Institute for International Economic Policy Working Paper Series
Elliott School of International Affairs
The George Washington University

Is There a Kuznets Curve for Intra-City Earnings Inequality?

IIEP-WP-2018-9

Haixiao Wu
George Washington University

November 2018

Institute for International Economic Policy
1957 E St. NW, Suite 502
Voice: (202) 994-5320
Fax: (202) 994-5477
Email: **iiep@gwu.edu**
Web: **iiep.gwu.edu**

Is There a Kuznets Curve for Intra-City Earnings Inequality?

Haixiao Wu
Department of Economics, The George Washington University
Washington DC, USA

November 10, 2018

Abstract

Many papers have found a positive relation between income inequality and city size in the US and other countries. This literature has assumed that the relation is linear. Tests performed here find that it is concave, resembling the classic Kuznets curve. A theoretical model based on the Income Elasticity Hypothesis (IEH), explains that inequality is a concave function of housing prices that tend to increase with city size. Further tests confirm the concavity of the relation between Gini and housing costs that is predicted by the IEH. Although for most cities, inequality still rises with housing costs, if housing costs continue to grow in large cities, inequality should eventually fall, resembling the Kuznets Curve at the country level.

1 Introduction

Inequality was once thought of as a byproduct of economic growth. Dating back to the 1950s, Kuznets(1955)[19] formalized a hypothesis about growth and income inequality, which was summarized as the Kuznets Curve. His hypothesis was that economic growth first increases and then decreases inequality. The Kuznets Curve initiated a lively literature including notable theoretical advancements that enhanced understanding of national-level patterns of inequality(e.g. Morrisson and Murtin, 2013[24]; Galbraith, 2007[12]).

In the last two decades, the study of earnings inequality in cities has gained attention. Economists represented by Glaeser et al.(2009)[14] find a strong positive correlation between inequality(the Gini coefficient) and city population in the U.S..[1] The positive correlation is also found in Canada(Bolton and Breau, 2012)[5] and Great Britain(Lee et al., 2016)[20].

There have been several suggestions about the determinants of city inequality. Empirically, researchers represented by Madden(2000)[22], Bolton

[1] The correlation becomes negative when they control for the initial distribution of skills.

and Breau(2012)[5], emphasize the demographic heterogeneity across cities, including race, gender, education,..., etc. Other determinants include unionization(Florida et al., 2016 [11]), local taxes, policies, and amenities(Gyourko et al., 2013[15]). A final possibility is industrial composition. Many of these factors vary with city size, which has given rise to several propositions to explain the positive correlation between city size and inequality.

One argument for the positive correlation is known as the monopoly argument. As initially proposed by Haworth et al. (1978)[16], economic agents unequally benefit from the development of cities, favoring highly specialized agents with monopoly advantage in the marketplace.

Another important view is the productivity-agglomeration hypothesis. This hypothesis first assumes agglomeration economies in cities mainly augment the productivity of the most educated or skilled workers so that pay inequality rises with city size. Because the supply of unskilled workers to cities is assumed to be perfectly elastic, it follows that returns to skill rise with city size, which is the reason for the positive relation between size and inequality. This view is supported by Wheeler(2005)[27] and Baum-Snow and Pavan(2013)[2]. Baum-Snow and Pavan(2013) [2] attribute the increasing returns to skill with city size to the disproportionate concentration of high skill industries in large cities. As an important addition, Behrens and Robert-Nicoud(2014)[3] develop a model that explains how agglomeration economies and firm selection could cause labor in large cities to be more productive.

However, as found by Black et al.(2009)[4], returns to college fall with city size. Specifically the ratio of earnings of college graduates to those of non-college graduates, called the skilled wage ratio (SWR) tends to fall with city size. At the same time, there is a tendency for skilled workers, measured by the ratio of college to non-college graduates or the skill intensity ratio (SIR) to concentrate in larger cities where returns to education are lower but house prices are higher. Moretti(2013)[23] has noted that this means that real wage inequality is smaller than nominal wage inequality. This finding is also supported by Wheeler(2006)[28] for within-job wage changes, and Glaeser et al. (2009)[14].It appears that understanding differences in earnings inequality across cities, requires a model that considers differences in both the SWR and SIR. Empirical tests based on differences in differences find that the SWR is lower in large cities and that this should promote earnings equality while the increasing SIR is a contravening force promoting higher earnings inequality in large cities. This suggests that, for inequality to rise with city size, the fall in the SWR with size must be outweighed by a rise in the SIR. If changing SWR and SIR are an important determinant of the change in city Gini coefficients with size, it appears possible that this change is not constant as has been implicitly assumed in the empirical literature. As SIR exceeds 1, the contribution of rising SIR to inequality actually changes sign from positive to negative because the majority of city residents are college educated. Thus, the rise in inequality with size may be limited and actually be concave, giving rise to a Kuznets curve in the city size - income inequality relation.

Because of the potential coexistence of forces promoting and discouraging a positive association between inequality and city size, there is no guarantee that the forces for rising inequality with city size should always dominate. However, empirical testing reported in the literature implicitly assumes a linear relation between inequality and city size, while arguments made above suggest a nonlinear function. Most discussion of the theoretical relation between city size and inequality has focused exclusively on theories that explain the positive correlation reported in the literature assuming linearity. These theoretical arguments for increasing inequality are generally plausible and will not be questioned in this research. However they are also incomplete because they ignore countervailing forces for which there is a clear theoretical rationale.

The academic literature suggesting that inequality rises with city size has led to policy recommendations to reverse its effects. Liu(2018)[21] calls for "driving the value of economic inclusion into organizational missions". Bolton and Breau (2012)[5] attribute rising inequality to structural changes in the labor force, and believe that policies that revitalize manufacturing industries in larger cities should help.

This paper extends, for the first time, the current literature on the relation among city size, housing cost, the SIR and the SWR to develop implications for the variation in the Gini coefficient. The result provides a new theoretical rationale for there being a relation between inequality and city size. What is different in these theoretical implications is the hypothesis that the relation is concave. While other forces identified in the literature may promote a monotonic positive association between size and inequality, the mechanism proposed in this paper indicates that rising inequality with city size is not permanent. Because the function is concave, there is a point where inequality starts to fall with city size, which resembles the Kuznets curve.

If there is a Kuznets curve for intra-city earnings inequality across cities, it has implications for policy making. Based on the current literature, there is a possibility that the rising inequality in large cities is due to some form of market failure. This might imply that some modification of migration incentives could reduce income inequality. Alternatively, if the Kuznets curve of intra-city inequality is concave and the association between size and inequality will eventually attenuate, the argument for interventions to correct a market failure is not appropriate. Indeed, the association between city size and inequality currently observed may indicate that intercity labor markets are working efficiently.

This paper begins by revisiting the relation between inequality and city size, followed by setting a theoretical framework for the analysis of inequality and housing costs on the basis of the Income Elasticity Hypothesis. Then a number of tests for the possibility of a concave Kuznets curve are conducted. Overall the evidence suggests that rather than the current linear relation between city size and inequality, the relation is concave and that housing cost effects play an important role in the shape of the curve.

2 Revisiting the Relation between Inequality and City Size

This section examines the positive correlation between inequality and city size, which has been noted in the literature discussed in the introduction. To be consistent with other studies, the Gini coefficient is used as the measure of inequality, and population as the measure of city size.

2.1 Method

In the cross-sectional estimation reported in Glaeser et al.(2009)[14], the Gini coefficient at city level is considered to be determined by the following model:

$$Gini_i = \alpha + X_i\beta + \varepsilon_i \tag{1}$$

where, the dependent variable is the Gini coefficient in MSA i; X_i is a vector of control variables(including MSA population) for MSA i.

Different variations of the specification in equation 1 are used to test the association between inequality in cities and city size. Similar models are found in Bolton and Breau (2012)[5], Florida et al.(2013)[11], Lee et al. (2016)[20], etc.

