UKCPR
University of Kentucky Center for Poverty Research

Discussion Paper Series
DP 2010-07
ISSN: 1936-9379

Poverty, Inequality and Cost of Living Differences

Enrico Moretti
Professor of Economics
Department of Economics
University of California - Berkeley

Preferred citation

Moretti, Enrico. (2010, Aug.). Poverty, Inequality and Cost of Living Differences. *University of Kentucky Center for Poverty Research Discussion Paper Series, DP2010-07.*

Author correspondence

Enrico Moretti, University of California - Berkeley, Department of Economics, 549 Evans Hall, Berkeley, CA 94720-3880. E-mail: moretti@econ.berkeley.edu

University of Kentucky Center for Poverty Research, 302D Mathews Building, Lexington, KY, 40506-0047
Phone: 859-257-7641; Fax: 859-257-6959; E-mail: mary.boulton@uky.edu

www.ukcpr.org

EO/AA

Poverty, Inequality and Cost of Living Differences

Enrico Moretti
UC Berkeley, NBER, CEPR and IZA

August 2010

I thank James Ziliak, David Wildasin and seminar participants at University of Kentucky for many useful comments.

1 Introduction

Metropolitan areas in the US are characterized by enormous differences in average income, earnings, and factor productivity. The income of individuals located in metropolitan areas at the top of the income distribution is more than double the income of observationally similar individuals located in metropolitan areas at the bottom of the distribution. These differences reflect, at least in part, variation in local productivity.

Metropolitan areas in the US are also characterized by enormous differences in cost of living. These differences in cost of living are mostly due to differences in the cost of land and therefore housing. For example, the average cost of housing in metropolitan areas like Anniston, AL or Decatur, AL is about half of the average cost of housing in metropolitan areas like Naples, FL or Atlanta, GA. The existing data on average income by city or state do not take into account these cost of living differences, and therefore provide a potentially inaccurate picture of the geographical distribution of households with a low standard of living. Additionally, even within a metropolitan area, different income groups may be exposed to different price levels, if for example, retailers are less available in poor areas, or if the poor have higher search costs because of lack of transportation. Also, to the extent that high and low income households are not distributed uniformly across metropolitan areas, differences in cost of living will also affect our existing aggregate measures of inequality.

When thinking about differences in income across localities, we are presumably interested in measures that reflect differences in standard of living across households. When designing public assistance programs that help the poor, we are presumably interested in targeting groups with a low standard of living. In this respect, measures of income and income inequality that account for differences in cost of living are important for policy because they are a arguably better measures of differences in standard of living than the nominal measures that are currently used. The measures of income and income inequality that are currently used fail to take into account the fact that different income groups face different cost of living trends.

In this report, I seek to re-examine how inequality across metropolitan areas and across skill groups is measured and how it should be interpreted. I investigate the effect that accounting for cost of living differences has on existing measures of differences across cities and between skill groups in income and wages.

Specifically, I propose two measures of local cost of living in order to adjust existing measures of differences in income levels across metropolitan areas and existing measures of inequality across skill groups.

My empirical finding indicate that accounting for cost of living differences significantly reduces income differences across metropolitan areas and significantly reduces measures of inequality across skill groups. I then present a theoretical framework that is useful in interpreting my empirical findings and draw policy implications.

2 Consumer Price Indexes and Income Distribution

There is vast heterogeneity in the cost of living experienced by different households in the US. Differences in cost of living arise from differences in cost of housing across areas and difference in the consumption of non-housing goods and services. Not all workers are exposed to the same changes in cost of living because (i) they may not live in the same areas and therefore they may experience different changes in housing prices; (ii) they may not consume the same bundle of goods and (iii) may not have access to the same set of retailers.

For example, some authors have hypothesized that poor households without cars have lower mobility than rich households and therefore face higher search costs for consumption goods. Lower mobility may result in more limited access to discount retailers and therefore higher cost of purchasing non-durable goods and services (for example: food and groceries) as well as durable goods (for example: household appliances, etc). If this is true, it implies that poor households face higher cost of living and their real earnings may be even smaller than their nominal earnings (relative to rich households).

