relationship changed its 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 and 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 next 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 ones. Sections 6 and 7 present robustness checks using alternative inequality and income measures. Section 8 discusses potential reasons for the change in sign of the income-inequality relationship while section 9 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 later on augmented this theory, stating that after a certain point, inequality starts increasing again with income for highly-developed economies (Castells-Quintana et al., 2015; Conceição & Galbraith, 2001).

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 lead to both economic growth and increased inequality (Autor & Dorn, 2013; Partridge & Weinstein, 2013; Rigby & Breau, 2008). 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 both lower economic growth and higher inequality (Li, Campbell, & Fernandez, 2013; Florida & Mellander, 2015) while crime and sociopolitical unrest hinder economic growth while both 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 the new situation. No explicitly short-run theory about the income-inequality relationship exists to the best of the author's knowledge. 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, Oechslin, & Zweimüller, 2014).

An MSA's population size, education levels, 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 (Baum-Snow & Pavan, 2012; Glaeser et al., 2009). 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 (Glaeser et al., 2009; Perugini & Martino, 2008). 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 when wage inequality is controlled for 2010 (Florida & Mellander, 2016). Higher per capita income growth appears to lead to lower end-of-period inequality in 1990 (Bhatta, 2001). However, cross-sections 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 in contrast with panel studies that assess how changes in income levels result in inequality changes for a given MSA (Atems, 2013; Partridge, 2005). 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. Per capita income 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 U.S. MSAs due to the differing

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

labor market and institutional context, influencing the income-inequality relationship. Furthermore, MSAs provide for both 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 both 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.

This study employs hence two distinct data sets. One with annual data from 2006–2016 and one with decennial data for 1980–2000.

For the 2006–2016 data set, the data stem from the 1-year ACSs collected by the U.S. Census Bureau. The data for all the main variables were retrieved from FactFinder (U.S. Census Bureau, n.d.-a). This data includes the pretax household income GINI at the MSA level. All ACS income variables are for the past 12 months prior to the interview moment, which is not publicly disclosed (Peters, 2013; IPUMS-USA, n.d.-b; U.S. Census Bureau, 2009). This paper converts all original income variables into 2010 U.S.-\$ using the conversion factors provided by the Integrated Public Use Microdata Series U.S.A. (IPUMS) to adjust for inflation (IPUMS-USA, n.d.-b). Table 1 presents descriptive statistics. The resulting panel dataset consists of 399 MSAs and 11 years. It is unbalanced due to the slight further delineation changes over the time period.

For the 1980–2000 data set, the data stem from U.S. Census via NHGIS and IPUMS (Ruggles et al., 2018; Manson et al., 2017). 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 popu-

²An MSA is a geographic entity delineated by the Office of Management and Budget for use by U.S. 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 (U.S. Census Bureau, n.d.-b).

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	0.561
per capita income	4069	24738	4423	4223	1175	51661
mean household income	4069	63444	11527	11193	2938	139718
median household income	4069	47992	8796	8529	2453	99965

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

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	0.532
per capita income	735	23906	4379	3742	2425	42928
median household income	735	51413	8406	7847	3087	97304

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

lations are incompletely identified in the IPUMS datasets (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 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.

The resulting unbalanced panel data set consists of 260 MSAs and 3 years. Table 2 presents descriptive statistics.

This paper estimates the income-inequality relationship in cross-sections and in 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 hence as follows:

$$g_{it} = \alpha + \beta y_{it} + \gamma X_{it} + \mu_i + \tau_t + \varepsilon_{it}, \quad (1)$$

where g_{it} is a measure of inequality for MSA i at time t , y_{it} is an income measure (in logs),

³A 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).

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. The cross-sections exclude the MSA and time FEs, and are only estimated for a given t .