The independent variable is either the Gini coefficient of annual income, or the Gini coefficient of annual labor earnings. The Gini coefficient of total income has the potential problem that it may reflect the concentration of the ultra-rich in big cities. Because income from capital and other sources is usually not from the location of residence, the spatial concentration of the ultra-rich can produce misleading results. This paper analyzes the inequality of wage earnings in both the theoretical framework and the empirical part.

The control variable of interest, city size, is usually represented by MSA population.

2.1.1 Problems of the Existing Empirical Model

The empirical model in equation 1 is widely used for testing the relation between inequality and city size. But the coefficient on city size is not not generally based on a formal theoretical model generating the equation specification. The important point of the testing here is to determine if the implicit assumption in previous literature that the relation between city size and city Gini is linear as opposed to some alternatives, specifically that it is concave.[2]

[2]The literature often includes additional control variables including MSA economic indicators(the percentage of labor force employment in primary /secondary/tertiary industries, unemployment rate, median income, tax and welfare, wage inequality, poverty, unionization) and demographic ratios(the education attainment ratios, age group ratios, race and immigrant ratios). These are bad controls as they are correlated with unobserved missing variables in the error term. With that in mind, this paper just empirically tests the correlation between Gini coefficient and city size, without controlling for possible forces that may be in play. Then use the theoretical model to justify that housing costs differentials lead to differentials in Gini coefficient.

The revised empirical model estimated in this paper is shown below:

$$Gini_{i,t} = \alpha + \sum_t \beta_t D_t * N_{i,t} + \sum_t \gamma_t D_t * N_{i,t}^2 + \sum_t \delta_t D_t + \varepsilon_{i,t} \tag{2}$$

where $N_{i,t}$ is the population in MSA i, in year t, and $D_t's$ are year fixed effect dummies($D_t = 1$ if year=t, t$\in$ [1980,1990,2000,2010]). The year fixed effects account for the changes in the national baseline level of Gini coefficient(or the constant term) over time. Interacting D_t with population allows for the coefficient on city size, β, to change over time. The square of population is added to the regression when testing the nonlinearity of the Gini coefficient as a function of population. α is the constant term.

In the empirical test, following the work of Wheeler(2004)[26], total employed population is used as the weight in the Weighted Least Squares estimation. To account for interdependency of MSAs within a year, errors are clustered by year.

2.2 Data

The IPUMS provides decennial census micro-data at individual level that includes information about education attainment, wage and salary income and geographic location. Because annual Data is available after 2000, for analysis over a longer period, decennial census data is more appropriate. In this paper, 4 rounds of Census micro-data are used: 1980 5% state; 1990 5% state; 2000 5%; 2010 ACS.

Based on definitions for metropolitan statistical areas (MSAs) from the U.S. Office of Management and Budget (OMB), metropolitan areas are given by variable "metarea". MSA statistics(SIR, SWR, population,..., etc) are summarized by metropolitan area for our city level analysis, based on individual level census data. The resulting dataset at MSA level consists of all MSAs defined by OMB.

2.2.1 Definition: Skilled and Unskilled

A worker is defined as a skilled if she/he has completed at least 4 years of college education, and unskilled otherwise. The skill intensity ratio (SIR) is the production input ratio of skilled to unskilled workers, so it is empirically calculated as the ratio of the number of employed skilled workers (s) over the number of employed unskilled workers(u). SIR=s/u.

In the theoretical model, skilled and unskilled workers are respectively homogeneous. In the empirical part, it is assumed that each worker earns the median yearly wage earnings of the corresponding skill level sub-population. $SWR = w_s/w_u$, where w_s and w_u are median wages of skilled and unskilled workers respectively.

2.3 Empirical Results

Descriptive statistics of key variables are shown in Table 1.

Table 1: Descriptive Statistics, pooling 4 rounds

Variable	Obs	Mean	Std. Dev.	Min	Max
Gini in %	1,070	13.68665	4.013528	4.451984	33.75492
SIR	1,070	.2989408	.1357844	.087949	1.103603
SWR	1,070	2.017934	.2882483	1.452801	4.082278
population in millions	1,070	.7299617	1.520396	.09966	17.76168
CEO rent index	791	3.00226	1.325659	1.72979	17.18437
Albouy index	240	.8960399	.226285	.5655255	2.247908

Table 2: The Relation between Gini and City Population

	(1) Gini in %	(2) Gini in %	(2') population at the peak
D1980 × population in millions	0.105***	0.260***	
D1980 × population squared		-0.0112***	11.6
D1990 × population in millions	0.0953***	0.313***	
D1990 × population squared		-0.0147***	10.6
D2000 × population in millions	0.285***	0.762***	
D2000 × population squared		-0.0294***	13
D2010 × population in millions	0.309***	0.860***	
D2010 × population squared		-0.0331***	13
Constant	10.46***	10.28***	
R-squared	0.614	0.639	
Year FE	YES	YES	
Observation	1070	1070	

Errors are clustered by year. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2 Column (1) reports results of OLS estimates pooling all observations at MSA level from decennial census 1980-2010. The relation is assumed to be linear in the model in Column (1). Consistent with the literature(Glaeser et al.(2009)[14], Bolton and Breau (2012)[5] and others), there is a strong correlation between inequality and population. On average, in 2010, an increase in MSA population of 1 million leads to a 0.31% of rise in the Gini coefficient.

In Table 2 Column (2), the square of population (in millions) is added to test the nonlinearity of the relation.This contrasts with previous literature which has assumed that the relation is linear. The results suggest significant concavity of the relation between population and Gini. As shown in Column (2'), the peak of the Gini coefficient as a function of city size is at the population of 11.6 million, 10.6 million, 13 million, 13 million respectively in year 1980, 1990, 2000, 2010. In 2010, for instance, the population in Los Angeles(12.8 million) is close to the peak of the curve, and population in New York is on the right-hand side of the peak. Other than Los Angeles and New York, most MSAs are on the left-hand side of the curve in which the Gini coefficient is increasing with population. However, the results in Table 2 demonstrate that the relation between inequality and city size in U.S. cities is concave rather than linear.[3]

In the theoretical framework below, it is argued that it is the housing cost that is the primary source of concavity, rather than population.

3 Theoretical Framework

The Income Elasticity Hypothesis (IEH) developed in a series of papers in the literature explains why SIR rises and SWR falls with city size. The extension of the the theory provided here traces implications of this process for changes in the Gini coefficient with city housing price and size to the extent that size and price are associated. The empirical literature has established that SWR and SIR vary inversely. The mechanism is based on the supply side of the labor market. On the surface it appears strange that skilled (unskilled) labor is differentially located in cities where its relative return is low. Possible explanations outside the IEH have appeared in the literature. Gyourko et al.(2013)[15] provides support by arguing that amenities in super cities are differentially attractive to more skilled/ educated workers. The industrial composition is also directly associated with SIR in cities. However, empirical papers do not find that the industrial composition plays a major role in the variations in SIR and SWR across cities. (See Elvery (2010)[10], Hendricks (2011)[17] and Brinkman (2014)[6].)

The IEH provides a theoretical rational for the decline in SWR and rise in SIR with city size, and the theoretical connection with earnings inequality is advanced here. The existing theoretical and empirical work provides a number of useful starting points. Wheeler(2005)[27] documents the association of earnings inequality and city size. Florida et al.(2016)[11] report that earnings inequality is composed of differences in SWR and SIR associated with city size. They

[3] To be consistent with the theoretical framework in this paper, the Gini coefficient calculated in the empirical part is based on the assumption that skilled and unskilled workers earn the median wage in their respective skill group. Specifically, the employed population is divided into two subgroups: people with 4 years of college education(skilled), and people without 4 years of college education(unskilled). This assumes people in each subgroup earn the same wage, represented by the median wage. Using median wage as the measure of wage in each subgroup takes account of the effect of the extremely rich workers on mean wage. Using mean wage as an alternative method does not affect the results much.