In this case, real poverty and inequality are even larger than the existing measures would suggest, because the existing measures are based on nominal income and earnings and do not account for differences in the price of consumption goods.

On the other hand, there may be differences in cost of living that arise from geographical differences. For example, skilled workers are overrepresented in metropolitan areas that have a high cost of housing, while unskilled workers are overrepresented in metropolitan areas that have low cost of housing. This implies that skilled workers are exposed to a higher cost of housing and that their real relative wage may be smaller than their nominal relative wage. In this case, real inequality is smaller than the existing measures of inequality.

2.1 The Official CPI

A cost of living index seeks to measure changes over time in the amount that consumers need to spend to reach a certain utility level or "standard of living." Changes in the official Consumer Price Index between period t and $t + 1$ as measured by the Bureau of Labor Statistics are a weighted average of changes in the price of the goods in a representative consumption basket. The basket is the original consumption basket at time t, and the weights reflect the share of income that the average consumer spends on each good at time t.

One well known problem with the CPI is the potential for substitution bias, which is the possibility that consumers respond to price changes by substituting relatively cheaper goods for goods that have become more expensive. While the actual consumption baskets may change, the CPI reports inflation for the original basket. Details of the BLS methodology are described in Chapter 17 of the Handbook of Methods (BLS, 2007), titled "The Consumer Price Index".

Appendix table 1 shows the relative importance of the main aggregate components of the CPI-U in 2000. The largest component by far is housing. In 2000, housing accounts for more than 42% of the CPI-U. The largest sub-components of housing costs are "Shelter" and "Fuel and Utilities". The second and third main components of the CPI-U are transportation and food. They only account for 17.2% and 14.9% of the CPI-U, respectively. The weights of all the other categories are 6% or smaller.

Although most households in the US are homeowners, changes in the price of housing are measured by the BLS using changes in the cost of renting an apartment (Poole, Ptacek and Verbugge, 2006; Bureau of Labor Statistics, 2007). The rationale for using rental costs instead of home prices is that rental costs are a better approximation of the user cost of housing. Since houses are an asset, their price reflects both the user cost as well as expectations of future appreciation.

Rental costs vary significantly across metropolitan areas. For example, in 2000, the average rental cost for a 2 or 3 bedroom apartment in San Diego, CA—the city at the 90th percentile of the distribution—is \$894. This rental cost is almost 3 times higher than the rental cost for an equally sized apartment in Decatour, AL, the city at the 10th percentile. Changes over time in rental costs also vary significantly across metropolitan areas. For example, between 1980 and 2000, the rental cost increased by \$165 in Johnstown, PA—one of the cities at the bottom of the distribution—and by \$892 in San Jose—one of the cities at the top of the distribution. The distribution of average rental costs and changes in average rental costs are shown in Figure 1.

Although the cost of living varies substantially across metropolitan areas, wage and income are typically deflated using a single, nation-wide deflator, such as the CPI-U calculated by the BLS. The use a nation-wide deflator is particularly striking in light of the fact that more than 40% of the CPI-U is driven by housing costs , and that housing costs vary so much across locations (Figure 1).

To investigate the role of cost of living differences on income and wage differences across cities and between skill groups, I propose two alternative CPI indexes that vary across metropolitan areas. I closely follow the methodology that the Bureau of Labor Statistics uses to build the official Consumer Price Index, but I generalize two of its assumptions.

2.2 Local CPI 1

First, I compute a CPI that allows for the fact that the cost of housing varies across metropolitan areas. I call the resulting local price index "Local CPI 1". Following the BLS methodology, I define Local CPI 1 as the properly weighted sum of local cost of housing—with the average across cities normalized to 1 in 1980—and non-housing consumption—normalized to 1 in 1980. I measure the cost of housing faced by an individual in metropolitan area c in two ways. In my preferred specification, I follow the BLS methodology and I use rental costs. I assign the cost of housing to residents in a metropolitan area based on the relevant average monthly rent. Specifically, I take the average of the monthly cost of renting a 2 or 3 bedroom apartment among all renters in area c. As an alternative way to measure cost of housing, in some models I use the price of owner occupied houses instead of rental costs. Specifically, I take the average reported value of all 2 or 3 bedroom owner occupied single family houses in area c. Both rental costs and housing prices are from the Census of Population. As I discuss later, empirical results are not sensitive to measuring housing costs using rental costs or housing prices. The price of non housing goods and services is assumed to be the same in a given year, irrespective of location. This assumption is relaxed in Local CPI 2.