Controls for population, education, and sector employment shares are included to avoid confounding factors. They have been shown to influence within-city inequality, as previously discussed.⁴

Reverse causality between income and inequality constitutes an issue in these regressions. Income influences inequality, but inequality, in turn, affects income and income growth. Convincing instruments for income levels 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.

4 Cross-Section Results

This section presents cross-section results using both data sets. These results present a first starting point to assess the income-inequality relationship across time.

The results can be found in table 3. The first three columns report regression results for the years 2006, 2010, and 2016. These regressions use the 2013 MSA delineations. The data stems from the ACSs via FactFinder. The last three columns report regression results for the years 1980, 1990, and 2000. These regressions use the 1990 MSA delineations. The data stem from the Census via NHGIS and IPUMS.

For the years 2000, 2006, 2010, and 2016, the income coefficient is not statistically significant even at the 10 % level. Per capita income levels appear not to influence inequality levels in these years: neither positively nor negatively.

The income coefficient is statistically significant at the 1 % level and negative in 1980 and 1990, however. Higher per capita income levels appear to be associated with reduced inequality levels in these years. A 1 % increase in per capita income 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.

Possible reasons for the divergence in results include differences in the database, the MSA delineations' changes, and qualitative changes in the income-inequality relationship over time. They are discussed more in detail in section 8.

The control variables' coefficients are typically of the expected signs. However, population has surprisingly a statistically insignificant coefficient. Thus, the MSA size does not seem to influence the inequality level in the considered context. The coefficient of the

⁴Quadratic terms for the income variables were also included in the regressions to test for quadratic relationships. Their coefficients are rarely statistically significant, however.

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(per capita income)	-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)
baplus	0.001*** (0.000)	0.002*** (0.000)	0.002*** (0.000)	0.077** (0.036)	0.101*** (0.038)	0.050** (0.023)
hsplus	-0.003*** (0.000)	-0.003*** (0.000)	-0.003*** (0.000)	-0.223*** (0.063)	-0.236*** (0.049)	-0.170*** (0.028)
sagr	-0.002*** (0.001)	-0.002** (0.001)	-0.002*** (0.001)	-0.249*** (0.088)	-0.140** (0.069)	-0.097*** (0.035)
sman	-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)
<i>MSAs</i>	382	366	359	251	245	239
<i>R</i> 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. baplus is the population share with a bachelor's degree or higher (in percent). hsplus is the population share with a high school diploma or higher (in percent). sagr is the share of persons 16 years and over employed in agriculture (in percent). sman is the share of persons 16 years and over employed in the manufacturing sector (in percent.)

Robust standard errors are in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$

Source: FactFinder, NHGIS and IPUMS as well as own calculations

share having a bachelor's degree or higher is statistically significant at the 5 % level and positive. Conversely, the coefficient of the share having a high school diploma or higher is statistically significantly negative. Thus, a better-educated population tends to be associated with higher inequality. These results correspond to the predictions and the findings obtained by Glaeser et al. (2009). The coefficients of both the share employed in agriculture and the share employed in manufacturing are statistically significant and negative. This indicates that an economic structure based on these sectors is associated with less inequality than a service-based economy. The coefficient sizes of all the control variables are tiny.

The control variables do not drive the results as similar results are obtained when excluding them from the regression. The negative income-inequality relationship persists for 1980 and 1990 when the GINI is only regressed on per capita income. The absolute coefficient size even increases slightly. The income coefficient is, in this case, also statistically significantly negative in 2000. For the years 2006, 2010, and 2016, the income coefficients remain not statistically significant as previously.

5 Panel Results

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

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

For the 2006–2016 panel, the income coefficient is statistically significant at the 1 % level and positive in both regressions. Increases in mean income appear to lead to increases in inequality. 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 of the GINI. These results correspond to the ones obtained for European regions in annual panels over the 1990s and 2000s (Castells-Quintana et al., 2015; Rodríguez-Pose & Tselios, 2009).