The Gini coefficient calculated in this paper is therefore based on two levels of wage earnings by education attainment. This assumption surely generates Gini coefficients that are underestimated, but it focuses on the portion of inequality that is solely due to wage earnings by skill, which is the purpose of this paper. The Gini coefficient of wage and salary earnings for a MSA without assuming two skill groups of workers is used as an alternative measure for robustness check. It is calculated based on census micro-data about individual wage and salary earnings. Limited by computation capability, the sample size used in the calculation of Gini coefficient is reduced to 0.1% of the population, as opposed to 5% in the full census sample. But 0.1% of the population is still a big sample size which generates estimates that are approximately equal to the correct value. Empirical tests based on this continuous Gini are performed with no substantial change in findings, as shown in Appendix D. The concave relation still exists in this case, though the coefficients are of less magnitude.

control for changes in median housing price and conclude that a positive increase in skilled wage inequality has contributed to the rise of income inequality with city size. Glaeser et al. (2009)[14] use a hedonic regression to create a housing price index, and find that SWR declines with city size. In sum, earnings and wage inequality and skill intensity are widely studied, but usually as separate topics. The variables usually said to account for these differences across cities are industry mix, amenities, unionization, and demographic characteristics.

The income elasticity hypothesis(IEH) in Kim, Liu and Yezer(2009)[18] provides an explanation for the differences in SWR and SIR across cities based on the Rosen(1974)[25] model of spatial indirect utility equilibrium. They provide empirical evidence that the income elasticity of demand for a primary residence is significantly less than unity, which means that housing costs are a greater proportion of income for unskilled workers in large cities. Inter-city utility equilibrium suggests that the compensating variation in wages for the high housing costs in large cities must be higher for unskilled workers. This explains the declining SWR without assuming differences in productivity, preferences, or demographic composition across cities. Broxterman and Yezer(2015)[7] further suggest that, in response to lower SWR in large cities, the IEH implies increasing SIR with housing costs as employers substitute skilled workers for unskilled workers.

The IEH has an important implication for earnings inequality within cities. The IEH suggests that housing cost rather than city population causes spatial differentials in earnings inequality. If housing cost in large cities rises relative to smaller cities, it first increases earnings inequality because of the rise in SIR, but eventually inequality should decrease as SWR decreases. In an extreme case, the economy can reach a point where SWR is extremely low and the majority of the population is skilled workers. In this case large cities would be uniformly high skill and the Gini coefficient of large cities in this extreme case would be low. This implies a bell-shaped curve of Gini coefficient as a function of city size(housing costs), much resembling the Kuznets curve.

This section constructs a theoretical framework based on the IEH to explain the underlying mechanism that leads to the concavity of the relation between the Gini coefficient and MSA population. Because city size is positively correlated with housing costs, the IEH suggests that it is housing price rather than population that causes the changes in Gini coefficient. The IEH relates the changes in SIR and SWR to changes in housing costs. The variation in Gini coefficient with housing costs is derived from combined changes in SIR and SWR. It will be shown here that the Gini coefficient is a bell-shaped curve with respect to housing costs. Other processes identified in the literature may also influence the relation between city size and inequality. However thus far these alternative explanations in the literature assume that the relation is linear and the distinguishing characteristic of the IEH is the prediction of concavity which was demonstrated empirically in the previous section of this paper.

3.1 Baseline City

In the economy, cities are indexed by i. Let i=m indicate the baseline(mean) city. Initially assume that all cities are identical. In the baseline city, identical firms produce a numeraire good with the following constant returns to scale CES production function.

$$F_m = A_m(\alpha s_m^{\frac{\sigma-1}{\sigma}} + \beta u_m^{\frac{\sigma-1}{\sigma}})^{\frac{\sigma}{\sigma-1}} \tag{3}$$

where s_m, u_m are inputs of skilled and unskilled workers, α and β are exogenous parameters, and σ is the absolute value of the elasticity of substitution. A_m denotes the productivity factor. The production function is identical in all cities.

Earnings of the skilled (unskilled) worker are given by $w_{s,m}(w_{u,m})$, as hours worked are assumed constant. Firms maximize profit so that

$$Max\Pi = F_m - w_{s,m}s_m - w_{u,m}u_m; \tag{4}$$

The FOCs for profit maximization are given by:

$$\begin{aligned} A_m^\sigma \alpha^\sigma(\alpha s_m^{\frac{\sigma-1}{\sigma}} + \beta u_m^{\frac{\sigma-1}{\sigma}})^{\frac{\sigma}{\sigma-1}} = w_{s,m}^\sigma s_m; \\ A_m^\sigma \beta^\sigma(\alpha s_m^{\frac{\sigma-1}{\sigma}} + \beta u_m^{\frac{\sigma-1}{\sigma}})^{\frac{\sigma}{\sigma-1}} = w_{u,m}^\sigma u_m. \end{aligned} \tag{5}$$

Taken together these FOCs imply that

$$(\frac{\beta w_{s,m}}{\alpha w_{u,m}})^{-\sigma} = \frac{s_m}{u_m}. \tag{6}$$

Let the national and representative city skill intensity ratio, $SIR_m = s_m/u_m$ and the national and representative city skilled wage ratio, $SWR_m = w_{s,m}/w_{u,m}$. Rearranging equation 6 yields

$$SWR_m = \frac{\alpha}{\beta} SIR_m^{-\frac{1}{\sigma}}; \tag{7}$$

With the production function given, the national SWR is a function of national SIR. The ratio of $\frac{\alpha}{\beta}$ can be calibrated with national SWR and SIR given by data, and the other parameter σ taken from the literature, which is explained in the simulation and calibration parts of the paper.

Because baseline city m is identical to other cities, the Gini in city m is equal to national average Gini coefficient.

The Gini coefficient is a function of SIR and SWR, given by: [4]

$$Gini_m = \frac{SWR_m * SIR_m}{SWR_m * SIR_m + 1} - \frac{SIR_m}{1 + SIR_m} = \frac{1}{1 + SIR_m} - \frac{\beta}{\beta + \alpha SIR_m^{1-\frac{1}{\sigma}}}. \tag{8}$$

As shown in equation 8, the average Gini coefficient is a nonlinear function of the national average SIR. The function is shown graphically in Section 4 as a bell-shaped curve (which is similar to the classic Kuznets Curve).

[4]Find derivation of the Gini coefficient expression in Appendix A.

3.2 City-level Gini Coefficient

In this section, the effects of changing housing price on the city labor market will be developed. Assume housing cost in city m increases by a factor of $(1+\pi)$ above housing cost in the baseline city. (A decrease in housing cost can be represented by cases where $-1 < \pi < 0$.) There must also be a compensating productivity or product price effect to compensate for the variation in housing cost above other cities.[5] Further assume that housing cost as a share of earnings is equal to H_u for unskilled workers in the average city. ε is the income elasticity of demand for a primary residence. The elasticities of SWR and SIR w.r.t. housing costs are determined according to the IEH [6]:

$$\begin{aligned} E_{SWR} &= \frac{\Delta(SWR)}{SWR}/\pi = (SWR_m - 1)(\varepsilon - 1)H_u/SWR_m; \\ E_{SIR} &= -\sigma E_{SWR} = -\sigma(SWR_m - 1)(\varepsilon - 1)H_u/SWR_m. \end{aligned} \tag{9}$$

Equation 9 is the central statement of the Income Elasticity Hypothesis, demonstrated in Kim et. al.(2009)[18]. As a result of the rise in housing cost, the SWR rises due to equation 9 and this generates a new SIR_i. Note that cities in the economy are assumed to be in spatial indirect utility equilibrium based on the Rosen model. The compensating variation in unskilled worker wages in response to a housing cost increase is relatively greater because the income elasticity of housing demand is well below unity. A change in city m's labor demand is assumed to be immediately satisfied by the perfectly elastic labor supply from the rest of the nation. Elasticity of SIR w.r.t. housing cost can also be expressed as:

$$E_{SIR} = \frac{SIR_i - SIR_m}{SIR_m}/\pi = -\sigma E_{SWR} \tag{10}$$

Combining equations 9 and 10 yields:

$$\begin{aligned} -\sigma(SWR_m - 1)(\varepsilon - 1)H_u/SWR_m &= \frac{SIR_i - SIR_m}{SIR_m}/\pi; \\ \Rightarrow SIR_i &= SIR_m - \pi\sigma SIR_m(1 - \frac{1}{SWR_m})(\varepsilon - 1)H_u \end{aligned} \tag{11}$$

