

Inequality and City Size*

Nathaniel Baum-Snow, Brown University & NBER

Ronni Pavan, University of Rochester

August, 2011

Abstract

Between 1979 and 2007 a strong positive monotonic relationship between wage inequality and city size has developed. This paper investigates the links between this emergent city size inequality premium and the contemporaneous nationwide increase in wage inequality. After controlling for the composition of the workforce across cities of different sizes, we show that at least 23 percent of the overall increase in the variance of log hourly wages in the United States from 1979 to 2007 is explained by the more rapid growth in the variance of log wages in larger locations relative to smaller locations. This influence occurred throughout the wage distribution and was most prevalent during the 1990s. More rapid growth in within skill group inequality in larger cities has been by far the most important force driving these city size specific patterns in the data. Differences in the industrial composition of cities of different sizes explain up to one-third of this city size effect. These results suggest an important role for agglomeration economies in generating changes in the wage structure during the study period.

*We gratefully acknowledge financial support for this research from National Science Foundation Award SES - 0720763. We thank Blaise Melly and Gregorio Caetano for helpful discussions. Cemal Arbatli and Ee Cheng Ong provided excellent research assistance. The paper has greatly benefited from discussions in seminars at the University of Chicago Booth School, UBC, CREI, and at the CURE and NARSC meetings.

1 Introduction

Juhn, Murphy & Pierce (1993), Card & DiNardo (2002), Lemieux (2006), and Autor, Katz & Kearney (2008) among others have documented a sharp rise in U.S. wage inequality since 1979, especially at the top end of the wage distribution. These studies discuss skill-biased technical change, capital-skill complementarity, the shifting composition of the workforce, the decline in the real value of the minimum wage and the decline of unionization as potential causes of this rise in wage inequality. It is less widely recognized that over this same time period, a strong positive relationship between wage inequality and city size has also developed. In the 2004 to 2007 period, the variance of log hourly wages was 0.28 in rural areas and roughly monotonically increasing to 0.53 in the largest three metropolitan areas. In contrast, in 1979 the variances of log hourly wages for rural areas and the three largest metropolitan areas were 0.19 and 0.24 respectively. Similar patterns are also seen in other commonly used measures of wage inequality.

In this paper, we investigate the mechanisms behind the emergence of the city size inequality premium from 1979 to 2007 and its relationship with the growth in overall wage inequality. After controlling for levels of and shifts in the composition of the workforce across cities of different sizes, we find that the overall increase in the variance of log wages in the United States would have been at least 23 percent smaller between 1979 to 2007, and 34 percent smaller between 1979 and 1999, had inequality in all cities grown at the same rate as in rural areas. We show that these influences of city size have occurred throughout the wage distribution.

Commensurate with Autor, Katz & Kearney's (2008) evidence using national data, we demonstrate that growth in within group inequality has been the most important force driving these city size specific patterns in the data. That is, most of the impact of city size on the increase in inequality nationwide derives from more rapidly rising within skill group log wage inequality in larger cities than in smaller cities. While this could reflect increased ability dispersion within observable groups in larger cities, we think it is more likely to reflect more rapid increases in the returns to unobserved skill in these locations. We also find that changes in the sorting of population subgroups with higher wage inequality toward larger cities has had no effect on overall inequality. For this reason, we suspect that increased sorting across locations on unobserved skill is also not an important explanation.

Up to one-third of the increase in the slope of the variance of wages with respect to city size is related to differences in one-digit industry composition across locations. Most of this industry effect comes from faster growth in the variance of wages within industry/skill groups that have always been disproportionately located in larger cities. This evidence is consistent with Autor, Katz & Krueger's (1998) evidence that skill upgrading, particularly in computer-intensive industries, has been an important mechanism behind the rise in wage inequality. It is also consistent with Bacold, Blum & Strange's (2009) evidence that while various measures of cognitive and noncognitive skills

are similar across cities of different sizes, the returns to certain soft and technical skills are higher in larger agglomerations.

Figure 1 documents recent trends in wage inequality for our sample of white men. It shows the change in log wages over each decade by percentile in the wage distribution in the base year. Figure 1 shows the large increases in dispersion experienced throughout the wage distribution during the 1980s and since 1999. During the 1990s, those in the top quarter of the wage distribution experienced wage increases and increasing inequality, with relative stability throughout most of the rest of the distribution. One lesson from the 1990s data is that it is impossible to fully understand trends in inequality without examining the upper and lower portions of the wage distribution separately.

The relationship between inequality and city size at recent points in time exhibits many of the same features as the nationwide evolution of inequality since 1979. In particular, even though larger cities exhibit greater log wage premia for higher skilled workers, they also have greater factor ratios of skilled to unskilled workers.¹ In the time series context, Katz & Murphy (1992) and Bound & Johnson (1992) argue that this positive correlation between relative prices and relative quantities is likely to be driven by increasing demand for skill precipitated by skill-biased technical change. Their empirical observations have led to an extensive theoretical literature, including Acemoglu (1998) and Galor & Moav (2000), that attempts to better rationalize the sources of skill-biased labor demand shifts. Krusell et al. (2001) propose that this pattern is attributable to a combination of capital-skill complementarity and declining capital rental rates relative to input costs of other factors, a hypothesis echoed in part by Autor, Katz & Kearney (2008). These analyses of time-series data are useful for understanding the patterns of prices and quantities of worker skill across locations in recent cross-sections as well. In particular, the cross-sectional patterns we document point to the existence of greater gaps in labor demand for skilled relative to unskilled workers in larger cities, where skilled workers are more abundant.

Although we can only partially identify mechanisms behind the emergent link between city size and wage inequality, it is clear that changes in the role of agglomeration economies are crucial to understanding this link, and consequently the recent increases in wage inequality nationwide. There are several factors that may have interacted with agglomeration economies to produce the emergent city size inequality premium. The economy may have experienced an increase in the importance of skill biased agglomeration forces. Alternatively, capital-skill complementarity joint with factor neutral agglomeration economies may have been a key ingredient. While labor is more expensive in larger locations (Baum-Snow & Pavan, 2011), the price of capital equipment has potentially become more uniform across different locations over time. Hence, cheaper capital equipment in larger locations would have increased the productivity of skilled workers as a result if the aggregate city level production function exhibits capital-skill complementarities.

¹Figure A1 presents data on skill premia and the skill composition of the workforce over time by city size.

Of course mechanisms that incorporate agglomeration economies are not likely the only drivers of the city size inequality premium. As our results suggest, it is also important to understand why the classes of workers that experienced greater increases in inequality are disproportionately located in larger cities. For this reason, we endeavor to accurately control for the skill composition of the workforce. After controlling for skill, however, we are left with a set of potential explanations for the city size inequality premium that all involve agglomeration economies one way or another. Either agglomeration economies are the reason that firms in industries that experienced more rapid growth in wage inequality have disproportionately located in larger locations or changes in the nature of agglomeration economies have directly caused changes in the wage structure. Whatever the exact mechanisms, it is clear that changes in the structure of labor demand are important for understanding the emergent role of city size in generating increases in wage inequality. Therefore, we hope that this study sparks further research investigating the mechanisms through which greater relative labor demand shifts for skilled workers have occurred in larger cities.

This paper proceeds as follows. Section 2 discusses in more detail the evolution of the city size inequality premium since 1979 and shows how we construct the data. Section 3 describes our empirical methodology. Section 4 investigates the roles of city size in generating growth in several measures of wage inequality. Section 5 characterizes the importance of accounting for industry composition. Finally, Section 6 concludes.

2 Wage Dispersion and City Size

2.1 Patterns in the Data

Table 1 presents a set of facts about the evolution of various measures of log hourly wage inequality over time. It shows that in each decade since 1979 the variance and 90-50 percentile gap of log wages have increased. In particular, the variance of log wages increased by 0.18 or 86 percent and the 90-50 gap increased by 0.28 or 53 percent between 1979 and 2007. The 50-10 percentile gap increased in every decade except the 1990s, with a total increase over the study period of 0.09 or 14 percent. While the 1970 census does not contain the requisite data to calculate hourly wages, the variance of weekly wages increased only slightly during the 1970s.

Columns 4 and 5 of Table 1 present a decomposition of the total variance in log wages into observed and residual components. The "Between" component is based on means of 825 age, education and city size cells². The "Residual" component is based on within-cell residuals from these means. We see that while the between variance increased at a faster rate between 1979 and 2007, the residual component of the variance increased by more in numerical terms. Columns 6 and 7 show the 90-50 and 50-10 percentile gaps in residuals. Both of these components increased

²We use 15 two-year age groups, 5 educational categories and 11 location types.

during the study period, with the 90-50 gap increasing much more quickly than the 50-10 gap.³ Changes over indicated time periods in the bottom block of Table 1 represent benchmarks against which we will compare counterfactual changes absent city size effects in Section 4 below.