For the 1980–2000 panel, the income coefficient is statistically significantly negative. Over these years, an increase in mean income seems to have decreased inequality. The absolute size of the income coefficient is smaller than previously. A 1 % increase in per capita income involves, *ceteris paribus*, a decrease in the GINI by 0.0007 points for a given

⁵Mean household income is not available for the latter sample.

Table 4: Panel Results Regressing Inequality on Income

	2006-2016		1980-2000
	(1) gini	(2) gini	(3) gini
ln(per capita income)	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)
baplus	-0.001** (0.000)	-0.000* (0.000)	0.150** (0.059)
hsplus	-0.001*** (0.000)	-0.001*** (0.000)	-0.043 (0.035)
sagr	-0.001 (0.000)	-0.000 (0.000)	-0.000 (0.072)
sman	-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 FEs	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 results for the 2006–2016 annual panel while the third column reports results for the 1980-2000 decennial panel.

baplus is the population share with a bachelor's degree or higher (in percent). hsplus is the population share with a high school diploma or higher (in percent). sagr is the share of persons 16 years and over employed in agriculture (in percent). sman is the share of persons 16 years and over employed in the manufacturing sector (in percent)

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

MSA. This decrement is equivalent to a decrease by about 0.2 % at the mean of the GINI. However, the within- R^2 increased considerably from 0.29 before to now 0.85.

This divergence in the obtained income-inequality relationships might be due to similar reasons as the divergence in cross-section results for these data sets: differences in the database, the changes in the MSA delineations, and qualitative changes in the income-inequality relationship over time. Besides, the 2006–2016 panel is a yearly one, whereas the 1980–2000 panel is a decennial one. The 2006–2016 panel has observations from 11 time periods, whereas the 1980–2000 one only has three.

The control variables' coefficients also change compared to the cross-section regressions. Population now exhibits a statistically significant negative coefficient. Thus, increases in MSA size seem to decrease inequality, whereas the population level per se does not affect an MSA's inequality level. The coefficient of the share having a bachelor's degree or higher is statistically significantly negative in the 2006–2016 panel but remains statistically significantly positive in the 1980–2000 panel. The coefficient of the share having a high school diploma or higher remains statistically significantly negative in the 2006–2016 panel but is not significant in the 1980–2000 one, providing for mixed results. The coefficient of the share employed in agriculture is not statistically significant in both panels, whereas the share employed in manufacturing remains statistically significantly negative in both panels. This coefficient indicates that deindustrialization is indeed associated with increasing inequality. The coefficient sizes of all the control variables remain tiny. The obtained results are again robust to excluding all control variables from the regression. Both the positive income-inequality relationship in the 2006–2016 panel and the negative one in the 1980–2000 panel persist.

6 Employing Alternative Inequality Measures

The obtained opposing results for the two data sets might result from a peculiarity of the GINI. Therefore, the previous regressions have been 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 have biased 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

⁶Robustness checks have also been run for the cross-sections with similar results, indicating that their results are overall robust as well. The results are available upon request.

⁷Regressions have also been conducted 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.

Table 5: Alternative Inequality Measures in the 2006–2016 Panel

	(1)	(2)	(3)	(4)	(5)	(6)
	gini	ge0	p90p10	p90p50	p50p10	s1
ln(per capita income)	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*

- 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 the bottom of the income distribution. The 90/50 percentile ratio measures the inequality within top incomes, while the 50/10 percentile ratio measures 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 2856 (from 4069 before) and in the 1980–2000 data set to 700 (735 before).

The alternative inequality measures are replacing the GINI as the dependent variable in the regressions. The results can be found in table 5 for the 2006–2016 panel and in table 6 for the 1980–2000 panel.