The motivation for using 2 and 3 bedroom apartments is to keep the size of the apartment roughly similar. I have experimented with variants of this selection rule. Estimates based only on 2 bedroom apartments are similar to the ones presented below. Estimates based on data from the American Housing Survey that use information on square footage to hold constant the exact size of the apartment also yield similar results.

It is important to note that this methodology ensures that the deflator that I use for a given worker does not reflect the increase in the cost of the apartment rented or the cost of the house owned by that specific worker. Instead, it reflects the increase in the cost of housing experienced by residents in the same city, irrespective of their own individual housing cost and irrespective of whether they rent or own.

2.3 Local CPI 2

In local CPI 1, changes in the cost of housing can vary across localities, but changes in the cost of non-housing goods and services are assumed to be the same everywhere. While the cost of housing is the most important component of the CPI, the price of other goods and services is likely to vary systematically with the cost of housing. In cities where land is more expensive, production and retail costs are higher and therefore the cost of many goods and services is higher. For example, a slice of pizza or a hair cut are likely to be more expensive in New York city than in Indianapolis, since it is more expensive to operate a pizza restaurant or a barber shop in New York city than Indianapolis.

Local CPI 2 allows for both the cost of housing and the cost of non-housing consumption to vary across metropolitan areas. Systematic, high quality, city-level data on the price of non-housing good and services are not available for most cities over a long time period. To overcome this limitation, I use two alternative approaches. First, in my preferred specification, I use the fact that the BLS releases a local CPI for a limited number of metropolitan areas. This local CPI is not ideal because of the 315 MSA's in the 2000 Census, the metropolitan-level CPI is made available by the BLS only for 23 MSA's in the period under consideration. Additionally, it is normalized to 1 in a given year, thus precluding cross-sectional comparisons. However, it can still be used to impute the part of local non-housing prices that varies systematically with housing costs. The local CPI computed by the BLS for city c in year t is a weighted average of housing cost (HP_{ct}) and non-housing costs (NHP_{ct}): $\text{BLS}_{ct} = w\text{HP}_{ct} + (1-w)\text{NHP}_{ct}$ where w is the CPI weight used by BLS for housing. Non-housing costs can be divided in two components:

$$\text{NHP}_{ct} = \pi\text{HP}_{ct} + v_{ct} \tag{1}$$

where πHP_{ct} is the component of non-housing costs that varies systematically with housing costs; and v_{ct} is the component that is orthogonal to housing costs. If $\pi > 0$ it means that cities with higher cost of housing also have higher costs of non-housing goods and services. I use the small sample of MSA's for which a local BLS CPI is available to estimate π.

To do so, I first regress changes in the BLS local index on changes in housing costs: $\Delta\text{BLS}_{ct} = \beta\Delta\text{HP}_{ct} + e_{ct}$. Estimating this regression in differences is necessary because BLS_{ct} is normalized to 1 in a given year. While cross-sectional comparisons based on BLS_{ct} are meaningless, BLS_{ct} does measure changes in prices within a city. Once I have an estimate of β, I can calculate $\hat{\pi} = \frac{\hat{\beta}-w}{1-w}$. Empirically, $\hat{\beta}$ is equal to .588 (.001) and $\hat{\pi}$ is equal to .35 in 2000.

I then impute the systematic component of non-housing costs to all MSA's, based on their housing cost: $E(\text{NHP}_{ct}|\text{HP}_{ct}) = \hat{\pi}\text{HP}_{ct}$. Finally, I compute "Local CPI 2" as a properly weighted sum of the cost of housing, the component of non-housing costs that varies with housing ($\hat{\pi}HP_{ct}$), and the component of non-housing costs that does not vary with housing.