Figure 2 demonstrates that a positive relationship between log wage inequality and city size emerged over the full distributions of log wages and city size concurrently with the growth in overall log wage inequality. For the purpose of this paper, we index metropolitan area size to be 0 in rural areas and 1 to 10 to represent deciles of the urban population distribution in year 2000. That is, in 2000 approximately 10 percent of the metropolitan area population nationwide resided in each of our city size categories. For other years, we maintain the same assignment of metropolitan areas to categories based on populations in 2000. We experimented with other similar indexes of metropolitan area size, including using contemporaneous deciles and/or fixed cutoff populations over time, and they all generate very similar results. We prefer our measure because it eliminates the possibility that changes in the relationship between city size and inequality could have been generated by a few metropolitan areas that changed locations in the city size distribution. In addition, our measure provides a clear way to assign metropolitan areas to city size categories in the 2004-2007 period for which there is no reliable MSA level population data. Incidentally, our measure also generates the steepest relationship between city size and inequality in 1979 of the four measures that we examined.^{4 5}

Figure 2 Panel A shows that while the variance of log hourly wages was almost flat as a function of city size in 1979, its slope increased in each subsequent decade. The variance of log hourly wages differed by 0.05 between rural areas and the largest metropolitan areas in 1979 whereas by 2004-2007 this gap had increased to 0.25. The variance of log weekly wages in 1969 was even flatter in city size than the variance of log hourly wages in 1979. Figure 2 Panels B and C show the evolution of log hourly wage distribution percentile gaps by city size over time. Panel C shows that the increase in the 50-10 percentile gap during the 1980s occurred approximately symmetrically across locations. The 50-10 gap experienced its greatest increase in slope with respect to city size during the 1990s, even though the average level changed very little during this period. Both slope and level increased after 1999. In contrast, Panel B shows that the level and slope with respect to city size of the 90-50 percentile gap increased in every period studied after 1979. These patterns are

³While the total variance can be decomposed into "Between" and "Residual" components, there is no natural decomposition possible for other measures of distribution spread.

⁴One potential concern with using population deciles rather than population levels is that if larger cities grow over time, our current strategy would misrepresent a stable relationship between inequality and city population as an increase in the city size inequality premium. However, a plot similar to Figure 2 Panel A that uses a flexible polynomial specification in city population instead of size deciles indicates a nearly identical pattern.

⁵Others including Ciccone & Hall (1996) use density rather than metropolitan area population as a way of capturing the extent of agglomeration forces. Depending on the importance of local transportation and communication costs, each measure can be justified by standard urban theory. We find population to be a more natural empirical measure as it does not require data on developed area. The correlation between year 2000 MSA population and population density in our data set is 0.44.

consistent with the evolution of the overall 90-50 gaps seen in Figure 1.

2.2 Data Construction

Our primary data source for demographic information and wages is the Census Public Use Microdata 5 Percent Samples from 1980, 1990 and 2000 plus the 2005-2007 American Community Surveys (ACS).⁶ We choose these data sets so as to achieve large enough samples within metropolitan areas in order to precisely estimate and decompose wage distributions by metropolitan area size categories. We limit our analysis to white men ages 25-54 who report working at least 40 weeks, 35 usual hours per week and who earn at least 75 percent of the federal minimum wage in each year. The full-time full-year limitation allows us to measure marginal products of labor for individuals who are less likely to be constrained in their residential locations by family or education considerations. We use white men only to limit the possibility that changes in discrimination and patterns of labor market attachment for women and non-whites influence our estimates. Our earnings measure is the log hourly wage calculated by subtracting log weeks times usual hours worked from log annual income.⁷ Annual income from the census is for the previous calendar year while that from the ACS is for the year ending in the (unobserved) survey month. Therefore, we sometimes report ACS wages as being for the period 2004-2007. To maintain comparability with the census data, we shift the wage distribution in each of the ACS sample years to have the same median as that for the 2006 sample. Many other studies that examine trends in inequality in the United States use the Current Population Surveys instead. We found that the CPS does not provide sufficient sample sizes and geographic detail to be the optimal data set for our purposes.

We consistently use year 1999 definition county based metropolitan area (MSA) geography throughout the analysis. Unfortunately, the most disaggregated census micro data geography of County Groups in 1980 and Public Use Microdata Areas in 1990 and 2000 in many cases does not match up exactly to MSA geography. As such, our spatial allocation of individuals reported as living in regions that straddle MSA boundaries is imperfect. We allocate those living in straddling county groups or PUMAs to the subregion with the largest population.

3 Measuring the Role of City Size

In this section, we develop a methodology for evaluating the effect of city size independent of observed skill on changes in various measures of log wage inequality. We begin with the nonparametric

⁶We do not use the 2008 ACS data because it only reports intervalled weeks worked.

⁷Measurement error is an additional justification for using full-time full-year workers. Baum-Snow and Neal (2009) demonstrate that there exists significant measurement error in hourly wages for part-time and part-year workers in the census. An additional potential measurement error concern explored by Lemieux (2006) involves changes in the propensity of workers to be paid by the hour. In Appendix A we show why this phenomenon has a minor impact on our analysis.

statistical decomposition of quantity and price components of changes in the log wage distribution proposed by DiNardo, Fortin & Lemieux (1996) and adopted by Autor, Katz & Kearney (2008) for analysis of U.S. data and Dustmann, Ludsteck & Schoenberg (2009) for analysis of German data.⁸ We utilize a framework that combines a quantity re-weighting procedure similar to that used in the existing literature with a version of the Changes-In-Changes (CIC) model introduced by Athey and Imbens (2006). The counterfactual distributions that we calculate using our framework allow us to study how different inequality measures would have looked had the city size-inequality gradient not emerged, while allowing demographic composition and skill prices to evolve over time as they did in equilibrium.

We construct our counterfactuals by adjusting log wage inequality for each demographic group to mimic the relationship relative to rural locations for each city size category that existed in 1979. This benchmarking to rural locations is a natural choice. As seen in Figure 2, rural log wage inequality consistently increased the least of all location types since 1979. Additionally, the industrial structure of the economy has changed the least in rural areas.⁹ Therefore, our counterfactual distributions capture how inequality would have evolved had larger cities' technological gaps with rural areas not expanded beyond their 1979 levels.

3.1 Fundamentals

We begin with the standard log wage decomposition that has been used extensively in other research, including Chay & Lee (2000) and Lemieux (2006). We denote the log wage of individual i of observed skill group x residing in location type s at time t as $w_{it}(x, s)$. We decompose each individual's observed log wage into mean and residual components:

$$w_{it}(x, s) = m_t(x, s) + \epsilon_{it}(x, s) \tag{1}$$

We calculate log wage residuals as observed log wages minus conditional means at each point in time. By construction, residual log wages $\epsilon_{it}(x, s)$ have mean zero.

Let $h_t(x, s)$ denote the joint distribution of observed skill and location. We extensively use the

⁸Machado and Mata (2005) and Melly (2005) are examples of papers that use a similar approach.

⁹Inspection of the modified Herfindahl index $\frac{1}{2} \sum_{k=1}^K (S_{kt} - S_{ks})^2$ by location confirms that rural areas have experienced the most stable industrial composition since 1979. In this expression, S_{kt} denotes the share of employment in industry k at time t . This index can be thought of as a monotonic transformation of the distance in K dimensional space between industry compositions at times t and s . It is straightforward to show that the index is bounded between 0 and 1. Using this index, we find that rural areas experienced the smallest change in industry composition between 1979 and 2007 and during each of the intervening study periods of any location size category. This pattern holds for all education groups as well except for high school dropouts in the 1990s and more than college in the 1980s. In addition, this pattern looks very similar when the index is instead calculated using three-digit industries. These results are reported in Table A1.

following decomposition for our analysis:

$$h_t(x, s) = h_{at}(s|x)h_{bt}(x)$$

The object $h_{at}(s|x)$ describes the probability of locating in a city of size s for an individual in observed skill group x , while $h_{bt}(x)$ describes the unconditional probability density of observed skills. This ordering is the only one that allows for determining the effect of city size on the quantity component of wage inequality since the distribution of observed skills $h_{bt}(x)$ can vary freely over time while we adjust the distributions of these groups across different locations through manipulation of $h_{at}(s|x)$.

The residual log wage distribution $f_t(\epsilon)$ can be expressed as follows:

$$f_t(\epsilon) = \int g_t(\epsilon|x, s)h_{at}(s|x)h_{bt}(x)dsdx \quad (2)$$

In this expression, $g_t(\cdot|x, s)$ is the residual log wage distribution for observed skill group x in location s at time t . To recover the unconditional distribution of residuals in year t , we must integrate these conditional distributions over x and s incorporating weights that represent quantities of workers in various demographic and city size categories. As we explain in detail in the following sub-section, the counterfactual residual distributions we examine are constructed by replacing the components $g_t(\epsilon|x, s)$ and $h_{at}(s|x)$ with analogs that reproduce 1979 relationships with respect to city size.

Following Lemieux (2006) and others, we treat each residual as the product of a price $\rho_t(x, s)$ and quantity $u(x, s)$ of unobserved skill. Assuming that unobserved skill quantity distributions within skill group and city size do not change over time, changes in the distributions $g_t(\epsilon|x, s)$ reflect changes in the price distributions of unobserved skills. This assumption is consistent with results discussed below showing that changes in observed skill quantities across locations had negligible effects on the structure of wages. Although this interpretation of the nature of residuals is useful, we stress that it is a convenient approximation that does not change the main contribution of this paper, which is to document the role of city size (through whatever mechanism) in generating increased overall log wage inequality.