The Gini coefficient in city i is given by

$$Gini_i = \frac{SWR_i * SIR_i}{SWR_i * SIR_i + 1} - \frac{SIR_i}{1 + SIR_i} = \frac{1}{1 + SIR_i} - \frac{1}{1 + \frac{\alpha}{\beta} SIR_i^{1-\frac{1}{\sigma}}} \tag{12}$$

Substituting equation 11 into Euqation 12 yields

$$\begin{aligned} Gini_i = f(\pi) &= \frac{1}{1 + SIR_m - \pi\sigma SIR_m(1 - \frac{1}{SWR_m})(\varepsilon - 1)H_u} \\ &- \frac{1}{1 + \frac{\alpha}{\beta}(SIR_m - \pi\sigma SIR_m(1 - \frac{1}{SWR_m})(\varepsilon - 1)H_u)^{1-\frac{1}{\sigma}}} \end{aligned} \tag{13}$$

[5] This is discussed in Appendix B.

[6] The IEH mechanism assumes that along with an increase in housing cost, the increase in local wage keeps the workers in interregional equilibrium.

As shown in equations 12 and 13, the Gini coefficient in city i is a nonlinear function of the local SIR, SIR_i, and is therefore determined by the differential in local housing costs, represented by π. Because SIR_i is monotonically increasing with housing costs, the Gini coefficient in city i is also a nonlinear function of housing cost changes(π). The concave functional relation between Gini and housing cost is shown graphically in Section 4.

4 Numerical Solution for the Effect of Housing Costs on Inequality

This section simulates equation 13 by assigning values to the parameters, and shows the results in a series of graphs. Parameters are calibrated on the basis of other empirical papers, as shown in Table 3.

Table 3: List of Parameter Values in Calibration

parameter	calibrated value	source
Hu	0.3	Glaeser et al.(2008)[13].
ε	0.3	Kim et al.(2009)[18], Broxterman and Yezer(2015)[7]. [1]
σ	1.6	Hendricks(2011)[17] [2]
SWR_m	2.18	National average SWR based on Census 2010
SIR_m	0.43	National average SIR based on Census 2010

[1] 0.3 is the value used in Kim et al.(2009)[18]. Broxterman and Yezer(2015)[7] summarized the estimates of income elasticity of demand for housing, with a wide range of 0.182-0.627. 0.3 is considered as the midpoint.

[2] The value of 1.6 is used in Hendricks(2011)[17], which is the midpoint of a range of estimates provided by Ciccone and Peri(2005)[9]. Estimating σ is a problem, and there are not many updated estimates in recent papers. Ciccone and Peri's(2005)[9] estimates are in the range of 1.3-2.2 based on micro-data of white males in 1950-1990 decennial census.

Assuming that cities share the same production function, the national average SIR and SWR are used to estimate the parameters of the production function in equation 3. Based on equation 6, given SWR_m and SIR_m, $\frac{\beta}{\alpha}$ is determined by:

$$\frac{\beta}{\alpha} = SWR_m^{-1} SIR_m^{-\frac{1}{\sigma}} = 0.78; \tag{14}$$

Note that the value of $\frac{\beta}{\alpha}$ depends on the choice of σ.

The values of σ, ε,Hu affect the peak of the Gini function, determined by equation 13. Numerical solution of equation 13 is accomplished using parameter values taken from the literature. The sensitivity of these solutions of equation 13 to these parameter values is tested by varying the values and resolving the model. This robustness exercise does not change the basic finding of concavity of the Gini function. Solutions are shown in figures and discussed below.

4.1 The "Kuznets Curve Relation" at the Country Level

The classic Kuznets Curve relates income per capita to the Gini coefficient. A supporting argument is that, the inequality in access to education is gradually eliminated when income per capita reaches higher levels. Morrisson and Murtin(2013), for instance, show empirically that human capital inequality within countries follows an "inverted U-shaped curve". This is also supported by the theoretical model in this paper.

The relation between national average inequality and national SIR is determined by equation 8. The relation is also solved numerically with the parameters given by the calibrations in Table 3, and shown graphically in Figure 1. When very few people in a country have a college education, the SIR is low, and consequently, the SWR is high. This is represented by the left end of the curve in Figure 1. As income per capita rises, SIR increases and SWR declines, so the economy moves from the left end to the right end of the figure along the curve. In this process, the Gini coefficient rises at first, and then falls.

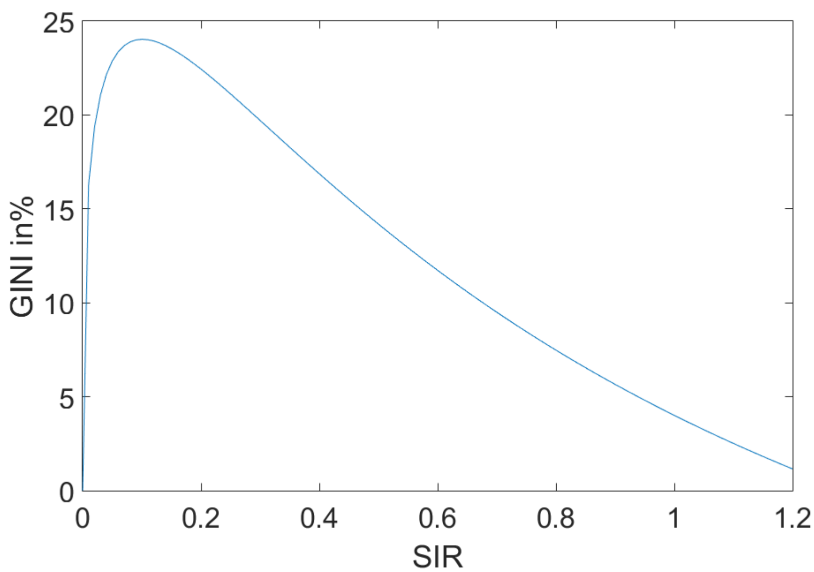


Figure 1: Change in National Average Gini Coefficient with Average SIR

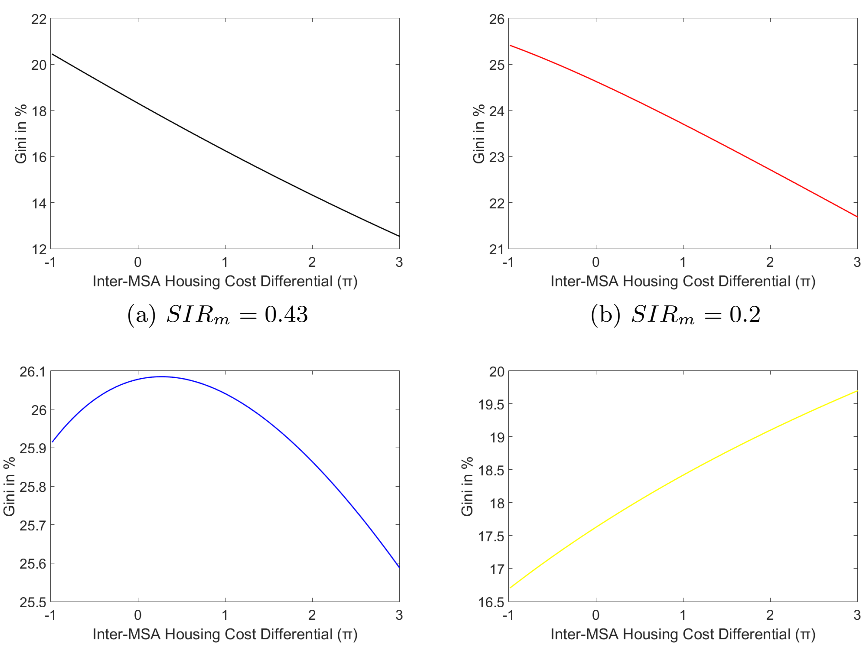
4.2 Concavity in the Gini - Housing Cost Relation: The "Kuznets Curve Relation" at the MSA Level