As an alternative strategy to measure local variation in non-housing prices, I use data on non-housing prices taken from the Accra dataset, which is collected by the Council for Community and Economic Research.[1] The Accra data have both advantages and disadvantages. On one hand, the Accra data are available for most cities, and therefore do not require any imputation. Furthermore, the detail is such that price information is available at the level of specific consumption goods and the price is not normalized to a base year. On the other hand, the Accra data are available only for a very limited number of goods.[2] Importantly, the sample size for each good and city is quite small, so that local price averages are noisy. Additionally, the set of cities covered changes over time. In practice, the empirical findings based on the version of local CPI 2 that uses the imputation and those based on the version of local CPI 2 that uses Accra data are similar.

In sum, local CPI 2 is more comprehensive than Local CPI 1 because it includes local variation in both housing and non-housing costs, but it is has the limitation that non-housing costs are imputed or come from Accra data. For this reason, in the next Section I present separate estimates for Local CPI 1 and Local CPI 2.

[1] The data were generously provided by Emek Basker. Basker (2005) and Basker and Noel (2007) describe the Accra dataset in detail.

[2] Only 48 goods have prices that are consistently defined for the entire period under consideration. The BLS basket includes more than 1000 goods.

2.4 Additional Details

Here I describe in more details on how I compute Local CPI 1 and Local CPI 2. As I mention above, I follow closely the BLS methodology, and take the properly weighted sum of changes in the cost of housing and non-housing consumption. Cost of housing is measured either using rental costs or housing prices. In the first case, my measure of rent is the "gross monthly rental cost" of the housing unit.

I limit the sample to 2 or 3 bedrooms rental units. This includes contract rent plus additional costs for utilities (water, electricity, gas) and fuels (oil, coal, kerosene, wood, etc.). This variable is considered by IPUMS as more comparable across households than "contract rent", which may or may not include utilities and fuels. The Department of Housing and Urban Development (HUD) also uses the "gross monthly rental cost" measure of rent to calculate the federally mandated "Fair Market Rent".

Rents are imputed for top-coded observations by multiplying the value of the top code by 1.3. Results do not change significantly when no imputation is performed or when I multiply the value of the top code by 1.4. For Local CPI 1, the cost of non-housing consumption is obtained by subtracting changes in the cost of housing from the nationwide CPI-U computed by the BLS:

$$\text{CPI Non-Housing} = (\text{CPI-U}/(1-w)) - (w/(1-w))\text{Housing} \tag{2}$$

where "Housing" is the average nationwide increase in cost of housing (from Census data) and w is the BLS housing weight in the relevant year.

The housing costs relevant for a worker living in metropolitan area c—whether he rents or own—is the average of the monthly cost of renting a 2 or 3 bedroom apartment among all renters in area c. When cost of housing is measured using housing prices, I use the property value reported by homeowners of 2 or 3 bedroom single family houses. In this case, the housing costs relevant for a worker living in metropolitan area c is then the average of housing values reported by all homeowners of 2 or 3 bedroom homes in area c.

Note that measured changes in cost of housing do not reflect the change in rental cost or changes in property values at the individual level. Instead, measured changes in cost of housing reflect an average for the local housing market, irrespective of an individual own housing cost and irrespective of whether she rents or owns.

As weights, in my baseline specifications I use the expenditure shares that the BLS uses to compute the official CPI. Since the basket is updated periodically, the BLS weights vary by year. One concern is that housing expenditure shares may vary across metropolitan areas because differences in housing prices. Additionally, it is possible that housing expenditure shares vary across skill groups if preferences are non homotetic. In practice, however, the use of BLS shares does not appear to introduce a significant bias in my estimates of the local CPI.

First, consider the possible differences in expenditure shares across metropolitan areas.