For the 2006–2016 panel, GE(0) shows very similar results to the GINI ones: 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 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 the expec-

Table 6: Alternative Inequality Measures in the 1980–2000 Panel

	(1)	(2)	(3)	(4)	(5)	(6)
	gini	ge0	p90p10	p90p50	p50p10	s1
ln(per capita income)	-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
<i>N</i>	700	700	700	700	700	700
<i>MSAs</i>	254	254	254	254	254	254
<i>within</i> – 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: NHGIS and IPUMS as well as own calculations

tations as they only consider parts of the income distribution.

The use of these alternative inequality measures also allows distinguishing between two hypotheses, which have been discussed for the rising inequality in the U.S.: a rise in the top income share and polarization (Autor, Katz, & Kearney, 2006; Essletzbichler, 2015; Piketty & Saez, 2003). Per capita income 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. Particularly the 2006–2016 panel could exhibit this pattern as it captures the time of technological change substituting middle-skill routine tasks, leading to polarization.

The obtained results hint towards polarization but cannot substantiate this hypothesis 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 per capita income increases disproportionately benefited the very top incomes. Concurrently, the 90/50 percentile ratio turned insignificant positive from being significant 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. 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 the bottom incomes as suggested by the polarization hypothesis.

7 Employing Median Income

This section evaluates whether controlling for the gap between mean and median income results in the changing signs in the income-inequality relationship across the two panels as well. This exercise can also enlighten further whether technological change leading to polarization drives this change in sign.

Only including median income into the regressions would not produce meaningful results as its relationship with inequality is statistically predetermined to be negative, unlike mean income's. MSAs exhibit right-skewed income distributions of their inhabitants' incomes. An increase in the median income of a right-skewed income distribution leads to a decrease in inequality by reducing the gap to the higher mean income. An increase in the mean income might lead in this context to higher inequality but not necessarily so, depending on which income group drives the increase.⁸ An increase in the difference between mean and median income as well as in the ratio of mean to median income should increase inequality in a right-skewed distribution. Thus, a positive coefficient is expected. In the case of polarization, there should be less income mass around the middle of the income distribution. Thus, both per capita and median income increases should increase inequality for a given gap or ratio between the two. Conversely, if more income accumulates around the middle, income increases reduce inequality for a given gap or ratio between per capita and median income. If only the top incomes increase with rising income, then this would be captured by the per capita-median gap, and the single income measures' coefficients should not be significant on their own.

The gap, respectively, the ratio between per capita and median income have been added to the regressions to assess the polarization hypothesis. The obtained results can be found in table 7 for the 2006–2016 panel and in table 8 for the 1980–2000 panel.

Columns 1 and 2 include the difference between per capita and median household income. This difference exhibits statistically significant positive coefficients in both panels, as expected. Conditional on this difference, per capita and median household income's coefficients are statistically significantly positive in the 2006–2016 panel. In contrast, they are statistically significant and negative in the 1980–2000 panel.

Columns 3 and 4 include the ratio between per capita and median household income. This ratio also exhibits statistically significant positive coefficients in both panels, as expected. Conditional on this ratio, the income coefficients are again statistically significantly positive in the 2006–2016 panel and negative in the 1980–2000 panel.

These opposing signs of the income coefficients confirm the opposing signs in the baseline panel analyses. The mean income coefficient was previously already positive in the 2006–

⁸This is confirmed empirically for both the cross-section and the panel analyses. On its own, median income always exhibits statistically significant negative coefficients. When both income types are included, the coefficients are as expected: always positive for per capita and always negative for median income. Results are available upon request.

Table 7: Mean and Median Income in the 2006–2016 Panel

	(1)	(2)	(3)	(4)
	gini	gini	gini	gini
ln(per capita income)	0.049*** (0.006)		0.033*** (0.006)	
ln(median household income)		0.049*** (0.006)		0.031*** (0.006)
difference	0.250*** (0.007)	0.299*** (0.008)		
ratio			2.851*** (0.075)	3.200*** (0.086)
Controls	yes	yes	yes	yes
MSA & Time FEs	yes	yes	yes	yes
Constant	yes	yes	yes	yes
N	4069	4069	4069	4069
$MSAs$	399	399	399	399
$within - R^2$	0.618	0.618	0.618	0.618

Difference is the difference between per capita and median household income.