The expression of the Gini function in equation 13 determines the Gini - housing cost relation. The function is complicated, but the graph of the function shown in Figure 2a looks straightforward. With national average SIR=0.43($SIR_m = 0.43$, the black curve), Gini coefficient is declining with housing costs at the

MSA level at current values of the calibrated parameters. Note that this does not mean that inequality always falls with housing costs monotonically. Housing costs affect inequality in MSAs by affecting SIR. The negative relation stems from the fact that the current SIR is very high. Referring to Figure 1, Gini is falling with SIR in the limited range of variation in SIR generated by housing cost differentials. (The SIR - housing cost relation has been studied in the literature. In Appendix C, the variation in SIR w.r.t. housing costs is shown in Figure 3.)

Some parameters in equation 13 have changed significantly over time. An alternative but plausible set of parameters allows the Gini coefficient to exhibit a bell-shaped curve w.r.t. housing costs. Section 5 explores the way changes in the values of the parameters affect the results.

Figure 2: Change in City Gini Coefficient with Housing Costs, at Different Levels of SIR_m



(a) $SIR_m = 0.43$ (b) $SIR_m = 0.2$

(c) $SIR_m = 0.1$ (d) $SIR_m = 0.01$

5 Robustness Test

The results in Section 4 show that the form of the relation between housing price and the Gini of cities is sensitive to changes in the parameters, this section allows the parameters to vary in value, and examines the circumstances under which variation in housing cost is sufficient to produce a concave Kuznets Curve between housing cost and Gini.

Because SWR is a function of SIR (equation 6), the shape of the Gini curve as a function of SIR is solely determined by the elasticity of substitution between skilled and unskilled workers, σ. The local SIR is endogenously determined by housing costs. City level Gini is a function of local SIR, where local SIR is an increasing function of housing costs differential, π, according to equation 11. If other parameters are given by Table 3, the Gini is concave w.r.t. SIR. But the shape of the Gini curve as a function of π depends on whether Gini is monotone in the range of SIR generated by the housing cost changes. The following tests illustrate how these relations change with parameters. Results with respect to SIR and σ are shown because they have the most significant effects. Other parameters also show some limited effects, but do not change the fundamental of the curve. They are therefore not discussed in detail here.

5.1 The Effect of National SIR (SIR_m) on the Kuznets Curve

Without modifying the production function parameters, assume that the national average SIR is respectively 0.01, 0.1, 0.2. These 3 cases, together with the case based on $SIR_m = 0.43$ in 2010, represent the evolution of skill intensity in the U.S..

As shown in Figure 2d, when national average SIR is very low(at $SIR_m = 0.01$), the Gini coefficient rises with city housing costs. In 1960-1970, the national average SIR was close to 0.1. Figure 2c shows that Gini coefficient first rises with housing costs, then falls with housing costs with $SIR_m = 0.1$. As national average SIR continues to grow, as shown by Figure 2b($SIR_m = 0.2$) and Figure 2a($SIR_m = 0.43$), the Gini coefficient should be falling.

Inequality at MSA level is indeed a concave function of housing costs, though the economy may be on the upward sloping, bell-shaped, or downward sloping portion of the curve, depending on the parameters. As the national SIR increases over time, the relation between inequality and city size changes accordingly, most probably from positive correlation to non-monotonic(bell shape), and then to negative correlation over time. As national average SIR is already at a relatively high level, the model indicates that the Gini coefficient should already be falling as a function of housing cost or population.

At first glance, this contradicts with the reported positive correlation between inequality and city size. There are two possible reasons for the contradiction. Firstly, note that in the theoretical framework in this paper, the housing cost is proposed to be the primary source of concavity, and this force is now countervailing other forces that increase inequality with city size. If housing price was the only factor influencing the relation between city size and inequality, the inequality should have begun to fall with housing cost and size when national average SIR was about 0.1-0.2. As noted above, the existing literature has identified a number of other forces that promote the positive relation reported in previous literature. It is likely that these forces have delayed the arrival of the turning point.Secondly, the change in Gini with housing costs is based on the assumption of interregional indirect utility equilibrium. Because the interregional labor market does not instantly reach equilibrium, there may be a lag in the housing cost effect.

5.2 The Effect of σ on the Kuznets Curve

The parameter σ is the absolute value of the elasticity of substitution of skilled for unskilled worker in the production function. It determines the Gini maximizing level of SIR for a MSA according to equation 12, and consequently determines the Gini - housing cost relation according to equation 13.

The Gini-SIR relation is shown in Figure 4 in Appendix C. The peak of the Gini curve moves rightward as σ rises. When demand for unskilled workers is more elastic, cities more actively substitute skilled workers for unskilled workers as SWR falls, which allows inequality to rise with SIR.

The Gini - Housing Cost Relation at Different σ Values is shown in Figure 5 in Appendix C, with $SIR_m = 0.1$. Inequality falls with housing cost when σ is low(Figure 5a), and rises with housing costs when σ is high(Figure 5c and Figure 5d). At some intermediate level of σ(Figure 5b), inequality as a function of housing costs is a bell-shaped curve.

In the theoretical model, it is assumed that production function, including parameter σ is identical across cities. The sensitivity test w.r.t. σ can be explained in two aspects. On one hand, if all cities share the same σ, then the Gini coefficient function is significantly affected by the value of σ. On the other hand, if σ (and hence the production function) varies across cities, then a city should experience falling/rising inequality if its elasticity of substitution is low/high, holding SIR at an intermediate level.

5.3 Summary of Model Calibration

The nonlinear relation between inequality and housing costs is confirmed in cases that allow SIR_m and σ to vary. Combining results in Section 5.1 and Section 5.2, it follows that a lower value of SIR_m and/or a greater value of σ can cause the function of Gini w.r.t. housing costs to be bell-shaped. If national average SIR rises over time, or σ falls over time, finally inequality should fall with housing costs, which is the downward sloping side of the Gini function.

Regardless of the other forces that increase inequality with city size, SIR should continue to rise with housing costs, and Gini coefficient should eventually fall. When most of the workers are skilled, then housing cost should act as a dominating force that finally overcomes other forces, leading to falling inequality. Section 6 provides empirical support for the concavity of Gini coefficient as a function of housing costs.

6 Empirical Test: the Relation between Inequality and Housing costs

The theoretical part of the paper uses housing costs instead of population as the determinant of MSA level inequality. This section reexamines the empirical test by replacing population by housing cost measures. The theoretical model based on the IEH predicts a nonlinear relation between inequality and housing costs that resembles the bell shape of the country level Kuznets Curve.

Previous research has used population rather than housing cost as the determinant of earnings or income inequality at the city or MSA level. As argued in Section 3, the IEH implies that housing costs and income elasticity of housing expenditure have caused the employers to substitute skilled workers for unskilled workers. Housing cost is a distinctive feature of cities that, while correlated with population, is predicted to have a concave relation to the Gini coefficient of cities. Therefore, it is important to use the level of housing costs in MSAs as the explanatory factor in the regression.

6.1 Alternative Measures of Housing Costs

There are 3 major ways to measure housing costs in MSAs: median house value, median annual rent, and repeat sales indexes like the FHFA Housing Price Index. In the census micro-data, respondents report their house value and yearly rent. Many studies use the medians of house value and rent in a MSA as indicators of the housing cost level. The FHFA Housing Price Index(HPI) is a widely used repeat sale price index.