The log wage distribution can be expressed analogously as follows:

$$a_t(w) = \int g_t(w - m_t(x, s)|x, s)h_{at}(s|x)h_{bt}(x)dsdx. \quad (3)$$

We treat means $m_t(x, s)$ as observed skill prices and distributions $h_{at}(s|x)$ as observed skill quantities in different locations.

3.2 Counterfactual Residual Log Wage Distributions

We now show how we construct the residual log wage distributions that we use to calculate growth in residual log wage inequality absent city size effects under three counterfactual environments. The first counterfactual experiment examines the role of quantity shifts across locations, the second experiment additionally examines changes in the relative prices of unobserved skill across locations over time and within observed skill groups while the third experiment examines the importance of accounting for demographic composition.

3.2.1 Quantities Only

In the first counterfactual experiment, we construct log residual wage distributions $f_t^q(\epsilon)$ that do not incorporate changes in the distribution of observed skills across location since 1979 but allow the relative quantities of observed skills in the economy as a whole to change as they did in equilibrium. Any reduction in inequality in these counterfactual distributions relative to actual residual distributions indicates that changes in the sorting of observed groups x across different locations s since 1979 has contributed to the increase in residual log wage inequality in the U.S. during this period. For each year t in Equation (2) above, we replace the distribution $h_{at}(s|x)$ with the distribution $h_{a1979}(s|x)$, resulting in the following overall residual distribution:

$$f_t^q(\epsilon) = \int g_t(\epsilon|x, s)h_{a1979}(s|x)h_{bt}(x)dsdx \quad (4)$$

Notice that the distribution of observed skills in the economy as a whole $h_{bt}(x)$ changes freely over time. The only modification is to how these skills are distributed across different locations.

3.2.2 Prices and Quantities

The second set of counterfactual residual distributions additionally maintain the 1979 conditional distributions of prices across city sizes within skill groups in later years while allowing the “marginal” distribution of prices (unconditional on location) to change over time and between skill groups. Our approach for obtaining these counterfactual distribution is a direct application of the CIC model of Athey and Imbens (2006). We adopt this approach because it is robust to monotonic transformations of wages and makes sense even if the distributions of unobserved skill are different across locations. Consistent with our anchoring discussion above, we allow the distributions of residual log wage inequality in rural area to vary freely over time. For each percentile in the residual distribution of each skill group in each location in 1979, we determine the percentile with the same residual in the rural distribution. The counterfactual price distributions for each skill group maintain this relationship in later years. In the language of the CIC model, rural areas are the control group and each other city size category is a different treatment group.

Consider the residual log wage ϵ for an individual in group x that resides in location s in 1979. This residual corresponds to the quantile $\phi = G_{1979}(\epsilon|x, s)$, where the function G_t is the cumulative distribution function of residuals. This same residual ϵ in the same group x corresponds to quantile $\phi' = G_{1979}(\epsilon|x, 0)$ in rural areas in 1979. The counterfactual residual ϵ_t^p associated with quantile ϕ at time t can then be written as:

$$\begin{aligned}\epsilon_t^p(\phi|x, s) &= \epsilon + G_t^{-1}(\phi'|x, 0) - G_{1979}^{-1}(\phi'|x, 0) = G_t^{-1}(\phi'|x, 0) \\ &= G_t^{-1}(G_{1979}(G_{1979}^{-1}(\phi|x, s)|x, 0)|x, 0)\end{aligned}$$

That is, this procedure reproduces the same change since 1979 in the residual log wage that was experienced by the rural worker who had the 1979 residual log wage associated with quantile ϕ' .

While 1979 residual distributions within skill groups are more similar across locations than in other years, they are slightly more dispersed in larger locations. This likely represents a combination of greater dispersion in the quantity of unobserved skill and higher returns to unobserved skill in larger locations. The counterfactual residuals apply the same rates of increase in the prices of unobserved skill experienced in rural locations to other locations in the same skill group as well. To see this, note that the 1979 mapping implies that:

$$\epsilon_{1979}(\phi|x, s) = \rho_{1979}(x, s)u(\phi|x, s) = \rho_{1979}(x, 0)u(\phi'|x, 0)$$

Substituting for $u(\phi'|x, 0)$ into the expression for the counterfactual residual in year t using the above equation achieves:

$$\epsilon_t^p(\phi|x, s) = \rho_{1979}(x, s) \frac{\rho_t(x, 0)}{\rho_{1979}(x, 0)} u(\phi|x, s)$$

Once we have the set of counterfactual residuals, inverting the function and differentiating produces the price component of the counterfactual residual distribution $g_t^p(\epsilon|x, s)$.^{10 11}

Putting the price and quantity components together, we construct counterfactual distributions of log wage residuals, fully decontaminated of changes in city size effects since 1979:

$$f_t^c(\epsilon) = \int g_t^p(\epsilon|x, s) h_{a1979}(s|x) h_{bt}(x) ds dx \quad (5)$$

¹⁰We demean all counterfactual residual distributions to allow them to capture residual components only. This demeaning has no effect on the results.

¹¹An alternative environment with many of the same features is what Athey and Imbens (2006) call the "quantile difference in differences" (QDID) model. Using this model, counterfactual residuals are generated by maintaining the residual gaps at each percentile between location s and location 0 that existed in 1979 for later years as well. In this environment, $\epsilon_t^p(\phi|x, s) = \rho_{1979}(x, s)u(\phi|x, s) + [\rho_t(x, 0) - \rho_{1979}(x, 0)]u(\phi|x, 0)$ which makes the most sense if the underlying distribution of unobservables is the same in locations 0 and s , a requirement that is not needed in the CIC model. Nevertheless, results using the QDID model are almost identical to those presented in the next section using the CIC model.

This equation represents the counterfactual distributions we use to measure the full effects of city size on residual log wage dispersion.

3.2.3 Excluding Observed Skill Composition

Our third experiment allows us to examine the extent to which more rapid growth in residual wage inequality in groups disproportionately located in larger locations contributes to the emerging city size inequality premium. These counterfactual distributions use residuals ϵ_t from Equation (1), but reflect prices and quantities adjusted across city size categories only, such that all observations are allocated to the same unified demographic group. By comparing inequality measures obtained using these benchmark counterfactuals with those obtained from the counterfactuals in the second experiment, we see how important accounting for observed skills across locations is for generating the city size effects on residual log wage inequality that we calculate.

Similarly to the second experiment, we compute the counterfactual residuals at each quantile ϕ in distributions indexed by location and time as:

$$\epsilon_t^n(\phi|s) = G_t^{-1}(G_{1979}(G_{1979}^{-1}(\phi|s)|0)|0)$$

Using these values, we recover the resulting counterfactual residual distributions $g_t^n(\epsilon|s)$. Denoting the unconditional distribution of city sizes in 1979 as $h_{1979}(s) = \int h_{1979}(x, s)dx$, we form the following counterfactual distributions that take out city size effects but do not account for differences in demographic composition across locations:

$$f_t^n(\epsilon) = \int g_t^n(\epsilon|s)h_{1979}(s)ds \tag{6}$$

Comparison of measures of inequality in $f_t^n(\epsilon)$ with those in $f_t^c(\epsilon)$ reveals the extent to which the prevalence in larger cities of demographic groups with greater increases in within group wage inequality has led to increases in overall residual inequality.

3.3 Counterfactual Log Wage Distributions

We turn now to the construction of log wage distributions that remove city size effects. This discussion follows closely that from the previous sub-section. The one difference is the addition of counterfactual observed skill distributions, or group means. Comparison with the residual counterfactual results allows us to determine the relative importance of prices of observed versus unobserved skills for generating the city size effect. This set of counterfactuals is analogous to those for the residuals with one additional experiment which separately evaluates the importance of increases in the gradient of the prices of unobserved skill only with respect to city size for understanding city size's influence on changes in the overall log wage structure.

3.3.1 Quantities Only

First, as with the residual distributions above, we adjust quantities to reflect their distributions across space as of 1979, resulting in the following counterfactual distributions:

$$a_t^q(w) = \int g_t(w - m_t(x, s) | x, s) h_{a1979}(s|x) h_{bt}(x) ds dx \quad (7)$$

3.3.2 Quantities and Prices of Unobserved Skills

Second, to evaluate the potential importance of changes in the prices of unobserved skill only as functions of city size for generating overall increases in inequality, we construct one set of counterfactual distributions that only adjust residual prices and quantities. These distributions can be written as follows:

$$a_t^x(w) = \int g_t^p(w - m_t(x, s) | x, s) h_{a1979}(s|x) h_{bt}(x) ds dx \quad (8)$$

where $g_t^p(\cdot | x, s)$ is the same as in Equation (5).