Ratio is the ratio between per capita and median household income.

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

2016 panel and negative in the 1980–2000 panel as well.

The observed pattern strongly hints at polarization occurring in the 21st century. In contrast, middle incomes appear to have disproportionately benefited from income increases before, as demonstrated by the 1980–2000 panel results.

8 Reasons for the Change in the Income-Inequality Relationship

There are four possible reasons why the mean 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 U.S. Census Bureau according to similar standards. Furthermore, the 2006–2016 results

Table 8: Mean and Median Income in the 1980–2000 Panel

	(1)	(2)	(3)	(4)
	gini	gini	gini	gini
ln(per capita income)	-0.070*** (0.016)		-0.081*** (0.016)	
ln(median household income)		-0.070*** (0.016)		-0.075*** (0.015)
difference	0.133*** (0.022)	0.062** (0.025)		
ratio			1.534*** (0.252)	0.655** (0.274)
Controls	yes	yes	yes	yes
MSA & Time FEs	yes	yes	yes	yes
Constant	yes	yes	yes	yes
N	735	735	735	735
$MSAs$	260	260	260	260
$within - R^2$	0.862	0.862	0.862	0.862

Difference is the difference between per capita and median household income.

Ratio is the ratio between per capita and median household income.

Standard errors clustered at the MSA level in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: NHGIS and IPUMS as well as own calculations

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 also 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 hence considerable. However, 260 MSAs are a large enough number of observations for regression analyses. Besides, the concept of MSAs remained constant across the data sets. 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. (Results are available upon request). Thus, delineation and sample size changes might play a role in the diverging results, but they appear unlikely to be the opposing results' sole cause.

Third, the time gaps and time dimensions of the panels differ as well. 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. The latter might result in statistical issues in FE-estimations as there might not be enough within-variation for proper estimation. The 10-year gap between observations also 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 also result in a positive income-inequality relationship. The latter comprise in contrast segregation, crime, and sociopolitical contrast and hence exactly those factors leading to a negative income-inequality relationship. Annual panel studies for European regions found as well 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 years (three). When regressing in this panel the GINI on mean income and the usual controls, the income coefficient remains statistically significant and positive for both per capita and mean household income. However, its size diminishes by about one third. (Results are available upon request.) Thus, there appears to be something special about the 2006–2016 time period rather than the time gap between observations and the number of observed years resulting in the positive income-inequality relationship. However, the latter cannot be excluded entirely 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 to 2006. The cross-sections also reflect this change. The negative income-inequality association stops already in 1990 and does not exist anymore for 2000 and all the further years.⁹ This timing corresponds to the sharp rise in inequality generally observed in the U.S. 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, not only inequality increased, but its relationship with income changed as well. The changed sign of the income-inequality relationship also hints at economic growth having become less inclusive over the years.¹⁰

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 segrega-

⁹The European panel studies finding a positive income-inequality relationship analyzed the 1990s and 2000s (Castells-Quintana et al., 2015; Rodríguez-Pose & Tselios, 2009).

¹⁰The 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 not statistically significant for the first time. The potential crisis effect cannot be ephemeral either as the positive income-inequality association also appears in a 2012–2016 panel, starting after the crisis years.

tion, crime, and sociopolitical unrest as detailed in section 2. Crime rates indeed declined for several offenses since the 1980s (Asher, 2017) but residential segregation increased during the considered time 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 increased since the 1980s due to globalization and technological change substituting middle-skill routine tasks (Autor & Dorn, 2013; Autor et al., 2006; Rigby & Breau, 2008). Unionization rates declined over the last decades (Hu & Hanink, 2018). All these developments would strengthen a positive income-inequality relationship. Combined, they might have lead to the observed change in the sign of the income-inequality relationship if the importance of these positive factors was 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 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.