For cross-sectional studies, because the characteristics of the median unit are very different across cities, the median values are not ideal measures of differential housing cost . On the other hand, FHFA HPI can only track price changes in each city over time, and thus fail as an indicator of cross section housing cost differences.

Because all the measures above are problematic for cross-sectional comparisons, housing costs indices based on hedonic regressions are chosen as the primary measures of housing cost levels in this paper. Hedonic regressions account for variation in physical characteristics of the unit and, in some cases, local amenities in the neighborhood, they are better measures of variation in the cost of a "standard" housing unit across locations.

Carrillo et al.(2013)[8] generate a panel of gross rent indexes for MSAs from 1982 to 2012, based on the following regression model:

$$log(Rent) = X\beta + D_i + \varepsilon \tag{15}$$

where MSAs are indexed by i, and D_i is the MSA fixed effect, which also serves as the gross rent index for MSAs. By controlling for an exhaustive list of housing, neighborhood, and location characteristics, they generate the MSA gross rent index based on about 173,000 housing units throughout the United States for period 1982-2012. This index is the principal measure of housing costs in this paper. This will be referred to as the CEO index.

An alternative hedonic housing cost index created by Albouy(2008)[1] is only available for year 2000, but has the advantage of including owner-occupied housing units in the hedonic regression, rather than just rental units. (Albouy imputes the estimated rent based on house value when a unit is owned by the respondent.) The Albouy's Index is based on a hedonic regression that has a restricted variable list compared to the CEO index and is used primarily as a robustness check in the empirical tests.

6.2 Inequality and CEO Rent Index at the MSA Level

In Table 4, a test similar to that in Table 2 is conducted, but population is replaced with the CEO gross rent index.

Table 4: The Relation between Gini and MSA CEO Rent Index

	(1) Gini in %	(2) Gini in %
D1990 × CEO Rent Index	0.797***	3.657***
D1990 × CEO Rent Index squared		-0.520***
D2000 × CEO Rent Index	1.224***	3.337***
D2000 × CEO Rent Index squared		-0.237***
D2010 × CEO Rent Index	0.652***	1.387***
D2010 × CEO Rent Index squared		-0.0509***
Constant	12.27***	8.517***
R-squared	0.500	0.514
MSA FE	NO	NO
Year FE	YES	YES
Observation	791	791

Errors are clustered by year. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Summary Statistics of CEO Rent Index by Census Year

Year	Obs	Mean	Std. Dev.	Min	Max	index at the peak
1990	243	2.1655	.3659	1.7298	4.2589	3.5
2000	274	2.8198	.7318	1.9778	7.7137	7.0
2010	274	3.9268	1.7079	2.4133	17.1844	13.6

The linear Model in Column (1) suggests that 1 unit increase in the CEO index increases the Gini coefficient by 0.652% in 2010. When the square of CEO index is added to the regression, Model (2) suggests that the Gini coefficient is a concave function of CEO rent index, which inequality first rises with housing costs, and then falls when housing costs are high. Inequality peaks at CEO rent index of 3.5, 7.0, 13.6 respectively in year 1990, 2000, 2010. As shown in Table 5, the rent index increases over time. San Francisco (CEO index=13.4) has been around the peak of the curve, and Stamford, CT (CEO index=17.2) is on the downward-sloping side of the curve.

Because the mean of CEO index in 2010 is about 3.9, inequality is at its peak when housing costs in a city are about 3 times of the national mean of all MSAs, and it starts to fall if housing costs continue to rise. Most MSAs still experience rising inequality with gross rent. However, if the housing costs continue to rise in the larger cities, they should eventually reach the downward portion of the curve, where inequality falls with housing costs. The major point here is that the concavity predicted for the relation between inequality and housing cost by the extended version of the IEH developed here holds in this housing cost data.

The continuous Gini coefficient of wage and salary earnings for a MSA without assuming two skill groups of workers is used as an alternative Gini coefficient for a robustness check in Table 8 in Appendix D. The feature of concavity is still significant in 2000 and 2010. Inequality peaks at CEO rent index of 6.6 and 15.1 respectively in year 2000, 2010.

6.3 Inequality and Albouy Index

In this section, the housing cost index in Albouy(2008)[1] is used as an alternative measure of housing cost levels. As discussed in Section 6.1, the Albouy index includes both owner-occupied houses and rental units, while the CEO Index includes only rental units. The linear Model in Column (1) of Table 6 indicates that 1 unit increase in the Albouy index increases Gini coefficient by 4.3%.

Table 6: The Relation between Gini and Albouy Index of Housing Costs, 2000

	(1) Gini in %	(2) Gini in %
Albouy Index	4.344***	10.51***
Albouy Index Squared		-2.329**
R-squared	0.287	0.306
Observation	240	240

Robust standard error. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Variable	Obs	Mean	Std. Dev.	Min	Max	index at the peak
Albouy index	240	.8960	.2263	.5655	2.2479	2.25

When the square of Albouy index is added to the regression, results in Column (2) once again confirm that Gini coefficient is a concave function of housing costs. Inequality peaks at Albouy index of about 2.25 in 2000. San Francisco-Oakland-Vallejo, CA(at Albouy index=2.248) is around the peak of the curve. For most cities, inequality varies directly with housing cost.

The continuous Gini coefficient of wage and salary earnings for a MSA without assuming two skill groups of workers is used as an alternative Gini coefficient for robustness check in Table 9 in Appendix D. Gini still exhibit a bell shape, peaking at Albouy index of about 2.06 in 2000. San Francisco-Oakland-Vallejo, CA (at Albouy index=2.248) has been on the downward portion of the curve.

6.4 Summary of the Empirical Relation between Inequality and Housing costs

Using either the CEO rent index or the Albouy Index as the explanatory variable, the empirical tests all indicate that the increasing inequality with city size is not permanent. Eventually housing cost through the IEH effect should act as a force that eliminates the monotonic relation between Gini and either population or housing cost. As a result, if housing costs continue to rise in the large cities, inequality should fall, assuming other factors(production function, national average SIR, etc.) are constant.

7 Conclusion

Based on the Income Elasticity Hypothesis proposed by Kim et al.(2009), this paper derives the MSA Gini coefficient as a function of housing costs in Section 3. In Section 4, by assigning values to the parameters in the theoretical model, the simulation of the model shows that Gini coefficient is a bell-shaped or concave function of housing costs. Given the parameters in Section 4, the economy is predicted by model to be on the downward-sloping side of the concave function, which means that inequality should fall with housing costs.

In Section 5, some alternative values for key parameters(SIR_m and σ) are used. it is shown that at a lower value of SIR_m and/or a greater value of σ the full bell-shaped curve of the relation between the Gini of cities and housing costs is clearly observed. As the national average of SIR rises or σ falls the relation tends to become monotonic decreasing.

Given the substantial literature suggesting reasons that Gini varies directly with city population, there appear to be other forces causing inequality to rise with city size. In that case housing cost is a countervailing force that generates concavity in the long run. The finding of concavity is supported by empirical results using housing cost as an independent role in determining MSA Gini coefficient in Section 6.

This paper provides a theoretical framework to explain the rising inequality with city size. The model derived from IEH provides a theoretical basis for the argument that housing costs rather than city population causes spatial differentials in earnings inequality. More importantly, for the first time in the literature, this theoretical model also suggests that, if the housing costs in large cities continue to grow relative to the smaller ones, inequality will eventually begin to fall. As skilled wage ratio(SWR) falls and skill intensity ratio(SIR) rises with housing costs, they are the forces that should prevent inequality from unlimitedly going up. This theoretical conclusion is confirmed by empirical tests using CEO gross rent index and Albouy Index as the measure of housing costs.

References

[1] David Albouy. *Are Big Cities Bad Places to Live? Estimating Quality of Life across Metropolitan Areas*. National Bureau of Economic Research, Cambridge, Mass, 2008.