3.3.3 Quantities and Prices

For our third experiment, full counterfactual price components of wage distributions must additionally adjust the mean component of wages. We specify these counterfactual means to allow the mean log wage within skill group x , $m_t(x) = \int m_t(x, s) h_{at}(s|x) ds$, to vary freely over time while rescaling mean log wages across locations within skill group to resemble 1979 profiles. The following rescaling imposes adjustments to gaps between overall group means and city size-specific group means to remain the same as they were in 1979:

$$m_t^c(x, s) = m_t(x) + (m_{1979}(x, s) - m_{t,1979}(x))$$

where $m_{t,1979}(x) = \int m_{1979}(x, s) h_{at}(s|x) ds$. The assumption underlying this formulation is that absent changes in the city size effect, mean log wages would have evolved the same except that the slope with respect to city size within demographic group would not have changed after 1979. This is a conservative specification of counterfactual means because overall group means may also have been affected by the changing impact of city size. Appendix B.1 specifies an alternative specification of counterfactual means that allows mean log wages in rural locations to vary freely but retains the gaps between means in other locations and rural means within skill group at 1979 levels in all subsequent years.¹²

¹²Implementing the CIC model on the full wage distributions analogously to our treatment of the residuals is very sensitive to changes in the right tails of rural distributions and imposes the unnecessarily strong assumption that observed skill price gaps between location s and location 0 within skill group rises at the same rate as the rise in the

The resulting counterfactual log wage distributions that fully adjust for both quantities and prices can then be written as:

$$a_t^c(w) = \int g_t^p(w - m_t^c(x, s) | x, s) h_{a1979}(s|x) h_{bt}(x) ds dx \quad (9)$$

3.3.4 Excluding Observed Skill Composition

Finally, we also construct counterfactual wage distributions that do not account for demographics. These are analogous to (6) and can be expressed as follows:

$$a_t^n(w) = \int g_t^n(w - m_t^c(s) | s) h_{1979}(s) ds \quad (10)$$

Note that these distributions utilize counterfactual means $m_t^c(s)$ that are built analogously to $m_t^c(x, s)$ except that they assign every individual to the same demographic cell.

4 Main Results

Tables 2 and 3 present results of each of the counterfactual experiments described in the previous section. For each experiment, we present percentage reductions in the growth of overall inequality reported in Table 1 under counterfactual scenarios described by the distributions detailed in the previous sub-section. We prefer this measure because it allows for direct comparison of effects on the variance and percentile gaps.

4.1 City Size Effects on Residual Log Wage Inequality

Table 2 presents results showing how city size has affected measures of residual log wage inequality through quantities only and prices plus quantities. It also shows the effects of city size absent controlling for demographic composition, indicating how important the propensity of groups with large increases in within group inequality to locate in larger cities has been for generating overall increases in residual log wage inequality. These counterfactual distributions are constructed using Equations (4), (5) and (6) respectively.

In Table 2 Column 1 we see that shifts in the composition of the workforce across city size categories had essentially no effect on residual wage inequality during any period between 1979 and 2007.¹³ For this reason, we suspect that changes in sorting on unobserved characteristics was also unimportant for understanding the role of city size in generating increases in residual wage log

price of unobserved skill in location 0.

¹³As in Lemieux (2006), we also investigated the role of overall quantity shifts after 1980, not just those across city sizes within demographic groups. These results are reported in Table A2 Column 1 and largely echo Lemieux's findings using the CPS data of an important role concentrated in the 1990s.

inequality. However, Panel A Column 2 shows that greater shifts in prices of unobserved skill in larger locations led to greater growth in residual log wage inequality in these locations. Had these shifts not occurred, the residual variance of wages would have grown 20 percent less quickly during the 1980s, 49 percent less quickly between 1979 and 1999 and 35 percent less quickly over our entire study period. Panels B and C show that while these influences of city size occurred throughout the wage distribution after 1989, they were more important in percentage terms for understanding trends in the 50-10 percentile gap than the 90-50 percentile gap during this time period. However, since the 50-10 residual gap was unchanged during the 1990s while the 90-50 gap grew, city size had similar estimated effects on these two measures in numerical terms. Between 1979 and 1989, city size's influence independent of observable demographics was more prevalent in the top part of the residual wage distribution.

Results in Table 2 Column 3 show that not accounting for demographic composition only increases the estimated effects of city size on the growth in residual variance by 23 percent, from a reduction of 35 percent to a reduction of 43 percent. This overall lack of importance of demographics is unbalanced in an interesting way. For the 90-50 percentile residual gaps, accounting for demographics is quite important. Not doing so increases the estimated effects of city size by more than 60 percent over the full sample period, from 30 percent to 50 percent. However, not accounting for demographics actually slightly reduces estimated influence of city size on 50-10 percentile residual gaps. These results are driven by the fact that increases in prices of unobserved skill for highly educated workers contributed markedly to increases in overall residual inequality while returns to unobserved skill for less educated workers changed much less, as seen in Table 1, joint with the fact that more educated workers are disproportionately located in larger locations.

4.2 City Size Effects on Total Log Wage Inequality

Table 3 presents results analogous to those in Table 2 but for total log hourly wages rather than just their residual component. To generate these results, we utilize counterfactual wage distributions in Equations (7), (8), (9) and (10). The counterfactual reductions in inequality reported in Table 3 are analogous to those from Table 2 with the addition of Column 2, which reports the effects of quantities and residual prices only, with no adjustment for city size's potential effects on group means.

Table 3 Column 1 shows that like for residuals, shifts in the composition of the workforce across cities of different sizes had virtually no effect on total wage inequality.¹⁴ However, as seen in Panel A Column 2, imposing only residual price adjustment reduces growth in the variance of wages by 10 percent during the 1980s, 31 percent between 1979 and 1999 and 21 percent between 1979 and 2007. These largest effects of changes in unobserved skill prices during the 1990s were primarily

¹⁴Table A2 Column 2 presents counterfactual reductions in inequality growth after imposing 1980 quantities of all demographic groups and size categories.

concentrated in the bottom part of the wage distribution, though larger influences occurred in the upper part of the wage distribution in the 1999-2007 period to generate roughly equal effects of city size throughout the wage distribution over our full sample period. As expected, results in Column 2 exhibit the same general patterns over time and across wage distribution percentiles as those for residual inequality seen in Table 2, but with smaller magnitudes.

Table 3 Column 3 presents the results that we wish to emphasize the most. It shows the full effects of city size taking into account both residual and mean components of prices in addition to quantities. These results indicate that an estimated 23 percent of the nationwide growth in the variance of wages between 1979 and 2007 can be attributed to city size independent of demographic composition. As a fraction of total growth, the effect of city size was greatest during the 1990s. This is seen by observing that between 1979 and 1999, city size accounted for 34 percent of the growth in the variance of wages relative to just 16 percent between 1979 and 1989. However, it should be noted that the variance of wages grew by 0.08 during the 1980s but only 0.05 during the 1990s. Results in Panels B and C show that city size increased the growth in the 50-10 percentile gap the most between 1979 and 1999, though over the full sample period the effects of city size were balanced throughout the wage distribution. Comparison of the results in Columns 2 and 3 reveals that the channel through which city size drove increases in wage inequality was primarily through changes in the prices of unobserved skill across cities rather than more rapid growth in the returns to observed skill in larger cities.¹⁵

Comparison of Columns 3 and 4 reveals that accounting for demographic composition is important for recovering city size's independent effect on log wage inequality. Failure to account for demographic composition mostly results in much larger reductions in counterfactual log wage inequality measures than those for residuals calculated from Table 2. This means that the skill groups with the largest increases in returns to observed skill were disproportionately located in larger cities while these groups experienced much smaller secular increases in returns to unobserved skill relative to other groups.

City size's effect on wage inequality during the 1980s was swamped by the large secular increases in the returns to observed and unobserved skill that have been documented in other research. During the 1990s we see city size's influence kicking in throughout the wage distribution. As with residual log wages, city size had a greater influence in percentage terms on growth at the bottom of the wage distribution but a similar influence on 50-10 and 90-50 gaps in numerical terms during this period. In contrast, between 1999 and 2007 city size had slightly negative estimated effects on the variance and 50-10 gaps but explains more than one-quarter of the increase in the 90-50 gap, an effect entirely driven by its influence on unobserved skill. The positive effects of city size on 50-10

¹⁵Table A2 Column 3 shows results for which we use a less conservative specification of counterfactual means. It shows city size effects on the growth in the variance of log wages over the full study period that are up 40 percent larger than those in Table 3, meaning that more rapid growth in prices of observed skill in larger cities could account for up to one-third of the city size effect on the growth in the variance of log wages.

gaps were thus concentrated between 1979 and 1999 while those on 90-50 gaps were concentrated between 1989 and 2007.

5 The Role of Industry

To better understand the mechanisms through which city size generated increases in wage inequality, this section examines the extent to which differing industrial composition of the workforce across locations accounts for the results in the previous section. Our examination proceeds analogously to that in the previous section with the addition of one-digit industries to the set of skill variables considered.¹⁶ Because some two and three-digit industries do not hire workers of all education levels in cities of some sizes, it would be impossible to disaggregate the set of industries examined much more than we do here. We find that up to one-third of the change in the impact of city size documented in the previous section is attributable to industry. Most of this effect comes through increases in residual log wage dispersion that occurred in industries disproportionately located in larger cities.

We view our estimates of the portion of the city size effect accounted for by industry composition primarily as informative about the mechanisms by which city size has caused increases in inequality. Firms producing tradeables that locate in larger cities do so despite higher input costs (Baum-Snow & Pavan, 2011). Additional productivity through agglomeration economies justifies these firm location choices. Therefore, city size itself likely has a role in guiding firms' location decisions and hence the location patterns of workers by industry.