Since housing costs vary across cities, it is in principle possible that the share of income spent on housing also vary, as consumers adjust their consumption bundles to local prices. Empirically, the demand for housing is not very price elastic and the share of income spent on housing appears to be higher in more expensive cities. In a recent AER paper, Lewbel and Pendakur (forthcoming) find that a housing price increase of 10 percent results in a 0.63 percentage points *higher* housing share, everything else constant. If this is true, it implies that the share of income spent on housing in expensive cities like New York is higher than the share of income spent on housing in less expensive cities like Indianapolis, everything else constant. Because college graduates are over-represented in expensive cities like New York and underrepresented in less expensive cities like Indianapolis, this should increase the housing share of college graduates relative to high-school graduates, everything else constant.

Second, consider the possibility that housing price elasticity vary by skill level (or income level). Lewbel and Pendakur find that high income individuals substitute less than low income individuals in the face of an increase in the price of housing. This should further increase the housing share of college graduates relative to high-school graduates, everything else constant.

Third, consider the possibility of non homotetic preferences. Most empirical studies find that housing is a normal good, with an income elasticity just below 1 when income is measured as permanent income.[3] If this is true, the share of income spent on housing should be slightly lower for college graduates than high-school graduates.

To account for these possibilities, I have replicated my results using different expenditure shares for different cities and different skill groups in different years. In particular, I use available estimates in the literature of price elasticity and income elasticity to impute shares that vary as a function of local housing prices and individual income. For housing, I assume a permanent income elasticity equal to .85, which is the mid-point in the range of estimates provided by Polinsky and Ellwood (1979). I also assume that the percent difference in permanent income between skilled and unskilled workers is 40% in 1980, 53% in 1990 and 60% in 2000. (These figures reflect estimates of the the nominal college premium.) To allow for differences across cities as a function of local housing prices, I use estimates of demand elasticity from Lewbel and Pendakur (AER forthcoming).

Estimates of the college premium based on expenditure shares that vary by MSA, skill group and year are similar to the ones obtained using BLS shares that vary only by year. Overall, using a common housing share for all individuals within a year appears not to be a bad approximation. This is consistent with what reported by Baum-Snow and Pavan (2009), who find that expenditures shares are generally similar across cities of different size (and therefore different price level).

[3] For example, Polinsky and Ellwood (1979) uncover estimates of permanent income elasticity ranging from 0.80 to 0.87.

3 Inequality Across Metropolitan Areas

I use my proposed measured of local cost of living to I investigate how income differences across metropolitan areas vary when cost of living are taken into account. In this Section, I focus on data from the 2000 Census of Population. I report my results on individual income. However, an analysis of yearly earning or monthly earnings or hourly wages all yield very similar conclusions.

Detailed results by metropolitan area are shown in Table 1. Column 2 reports average nominal personal income, while column 3 reports average real personal income based on Local CPI 1. Column 4 reports the percent difference between real and nominal income. The table illustrates that in cities like San Jose, Stanford and San Francisco, where cost of living is particularly high, real income is more than 20% below nominal income. By contrast, in cities like Gadsden, Anniston or Johnstown where cost of living is particularly low, real income is more than 15% above nominal income.

Because cities with high nomina income tend to have higher cost of living on average, deflating by local cost of living tends to dramatically lower the geographical dispersion in mean income across geographical areas.

To see this more systematically, consider that in the 332 metropolitan areas identified in the 2000 Census of population, the mean personal income by metropolitan area in nominal dollars is 32,144, while the median personal income is 31575. The standard deviation of the distribution of nominal income is 6123.1.

The difference between the metropolitan area with the highest average personal income and the city with the lowest is 57654. The difference between the metropolitan area in top 1% of the average personal income distribution and the metropolitan area in bottom 1% of the average personal income distribution is 26986. The difference between the metropolitan area in top 10% of the average personal income distribution and the metropolitan area in bottom 10% of the average personal income distribution is 14776.

By contrast, the distribution in real terms is much more compressed. For example, deflating nominal income using Local CPI 1—where the average across cities of Local CPI 1 is set equal to 1—I find a standard deviation of the distribution of nominal income equal to 4127.0. The difference between the metropolitan area with the highest average personal income and the city with the lowest is 40004. The difference between the metropolitan area in top 1% of the average personal income distribution and the metropolitan area in bottom 1% of the average personal income distribution is 18742. The difference between the metropolitan area in top 10% of the average personal income distribution and the metropolitan area in bottom 10% of the average personal income distribution is 9812. When using local CPI 2, the increased in the compression of the income distribution across cities is even larger.