[2] Nathaniel Baum-Snow and Ronni Pavan. Inequality and city size. *Review of Economics and Statistics*, 95(5):1535–1548, 2013.

[3] Kristian Behrens and Frederic Robert-Nicoud. Survival of the fittest in cities: urbanisation and inequality. *The economic journal*, 124(581):1400; 1371–1400, 2014.

[4] Dan Black, Natalia Kolesnikova, and Lowell Taylor. Earnings functions when wages and prices vary by location. *Journal of Labor Economics*, 27(1):21–47, 2009.

[5] Kenyon Bolton and Sébastien Breau. Growing unequal? changes in the distribution of earnings across canadian cities. *Urban Studies*, 49(6):1377–1396, 2012.

[6] Jeffrey Brinkman. The supply and demand of skilled workers in cities and the role of industry composition. *IDEAS Working Paper Series from RePEc*, 2014. ID: proquest1698079009.

[7] Daniel A. Broxterman and Anthony M. Yezer. Why does skill intensity vary across cities? the role of housing cost. *Regional Science and Urban Economics*, 55:14–27, 2015.

[8] Paul E. Carrillo, Dirk W. Early, and Edgar O. Olsen. A panel of interarea price indices for all areas in the united states 1982-2012. *Journal of Housing Economics*, 26:81, 2014.

[9] A. Ciccone and G. Peri. Long-run substitutability between more and less educated workers: Evidence from us states, 1950-1990. *Review Of Economics And Statistics; Rev.Econ.Stat.*, 87(4):652–663, 2005. ID: wos000233871900007.

[10] Joel A. Elvery. City size and skill intensity, November 2010 2010. ID: 271675.

[11] Richard Florida, Charlotta Mellander, Högskolan i Jönköping, Centre of Excellence for Science IHH, Innovation Studies (CESIS), Nationalekonomi IHH, and Internationella Handelshögskolan. The geography of inequality: Difference and determinants of wage and income inequality across us metros. *Regional Studies*, 50(1):79–92, 2016.

[12] James K. Galbraith. Global inequality and global macroeconomics. *Journal of Policy Modeling*, 29(4):587–607, 2007.

[13] Edward L. Glaeser, Matthew E. Kahn, and Jordan Rappaport. Why do the poor live in cities? the role of public transportation. *Journal of Urban Economics*, 63(1):1–24, 2008.

[14] Edward L. Glaeser, Matt Resseger, and Kristina Tobio. Inequality in cities. *Journal of Regional Science*, 49(4):617–646, 2009.

[15] Joseph Gyourko, Christopher Mayer, and Todd Sinai. Superstar cities. *American Economic Journal: Economic Policy*, 5(4):167–199, 2013.

[16] Charles T. Haworth, James E. Long, and David W. Rasmussen. Income distribution, city size, and urban growth. *Urban Studies*, 15(1):1–7, 1978.

[17] Lutz Hendricks. The skill composition of u.s. cities.(report). *International Economic Review*, 52(1):1, 2011.

[18] Dongsoo Kim, Feng Liu, and Anthony Yezer. Do inter-city differences in intra-city wage differentials have any interesting implications? *Journal of Urban Economics*, 66(3):203; 209–209, 2009.

[19] Simon Kuznets. Economic growth and income inequality. *The American Economic Review*, 45(1):1–28, 1955.

[20] N. Lee, P. Sissons, and K. Jones. The geography of wage inequality in british cities. *Regional Studies*, 50(10):1714–1727, 2016.

[21] Amy Liu. The urgency to achieve an inclusive economy in the bay area. Technical report, Brookings, 2018.

[22] Janice Fanning Madden. *Changes in income inequality within U.S. metropolitan areas*. W.E. Upjohn Institute for Employment Research, Kalamazoo, Mich., 2000.

[23] Enrico Moretti. Real wage inequality. *American economic journal.Applied economics*, 5(1):103; 65–103, 2013.

[24] Christian Morrisson and Fabrice Murtin. The kuznets curve of human capital inequality: 1870–2010. *The Journal of Economic Inequality*, 11(3):283–301, 2013.

[25] Sherwin Rosen. Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1):34–55, 1974.

[26] Christopher H. Wheeler. On the distributional aspects of urban growth. *Journal of Urban Economics*, 55(2):371–397, 2004.

[27] Christopher H. Wheeler. Cities, skills, and inequality. *Growth and Change*, 36(3):329–353, 2005.

[28] Christopher H. Wheeler. Cities and the growth of wages among young workers: Evidence from the nlsy. *Journal of Urban Economics*, 60(2):162–184, 2006.

Appendix A Derivation of Gini Coefficient

For a simplified economy where there are only 2 levels of income, the rule of thumb formula for Gini coefficient is given by:

$$Gini = f - u \tag{16}$$

where f is "fraction of total income earned by the high income group" and u is "fraction of the high income population out of total population".

In the model, we have

$$\begin{aligned} f = w_s S/[w_s S + w_u U] &= SWR * SIR/[SWR * SIR + 1]; \\ u = S/[S+U] &= SIR/[SIR+1]; \\ \Rightarrow Gini = f - u &= \frac{SWR * SIR}{SWR * SIR + 1} - \frac{SIR}{1 + SIR} \end{aligned} \tag{17}$$

Appendix B Interregional Equilibrium and Zero Profit Condition of Employers

In the baseline model, we can see that a change in housing cost requires compensating changes in wages. In an interregional equilibrium, there should be compensations for the firms in high housing cost cities. In this paper, high housing costs are assumed to be the result of high productivity, A_i.[7] Firms keep entering the productive cities, driving up housing costs, until they earn zero economic profit. Note that housing costs not only enter the cost of land, but also the cost of workers implicitly for employers.

If housing costs increase by a proportion of π from the baseline city to city i, SIR and SWR changes according to equation 11. And the IEH suggests that

$$\begin{aligned} w_{u,i} &= (1 + \pi H_u) w_{u,m}; \\ w_{s,i} &= \theta_i w_{u,i}; \end{aligned} \tag{18}$$

The interregional equilibrium requires zero profit for firms in both cities:

$$\begin{aligned} \Pi_m^* &= A_m(\alpha s_m^{\frac{\sigma-1}{\sigma}} + \beta u_m^{\frac{\sigma-1}{\sigma}})^{\frac{\sigma}{\sigma-1}} - w_{s,m}s_m - w_{u,m}u_m = 0; \\ \Pi_i^* &= A_i(\alpha s_i^{\frac{\sigma-1}{\sigma}} + \beta u_i^{\frac{\sigma-1}{\sigma}})^{\frac{\sigma}{\sigma-1}} - w_{s,i}s_i - w_{u,i}u_i = 0. \end{aligned} \tag{19}$$

[7] One can also assume that prices are higher in big cities, which does not change the equations much.

where the input of skilled and unskilled workers (s_m, u_m, s_i, u_i) are functions of wages $(w_{s,m}, w_{u,m}, w_{s,i}, w_{u,i})$, determined by the FOCs of profit maximization. An increase in productivity from A_m to A_i is accompanied by an increase in housing costs by a factor of $1+\pi$, and consequently rise in wages from $w_{s,m}, w_{u,m}$ to $w_{s,i}, w_{u,i}$, according to equation 18. Because the production function and housing space input are not observable, equation 19 is just a theoretical hypothesis about the relation between housing costs and productivity, which is not to be tested in the empirical part of this paper. The empirical part focuses on the changes in SIR, SWR and Gini coefficient w.r.t. housing costs.