In order to operationalize the addition of industries to our analysis, we are required to make some mild parametric assumptions and limit ourselves to using variance as a measure of inequality. Even the 5 percent census samples do not have sufficient sample sizes to allow for nonparametric analysis of wage distributions by age, education, industry and city size. As such, we decompose the variance of hourly wages into "Between" and "Residual" components based on the following regression equation:

$$\ln w_{idjst} = \alpha_{dst} + \beta_{djst} + \delta_{jst} + \varepsilon_{idjst} \quad (11)$$

In this equation, i indexes individual, d indexes age/education (skill) group, j indexes one-digit industry, s indexes location size and t indexes time.¹⁷ Note that the full nonparametric specification used for the analysis of skill groups and city size in the previous section nests the specification in Equation (11). Based on this empirical formulation, we calculate counterfactual variances absent city size effects using a similar method as in the previous section. In particular, for log wages and their residuals we calculate inequality measures of the counterfactual distributions

¹⁶In fact, we use a set of industry categories that additionally disaggregates non-durable and durable manufacturing, transportation from communications and public utilities, and professional services from other types of services.

¹⁷To attain sufficient variation within education and industry, we use five year age ranges for this analysis rather than the two year age groups used for the analysis in the previous section.

$f_t^q(\varepsilon)$, $f_t^c(\varepsilon)$, $a_t^q(w)$ and $a_t^c(w)$ given by Equations (4), (5), (7) and (9) respectively. In addition, we calculate counterfactual "between" variances, which is based only on the mean component $\alpha_{dst} + \beta_{djt} + \delta_{jst}$. The details of our counterfactual calculations, particularly the procedure we use to calculate counterfactual between variances, are reported in Appendix B.2.

Table 4 presents the fraction reduction in the growth of between, residual, and total variances of log wages under various counterfactual environments. In each panel of Table 4, we adjust for a different set of observable characteristics when constructing counterfactual variances. Results in Panel A are constructed using the underlying regression specification in (11). Because it includes a more saturated specification, this empirical model generates residual variances that are 0.01 smaller in all years than those reported in Table 1. Because the set of regressors is different, entries that apply to residuals in Table 4 cannot be compared to those in Table 2 while those applying to total log wages can be compared to results in Table 3.

The counterfactual exercises in Panel B only control for city size and demographic characteristics. For these exercises, the between component, $\alpha_{dst} + \beta_{djt} + \delta_{jst}$, is further decomposed by projecting the full set of group means onto this smaller set of observable characteristics, yielding a new set of group means and a second residual. We calculate separate counterfactual variances using the same logic as in Panel A for these two elements of the between component and recombine these two sets of counterfactual variances such that the same set of residuals apply to the reported between and residual components in Panel B as in Panel A.

Results in columns 1-3 of Panel B show that changing distributions of demographics across city sizes cannot account for increases in the variance of log wages. Columns 4-6 examine the role of quantities and prices together. Column 4 shows that city size had only a 9 percent impact on growth of the between component of the variance over the full sample period. City size accounts for 31 percent of the growth of residual variance, adding up to a total of 22 percent. These numbers are slightly smaller than those reported in Table 3 because the construction of counterfactual residual variances are more constrained and the age cells are larger.

Comparisons of Panels A and B reveal that one-fifth to one-third of the growth in total variance accounted for by city size independent of skill can be attributed to industry, depending on the time period. Because the residual component is the dominant source of the effects of city size on the variance of log wages in Panel B, industry's contribution primarily comes from the residual component of log wages. This means that industries with greater wage dispersion within skill groups were more concentrated in larger cities. Shifts in the distributions of workers across location within industry were not responsible for any change in the variance.

Table 5 summarizes our results for each of the time periods studied. The first row of each panel presents the fraction of the growth in the variance of each component of log wages due to all factors related to city size. These results are based on an analysis that is analogous to that used for Table 4 Panel B, but without any controls for demographics included in the regression

specification. Therefore, these results are very similar to those in Panel A of Table 3, Column 4. Based on results in Table 4, the subsequent two rows break down this effect of city size into components due to skill and industry compositions of the workforce. The final row in each panel, marked "Remainder", indicates the portion of the growth in the variance that is due to city size and that we cannot attribute to another factor correlated with city size. Therefore, agglomeration economies are likely centrally involved in generating these "Remainder" effects. These results also clearly indicate that understanding changes in wage dispersion within skill groups and industries disproportionately located in larger cities is key to explaining the growth in U.S. wage inequality since 1979.

6 Conclusions

Cities have played an important role in the rise of wage inequality over time. In 1979, there was only a weak positive relationship between inequality and city size while by 2007 a much stronger relationship between these two variables had developed. We demonstrate that city size specific factors can explain at least 23 percent of the overall increase of the variance in wages between 1979 and 2007 independent of observed skill. Although city size accounts for a greater amount of the increase in inequality in the top half of the wage distribution, in percentage terms its influence is balanced across the wage distribution between 1979 and 2007. The city size effect in the 1990s exceeds those in other periods studied.

An important factor generating the city size specific component of inequality growth is that demographic groups and industries disproportionately located in larger cities experienced larger increases in their wage dispersion in larger cities than in smaller cities. Even after controlling for these demographic differences, a sizeable fraction of this effect remains, and we believe that this residual effect is likely to be strongly connected with a change in the role of agglomeration economies. City size has become particularly more complementary with wage dispersion within observed skill groups.

We hope that our analysis sparks further research examining reasons for changes in the structure of labor demand using metropolitan area level data. In particular, while we provide evidence that agglomeration has interacted with technical change of some sort, whether skill-biased or skill-neutral, there remains much to be learned about the extent to which the increase in wage inequality that has been driven by this technical change has been augmented by capital-skill complementarity and potentially declining capital costs. As such, a ripe area for future research is to understand how increases in inequality attributable to movements along labor demand schedules have augmented those caused by technical change particularly oriented toward larger cities.

References

- Acemoglu, Daron. 1998. "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality." *Quarterly Journal of Economics*, 113:3, 1055-1089.
- Athey, Susan and Guido W. Imbens. 2006. "Identification and Inference in Nonlinear Difference-in-Differences Models." *Econometrica*, 74:2, 431-497.
- Autor, David H., Lawrence F. Katz and Melissa S. Kearney. 2008. "Trends in U.S. Wage Inequality: Revising the Revisionists." *Review of Economics and Statistics*, 90:2, 300-323.
- Autor, David H., Lawrence F. Katz and Alan B. Krueger. 1998. "Have Computers Changed the Labor Market?" *Quarterly Journal of Economics*, 113:3, 1169-1213.
- Baum-Snow, Nathaniel and Derek Neal. 2009. "Mismeasurement of Usual Hours Worked in the Census and ACS" *Economics Letters*, 102:1, 39-41.
- Baum-Snow, Nathaniel and Ronni Pavan. 2011. "Understanding the City Size Wage Gap" *Review of Economic Studies*.
- Bacold, Marigee, Bernardo S. Blum and William S. Strange. 2009. "Skill and the City" *Journal of Urban Economics*, 65:2.
- Bound, John and George Johnson. 1992. "Changes in the Structure of Wages in the 1980s: An Evaluation of Alternative Explanations" *American Economic Review*, 82:3, 371-392.
- Card, David and John E. DiNardo. 2002. "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles" *Journal of Labor Economics*, 20:4 733-783.
- Chay, Kenneth and David S. Lee. 2000. "Changes in Relative Wages in the 1980s: Returns to Observed and Unobserved Skills and Black-White Wage Differentials" *Journal of Econometrics*, 99:1, 1-38.
- Ciccone, Antonio & Robert Hall. 1996. "Productivity and the Density of Economic Activity" *American Economic Review*, 86:1, 54-70.
- DiNardo, John, Nicole Fortin and Thomas Lemieux. 1996. "Labor Market Institutions, and the Distribution of Wages 1973-1992: A Semiparametric Approach," *Econometrica*, 64, 1001-1044.
- Dustmann, Christian, Johannes Ludsteck & Uta Schoenberg. 2009. "Revisiting the German Wage Structure," *Quarterly Journal of Economics*, 124:2, 843-881.
- Galor, Oded and Omer Moav. 2000. "Ability-Biased Technological Transition, Ability Bias and Economic Growth," *Quarterly Journal of Economics*, 115:2, 469-497.
- Glaeser, Edward L., Matt Resseger and Kristina Tobio. 2008. "Urban Inequality", NBER Working Paper #14419.
- Juhn, Chinhui, Kevin Murphy & Brooks Pierce. 1993. "Wage Inequality and the Rise in Returns to Skill" *Journal of Political Economy*, 101:3, 410-442.
- Katz, Lawrence and Kevin Murphy. 1992. "Changes in Relative Wages 1963-1987: Supply and Demand Factors" *Quarterly Journal of Economics*, 107:1 35-78.

Krusell, Per, Lee E. Ohanian, Jose-Victor Rios-Rull, and Giovanni L. Violante. 2000. "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis." *Econometrica*, 68:5 1029-1053.

Lemieux, Thomas. 2006. "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?" *American Economic Review*, 96:3 461-498.

Melly, Blaise. 2005. "Decomposition of Differences in Distribution Using Quantile Regression." *Labour Economics*, 12:4, 577-590.

Machado, José A.F. and José Mata. 2005. "Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression." *Journal of Applied Econometrics*, 20:4, 445-465.