9 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 studying the income-inequality relationship within MSAs over a more extended period than previously possible as well as employing cross-section and panel regression techniques.

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

The main explanations for this change in sign consist of MSA delineation changes and different time dimensions in the panels as well as qualitative changes in the income-inequality relationship. The latter are most notably due to polarization resulting from technological change substituting middle-skill routine tasks in line with Autor and 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 upon the influence of specific factors on this relationship at different periods.

References

- Asher, J. (2017). *The U.S. murder rate is up but still far below its 1980 peak*. Retrieved 2021-02-05, from <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 U.S. county-level data. *Journal of Regional Science*, *53*(4), 656–671. doi: 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. doi: 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. doi: 10.1257/000282806777212620
- Baum-Snow, N., & Pavan, R. (2012). Inequality and city size. *Review of Economics and Statistics*, *95*(5), 1535–1548. doi: 10.1162/REST_a_00328.
- 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. doi: 10.1111/0735-2166.00093
- Bischoff, K., & Reardon, S. F. (2014). Residential segregation by income, 1970-2009. In J. Logan (Ed.), *Diversity and disparities: America enters a new century*. New York, NY, USA: The 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. doi: 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. doi: 10.1007/s10037-015-0098-4
- 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, UK: Cambridge University Press. doi: 10.1017/CBO9781139175210.008

- Essletzbichler, J. (2015). The top 1% in U.S. metropolitan areas. *Applied Geography*, *61*, 35–46. doi: 10.1016/j.apgeog.2015.01.019.
- Florida, R., & Mellander, C. (2015). *Segregated city : the geography of economic segregation in America's metros*. Toronto, Canada: 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. doi: 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. doi: 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. doi: 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. doi: 10.1007/s10887-013-9099-8
- Hu, Q., & Hanink, D. M. (2018). Declining union contract coverage and increasing income inequality in U.S. metropolitan areas. *The Professional Geographer*, *70*(3), 453–462. doi: 10.1080/00330124.2017.1416298
- IPUMS-USA. (n.d.-a). *IPUMS USA: incompletely identified metropolitan areas*. Retrieved 2021-02-05, from <https://usa.ipums.org/usa/volii/incompmetareas.shtml>
- IPUMS-USA. (n.d.-b). *IPUMS-USA: INCTOT*. Retrieved 2021-02-05, 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. (<http://www.jstor.org/stable/1811581>)
- Li, H., Campbell, H., & Fernandez, S. (2013). Residential segregation, spatial mismatch and economic growth across US metropolitan areas. *Urban Studies*, *50*(13), 2642–2660. doi: 10.1177/0042098013477697
- Madden, J. F. (2000). *Changes in income inequality within U.S. metropolitan areas*. Kalamazoo, MI, USA: 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]*. Minneapolis, MN, USA: University of Minnesota. doi: 10.18128/D050.V12.0.
- Partridge, M. D. (2005). Does income distribution affect U.S. state economic growth? *Journal of Regional Science*, *45*(2), 363–394. doi: 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. doi: 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. doi:

- 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. doi: 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. doi: 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. doi: 0.1080/00045600802219786
- 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. doi: 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. doi: 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]*. Minneapolis, MN, USA: IPUMS. doi: 10.18128/D010.V6.0
- U.S. Census Bureau. (n.d.-a). *American FactFinder*. Retrieved 2020-07-23, from <http://factfinder2.census.gov>
- U.S. Census Bureau. (n.d.-b). *Glossary*. Retrieved 2021-02-05, from <https://www.census.gov/programs-surveys/metro-micro/about/glossary.html>
- U.S. Census Bureau. (2009). *A compass for understanding and using American Community Survey data: what researchers need to know*. Washington, DC, USA: U.S. Government Printing Office. (<https://www.census.gov/library/publications/2009/acs/researchers.html>)