Appendix C Figures

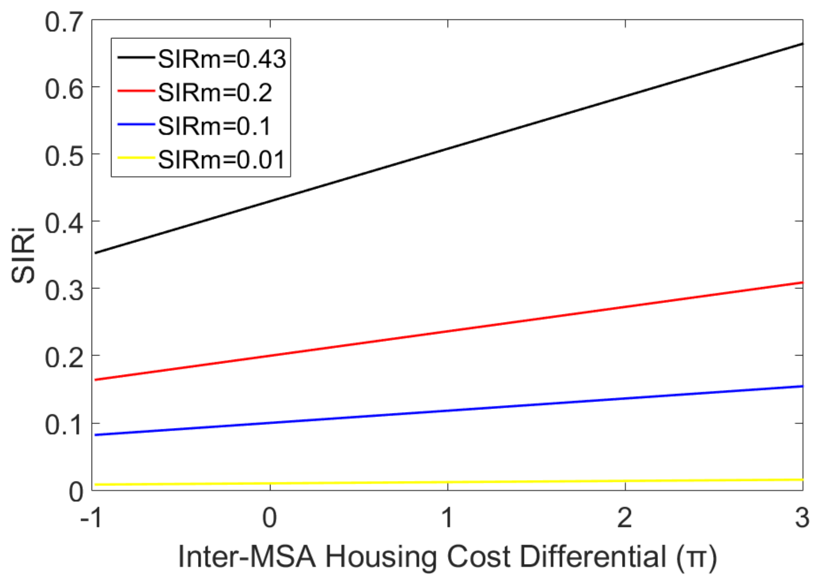


Figure 3: Change in CIty Level SIR with Housing costs

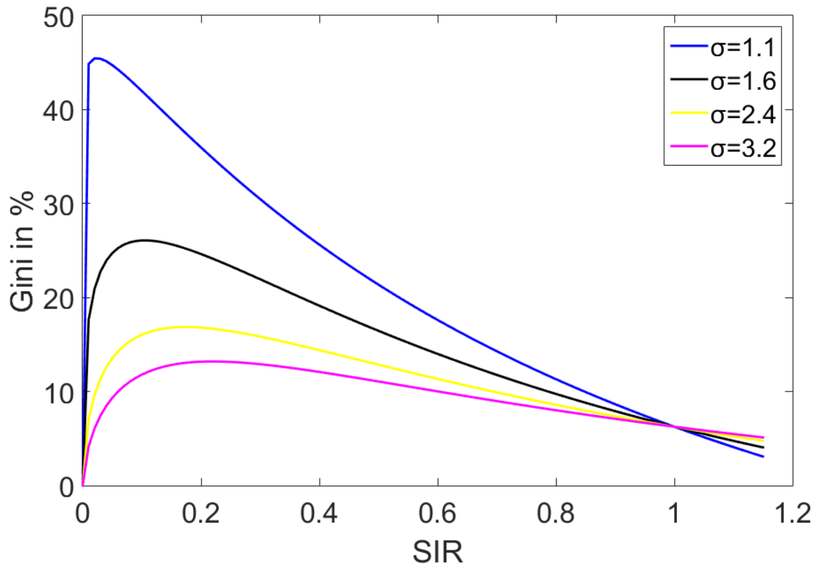
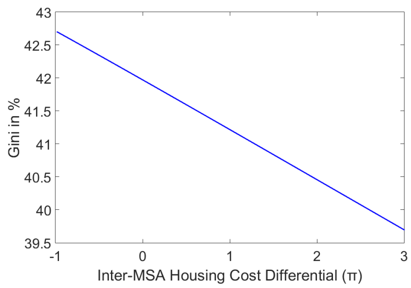
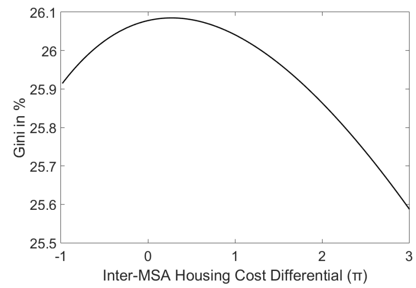


Figure 4: Change in City Level Gini Coefficient with City SIR, at Different σ Values

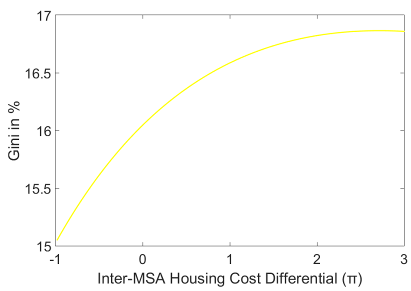
Figure 5: Change in City Gini Coefficient with Housing Costs, at Different Levels of $\sigma(SIR_m$ fixed at 0.1)



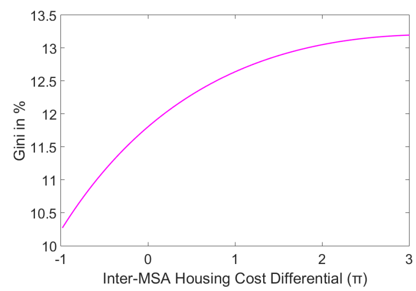
(a) $\sigma = 1.1$



(b) $\sigma = 1.6$



(c) $\sigma = 2.4$



(d) $\sigma = 3.2$

Appendix D Result Tables with Continuous Gini

Table 7: The Relation between Gini and City Population

	(1) Gini in %	(2) Gini in %
1980 × population in millions	-0.0173***	-0.0908***
1980 × population squared		0.00533***
1990 × population in millions	0.0444***	0.0603***
1990 × population squared		-0.00107***
2000 × population in millions	0.277***	0.305***
2000 × population squared		-0.00176***
2010 × population in millions	0.177***	0.186***
2010 × population squared		-0.000542***
Constant	44.95***	45.04***
R-squared	0.301	0.301
Year FE	YES	YES
Observation	1070	1070

Error is clustered by year. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: The Relation between Gini(alternative measure) and MSA CEO Rent Index

	(1) Gini coefficient	(2) Gini coefficient
1990 × CEO Rent Index(Hedonic)	-0.186***	-3.262***
1990 × CEO Rent Index squared		0.559***
2000 × CEO Rent Index(Hedonic)	0.960***	2.988***
2000 × CEO Rent Index squared		-0.227***
2010 × CEO Rent Index(Hedonic)	0.348***	0.665***
2010 × CEO Rent Index squared		-0.0220***
Constant	46.72***	50.76***
R-squared	0.199	0.208
Year FE	YES	YES
Observation	791	791

Error is clustered by year. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: The Relation between Gini(alternative measure) and Albouy Index of Housing Costs

	(1) Gini coefficient	(2) Gini coefficient
Albouy Index	3.375***	9.426***
Albouy Index Squared		-2.287**
R-squared	0.117	0.129
Year FE	NO	NO
Observation	240	240

Robust standard error. * p < 0.1, ** p < 0.05, *** p < 0.01.

Appendix E Question about Comment 6

According to equations 8 and 14:

$$\frac{\beta}{\alpha} = SWR_m^{-1} SIR_m^{-\frac{1}{\sigma}};$$
$$Gini_m = \frac{1}{1+SIR_m} - \frac{\beta}{\beta + \alpha SIR_m^{1-\frac{1}{\sigma}}}. \tag{20}$$

substitute for β/α in the Gini equation above yields

$$Gini_m = \frac{1}{1+SIR_m} - \frac{1}{1+SIR_m SWR_m}$$

According to the last equation, if we allow β/α to vary with σ, the resulting Gini coefficient will be invariant with respect to σ. Then the Gini coeffient is not affect by σ, which makes the tests unnecessary. So I have been using the same values of β/α in the sensitivity tests. So the sensitivity tests examines the effect of a change in σ on SWR, SIR and hence the Gini coefficient, assuming that the other production function parameters are constant.

My argument is that, if we want to find the partial derivative effect of a change in σ, we cannot allow β/α to change, as they are taken as parameters, and estimated based on actual data in 2010 census. similarly, we assume counter-factual changes in SIR_m in our sensitivity tests with respect to SIR_m, keeping all the production function parameters constant. Otherwise, when we change SIR_m to the 1970 level, we have to re-estimate β/α because both SIR_m and SWR_m change.

So I think we cannot reestimate β/α every we change another parameter. If we have to, then the sensitivity tests should be deleted because the Gini coefficient expression will change because β/α is now not constant parameters, but functions of SWR and SIR.