A Accounting for Changes in Being Paid by the Hour

One potential concern about our evidence on the importance of city size for generating changes in wage inequality is changes in measurement error of wages that is correlated with city size. Lemieux (2006) emphasizes that misreporting of annual earnings by workers paid by the hour, a group that has expanded in recent decades, is a potentially important source of wage measurement error in data sets like the census that ask about annual wage and salary income. Because hourly employees may have more difficulty determining their annual earnings than salaried workers, their computed hourly wages may be more unreliable.¹⁸ Alternatively, salaried workers may not accurately recall the true number of hours worked. Unfortunately, the census and ACS do not ask whether a worker is paid by the hour, so we evaluate this potential source of measurement error using the CPS outgoing rotations data.¹⁹ In this data set, the fraction of workers with our sample characteristics that were paid by the hour increased from 0.48 in 1979 to 0.52 in the 2004-2007 period. However, this increase has no strong relationship with city size. Though rural areas experienced the greatest increase at 10 percentage points, relatively large increases of over 0.05 also occurred in city size groups 8 and 10.

We confirm these indications from the raw data by incorporating adjustments for propensity to be paid by the hour into the estimated variance of wage residuals. Following Lemieux (2006), assume that the log wage of a worker paid by the hour is the same as that for an identical worker not paid by the hour with the addition of a linear idiosyncratic error term v . Denote the fraction of demographic group x in location s at time t paid by the hour as $P_t(x, s)$. The variance of the residual component of wages in location s can be decomposed as follows:

$$V_t(s) = \int [P_t(x, s) (V_t(\varepsilon|x, s) + V_t(v|x, s)) + (1 - P_t(x, s))V_t(\varepsilon|x, s)] f_t(x, s) dx$$

Using evidence from Lemieux (2006) that $V_t(v|x, s)$ is about equal over time at 0.022, we can write this as:²⁰

$$V_t(s) = \int V_t(\varepsilon|x, s) f_t(x, s) dx + 0.022 \times \int P_t(x, s) f_t(x, s) dx \quad (12)$$

The first term is what we wish to recover, as it captures the variance of wages over all individuals if they all do not get paid by the hour. The second term is the additional amount of residual error

¹⁸Our sample restriction to include only full-time full-year workers is intended to help reduce this potential measurement error problem.

¹⁹Unfortunately MSA of residence is only partially observed in the 1979 CPS. In this year, we were able to identify only if individuals lived in one of the 5 largest city size categories, a remaining smaller city, or in rural areas.

²⁰Table 3 of Lemieux's (2006) paper indicates that the relative variance of measurement error in March CPS data for men paid by the hour relative to men not paid by the hour is 0.022. His Appendix Figure 3A indicates that the gap between these two measurement error variances did not change much over time. We additionally assume that this value does not vary much with x or s .

that we want to account for.

Our estimates of the second term using CPS data yields additional large city to rural area gaps in the variances of log wages from being paid by the hour of 0.003 in 1979 rising to 0.005 by 2007, an increase that was dwarfed by the overall relative increase in residual variance. Therefore, the increase in the positive relationship between city size and the variance of log wages cleaned of this measurement error is even greater than that for observed log wages documented in Figure 2. Our estimated increases in these gaps as functions of city size are entirely driven by growth in smaller locations where the propensity to be paid by the hour grew the most, belying the stylized facts in Table 1 which shows that the smallest increases in overall log wage inequality occurred in the smallest locations. For these reasons, it is unlikely that changes in the propensity to be paid by the hour has an important effect on our analysis.

B Further Details on Construction of Counterfactual Distributions

B.1 Alternative Counterfactual Mean Wages

This is an alternative specification for counterfactual mean wages to the one presented in Section 3.3. In this alternative approach we assume that, had the city size effect not emerged, log wages in rural areas would not have changed while the average log wages in all other locations would have maintained the same relationship to those in rural areas as existed in 1979:

$$m_t^b(x, s) = m_t(x, 0) + m_{1979}(x, s) - m_{1979}(x, 0)$$

As we show in Table A2, the benchmark counterfactual log wage distribution is only slightly less spread out than actual wage distributions while alternative counterfactual distributions built using these alternative means are markedly tighter than those observed in equilibrium.

B.2 Derivations of Counterfactual Variances Controlling for Industry

As a basis for these counterfactuals, we can write the variance of log wages in year t as

$$V(\ln w_t) = \sum_{d,j,s} \theta_{djst} V(\alpha_{dst} + \beta_{djt} + \delta_{jst}) + \sum_{d,j,s} \theta_{djst} V(\varepsilon_{djst}) \quad (13)$$

where θ_{djst} denotes the share of the total population in year t that is in skill/industry/location cell djs . It is possible to write the variance in this way because membership in any cell djs is exclusive from membership in any other cell. Given small sample sizes in many cells, we apply to our estimates of the within variance the same methodology that we applied to the estimation of

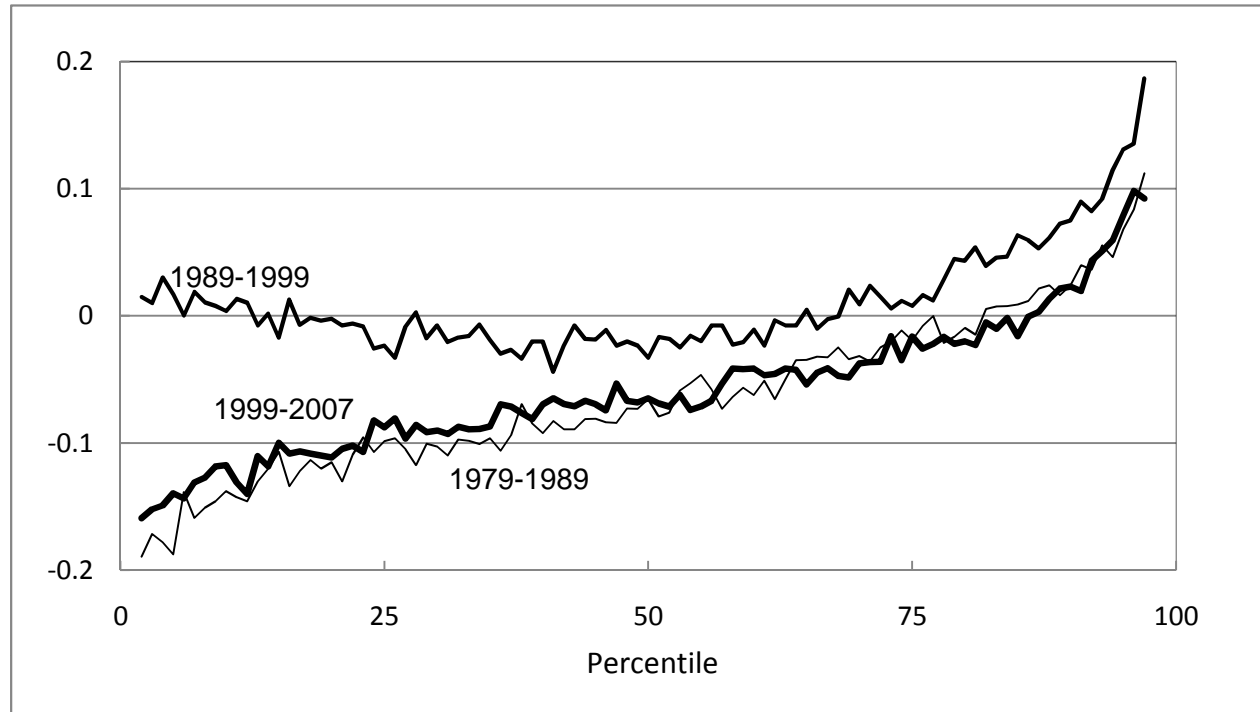
the between components. We calculate $V(\varepsilon_{djst})$ by regressing the squared errors estimated using Equation (11) back on the same set of indicator variables as in Equation (11).

We replace any of the three elements in Equation (13) either with values from 1979 or values that adjust for city size effects that emerged after 1979. To calculate counterfactual variances using 1979 quantities, we replace θ_{djst} with $\theta_{djs1979}$. We apply the same mean adjustment specified in Section 3. We calculate counterfactual residual variances as follows.

$$V^c(\varepsilon_{djst}) = V(\varepsilon_{dj0t}) + V(\varepsilon_{djs1979}) - V(\varepsilon_{dj01979})$$

This method is similar to the nonparametric method used to generate the results in Section 4. Because it is more constrained, this procedure generates counterfactual variances that respond slightly less to city size effects than the results reported in Tables 2 and 3.

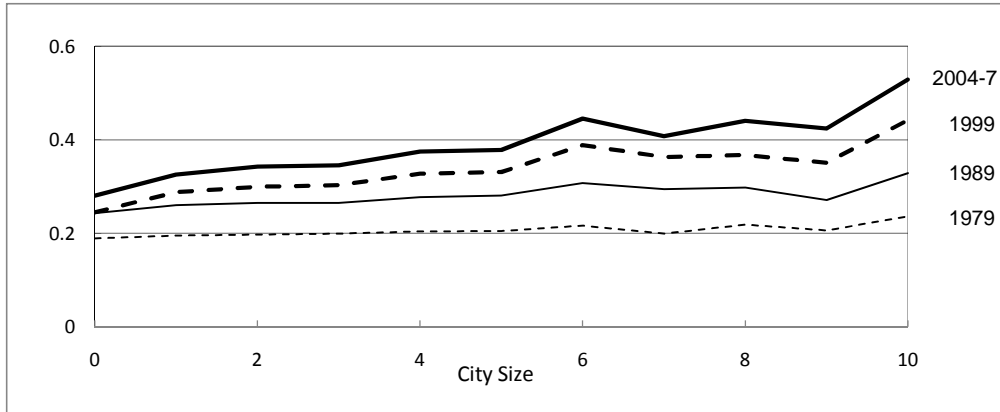
Figure 1: Log Hourly Wage Growth by Percentile, 1979-2007



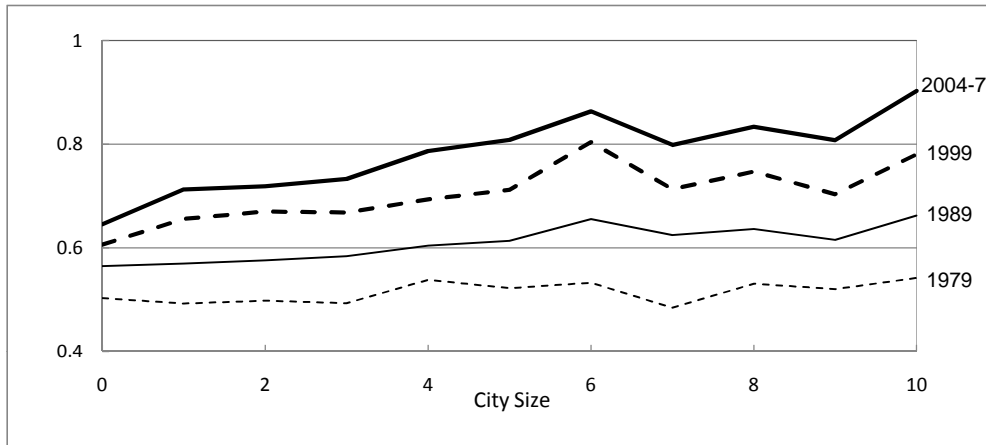
Notes: The sample includes all full-time white male workers ages 25-54 working at least 40 weeks in the listed years. Data is from the census 5% PUMS in 1980, 1990 and 2000 and 1% American Community Surveys (ACS) in 2005, 2006 and 2007. Hourly wages are deflated by the CPI-U and calculated as the logarithm of wage and salary income divided by the product of weeks worked and usual hours worked per week. Observations with imputed demographics, labor supply or wages, the self-employed and those who earned less than 75% of the federal minimum wage in the earnings year are excluded from the sample. Calculations are weighted by sampling weights except for those using the 1980 census which is an unweighted sample. Data listed as being for 2007 actually represents data from full years ending in 2005, 2006 or 2007 with distributions from each year recentered to have a common median.

Figure 2: Wage Inequality by City Size

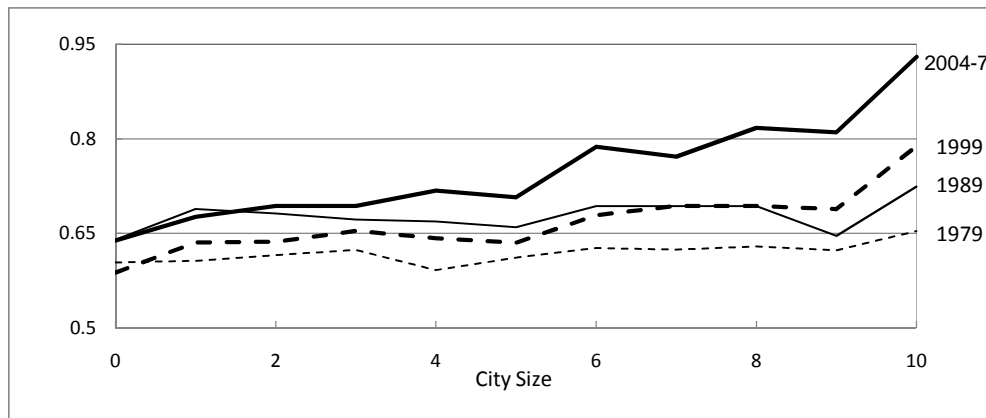
Panel A: Variance of Hourly Wages



Panel B: 90-50 Percentile Gap



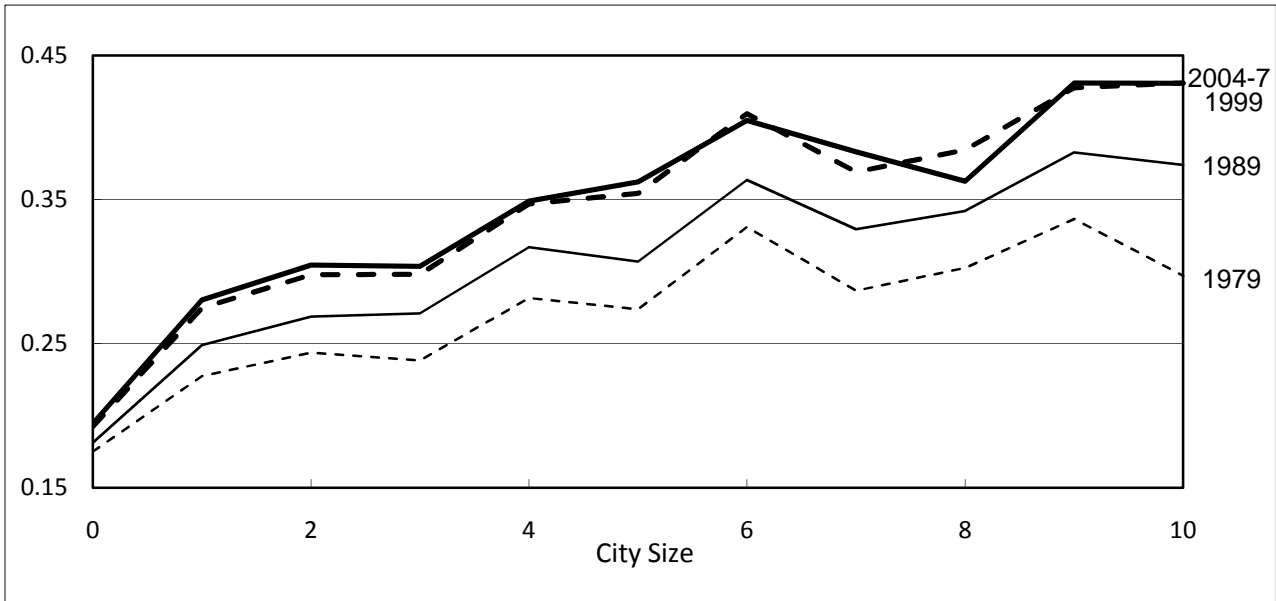
Panel C: 50-10 Percentile Gap



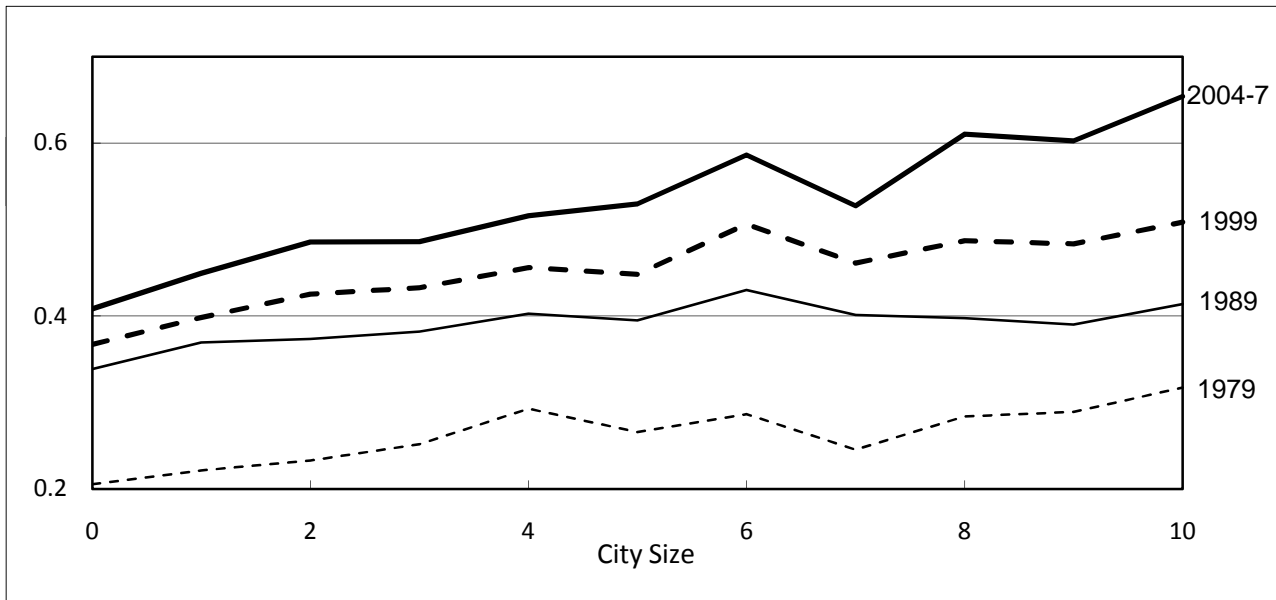
Notes: See the notes to Figure 1 for a description of the sample. City size categories are based on 2000 metro area populations. Size 0 corresponds to non-MSA locations. Sizes 1-10 correspond to ten-percentile bins from the year 2000 MSA population size distribution.

Figure A1: Relative Skill Levels and Wages by City Size Over Time

Panel A: Fraction College or More by City Size



Panel B: College Log Wage Premium by City Size



Notes: Because of a high rates of allocated income in the American Community Surveys for those with a college degree, we do not drop observations with allocated income or labor supply information in any year for the purpose of generating this graph.

Table 1: Trends in Log Wage Inequality

	1	2	3	4	5	6	7
Year	Variance	Total 90-50 Gap	50-10 Gap	Between Variance	Variance	Residual 90-50 Gap	50-10 Gap
1979	0.21	0.53	0.63	0.05	0.16	0.45	0.55
1989	0.29	0.62	0.70	0.08	0.21	0.50	0.60
1999	0.34	0.72	0.67	0.10	0.25	0.56	0.60
2004-7	0.39	0.81	0.72	0.12	0.27	0.61	0.64
1979 to 1989	0.08	0.09	0.07	0.03	0.05	0.05	0.05
1979 to 1999	0.13	0.19	0.04	0.05	0.09	0.11	0.05
1979 to 2004-7	0.18	0.28	0.09	0.07	0.11	0.16	0.09

Notes: See the notes to Figure 1 for a description of the sample. Residuals are calculated using fully interacted age, education and city size cell means of log wages. We use 5 education cells, 15 age cells and 11 city size cells. Residual variance from a more saturated semiparametric specification that also includes 1-digit industry but only 6 age categories is 0.01 smaller in all years.

Table 2: Contributions of City Size to Residual Log Wage Inequality

Calculated Using X Set Adjustment	1 $f_i^q(\varepsilon)$ Full Demog Quantities	2 $f_i^c(\varepsilon)$ Full Demog Residual Prices & Quantities	3 $f_i^n(\varepsilon)$ No Demog Residual Prices & Quantities
Panel A: Variance			
1979 to 1989	0%	20%	26%
1979 to 1999	-2%	49%	59%
1979 to 2004-7	-2%	35%	43%
Panel B: 90 - 50 Percentile Gap			
1979 to 1989	0%	21%	46%
1979 to 1999	-1%	36%	62%
1979 to 2004-7	-1%	30%	50%
Panel C: 50 - 10 Percentile Gap			
1979 to 1989	0%	7%	-2%
1979 to 1999	-3%	104%	88%
1979 to 2004-7	-3%	58%	51%

Notes: Entries indicate the percent reduction in growth of residual log wage inequality measures shown in Table 1, Columns 5-7 due to each of the factors listed in column headers. These fractions are calculated by comparing growth in counterfactual residual inequality to growth in actual residual inequality. In Column 1 we maintain the 1979 distribution of individuals across locations within each demographic group. In Column 2, we additionally maintain the 1979 residual profile with respect to city size within demographic group. For Column 3, we use analogous distributions to those in Column 2 except we assign each person to the same demographic cell. Section 3.2 of the text mathematically specifies how we calculate each counterfactual residual distribution.

Table 3: Contributions to Total Log Wage Inequality

Calculated Using X Set Adjustment	1 $a_t^q(w)$ Full Demog Quantities	2 $a_t^\varepsilon(w)$ Full Demog Residual Prices & Quantities	3 $a_t^c(w)$ Full Demog Total Prices & Quantities	4 $a_t^n(w)$ No Demog Total Prices & Quantities
Panel A: Variance				
1979 to 1989	-2%	10%	16%	26%
1979 to 1999	-2%	31%	34%	50%
1979 to 2004-7	-1%	21%	23%	43%
Panel B: 90 - 50 Percentile Gap				
1979 to 1989	-4%	-14%	-10%	17%
1979 to 1999	2%	14%	17%	50%
1979 to 2004-7	-2%	17%	20%	51%
Panel C: 50 - 10 Percentile Gap				
1979 to 1989	0%	6%	16%	14%
1979 to 1999	-19%	71%	78%	102%
1979 to 2004-7	0%	19%	20%	41%

Notes: Entries indicate the percent reduction in growth of total log wage inequality measures shown in Table 1, Columns 1-3 due to each of the factors listed in column headers. These fractions are calculated by comparing growth in various counterfactual wages as compared to growth in actual wages. See Section 3.3 of the text for complete explanations and mathematical expressions showing how we construct each counterfactual.

Table 4: Counterfactual Reductions in Variance Growth Incorporating Industry

Calculated Using	1	2	3	4	5	6
		$f_t^q(\varepsilon)$	$a_t^q(w)$		$f_t^c(\varepsilon)$	$a_t^c(w)$
		Quantities			Prices & Quantities	
	Between	Residual	Total	Between	Residual	Total
Panel A: Demographics, Industry and City Size						
1979 to 1989	-2%	0%	-1%	9%	16%	13%
1979 to 1999	0%	-2%	-1%	7%	36%	26%
1979 to 2004-7	2%	-2%	0%	4%	22%	15%
Panel B: Demographics and City Size						
1979 to 1989	-4%	-1%	-2%	8%	21%	15%
1979 to 1999	-1%	-2%	-2%	10%	45%	33%
1979 to 2004-7	1%	-2%	-1%	9%	31%	22%

Notes: Entries are calculated analogously to those in which quantity or price and quantity components of the city size effect are taken out in Tables 2 and 3, except using a baseline regression model that also includes one-digit industry indicators interacted with age and education and separately interacted with city size categories. Because this model is more richly specified, the residual variance is 0.01 smaller than that reported in Table 1. While results for Panel B are based on the same decomposition of total variance into between and residual components as in Panel A, they use only demographics and city size cells in the construction of counterfactuals.

Table 5: Percent of Variance Growth Due to Various Factors

	Between	Residual	Total
Panel A: 1979-1989			
Total City Size-Specific	37%	29%	33%
Skill Sorting Across Locations	79%	28%	53%
Industry Sorting Across Locations	-4%	18%	7%
Remainder	25%	54%	40%
Panel B: 1979 to 1999			
Total City Size-Specific	55%	58%	57%
Skill Sorting Across Locations	82%	22%	42%
Industry Sorting Across Locations	5%	17%	13%
Remainder	13%	62%	45%
Panel C: 1979 to 2007			
Total City Size-Specific	62%	41%	49%
Skill Sorting Across Locations	86%	24%	56%
Industry Sorting Across Locations	7%	23%	15%
Remainder	7%	53%	29%

Notes: Each entry labeled "Total City Size-Specific" gives the fraction of the total growth in the variance of hourly wages during the time period indicated in panel headers due to factors correlated with city size. Remaining entries give the fraction of the growth in variance related to city size due to the factors listed at left. Entries are calculated using numbers in Table 4 Columns 7-9.

Table A1: Shifts in Industry Composition by City Size

City Size Index	Sum of Squared Industry Share Changes			
	1979-1989	1989-1999	1999-2004/7	1979-2004/7
Rural	0.000	0.001	0.001	0.006
1	0.002	0.001	0.001	0.011
2	0.002	0.002	0.001	0.015
3	0.003	0.002	0.001	0.017
4	0.002	0.002	0.001	0.015
5	0.003	0.003	0.001	0.017
6	0.003	0.004	0.001	0.016
7	0.003	0.003	0.001	0.018
8	0.002	0.003	0.002	0.015
9	0.004	0.003	0.001	0.017
10	0.004	0.004	0.001	0.020

Notes: Each entry is one-half of the sum of the squared difference in 1-digit industry shares between the years in the column headers for the city size category indicated in each row.

Table A2: Additional Counterfactual Calculations

	1	2	3
Counterfactual	Residuals	Log Wages	Log Wages
Full 1979 Quantities	Yes	Yes	No
City Size Adjustment	None	None	Total Prices b & Quantities

Panel A: Variance

1979 to 1989	0%	0%	16%
1979 to 1999	19%	14%	40%
1979 to 2004-7	13%	12%	33%

Panel B: 90 - 50 Percentile Gap

1979 to 1989	-3%	-20%	-7%
1979 to 1999	11%	7%	24%
1979 to 2004-7	9%	11%	29%

Panel C: 50 - 10 Percentile Gap

1979 to 1989	-14%	-15%	14%
1979 to 1999	19%	25%	102%
1979 to 2004-7	12%	9%	40%

Notes: Column 1 gives reductions in the growth of indicated residual inequality measures holding the full distribution of observables at 1979 quantities. Column 2 gives reductions in the growth of indicated log wage inequality measures holding the full distribution of observables at 1979 quantities. Column 3 is analogous to Table 3 Column 3 except that it uses the less restrictive mean adjustment detailed in Appendix A.2. Counterfactual distributions used to calculate numbers in the three columns are $\int g_t(\varepsilon | x, s)h_{1979}(x, s)dsdx$, $\int g_t(w - m_t(x, s) | x, s)h_{1979}(x, s)dsdx$ and $\int g_t^p(w - m_t^b(x, s) | x, s)h_{a1979}(s | x)h_{b_t}(x)dsdx$ respectively.