4 Inequality Across Skill Groups

In this Section, I estimate how much of the increase in nominal wage differences between skilled and unskilled individuals is accounted for by differences in the cost of living.

I use data from the Census of Population and focus on changes between 1980 and 2000 in the difference in the average hourly wage for workers with a high school degree and workers with college or more.

I begin by quantifying the changes in the cost of living experienced by high school and college graduates between 1980 and 2000. The top panel of Table 2 shows changes in the official CPI-U, as reported by the BLS, and normalized to 1 in 1980. This is the most widely used measure of inflation, and it is the measure that is almost universally used to deflate wages and incomes. According to this index, the price level doubled between 1980 and 2000. This increase is—by construction—the same for college graduates and high school graduates.

The next panel shows the increase in the cost of housing faced by college graduates and high school graduates. College graduates and high school graduates are exposed to very different increases in the cost of housing. In 1980 the cost of housing for the average college graduate is only 4% more than the cost of housing for the average high school graduate. This gap grows to 11% in 1990 and reaches 14% by 2000. Column 4 indicates that housing costs for high school and college graduates increased between 1980 and 2000 by 127% and 147%, respectively.

The third panel shows "Local CPI 1", normalized to 1 in 1980 for the average household.[4] The panel shows that in 1980 the overall cost of living experienced by college graduates is only 2% higher than the cost of living experienced by high school graduates. This difference increases to 6% by year 2000. The difference in Local CPI 1 between high school and college graduates is less pronounced than the difference in monthly rent because Local CPI 1 includes non-housing costs as well as housing costs.

The differential increase in cost of living faced by college graduates relative to high school graduates is more pronounced when the price of non-housing goods and services is allowed to vary across locations, as in the bottom panel. In the case of Local CPI 2, the cost of living is 3% higher for college graduates relative to high school graduates in 1980 and 9% in 2000. Column 4 indicates that the increase in the overall price level experienced by high school graduates between 1980 and 2000 is 108%. The increase in the overall price level experienced by college graduates between 1980 and 2000 is 119%.

The relative increase in the cost of housing experienced by college graduates between 1980 and 2000 can be decomposed into a part due to geographical mobility and a part due to the fact that already in 1980 college graduates are overrepresented in cities that experience large increases in costs. Specifically, the 1980-2000 nationwide change in the cost of housing experienced by skill group j (j=high school or college), can be written as

[4] Here I use rental costs to measure housing costs. Using property values for owner occupied houses yields similar results.

$$P_{j2000} - P_{j1980} = \sum_c \omega_{jc2000} P_{c2000} - \sum_c \omega_{jc1980} P_{c1980}$$
$$\sum_c (\omega_{jc2000} - \omega_{jc1980}) P_{c2000} + \sum_c \omega_{jc1980} (P_{c2000} - P_{c1980})$$

where ω_{jct} is the share of workers in skill group j who live in city c in year t and P_{ct} is the cost of housing in city c in year t. The equation illustrates that the total change in cost of housing is the sum of two components: a part due to the the change in the share of workers in each city, given 2000 prices ($\sum_c (\omega_{jc2000} - \omega_{jc1980}) P_{c2000}$); and a part due to the differential change in the cost of housing across cities, given the 1980 geographical distribution ($\sum_c \omega_{jc1980} (P_{c2000} - P_{c1980})$). The change in cost of housing of college graduates *relative* to high school graduates is therefore the difference of these two components for college graduates and high school graduates.

Empirically, I find that both factors are important. About 43% of the total increase in cost of housing of college graduates relative to high school graduates is due to the first component (geographical mobility of college graduates toward expensive cities), and 57% is due to the second component (larger cost increase in cities that have many college graduates in 1980).

Model 1 in the top panel of Table 3 estimates the conditional <u>nominal</u> wage difference between workers with a high school degree and workers with college or more, by year. Estimates in columns 1 to 4 are from a regression of the log nominal hourly wage on an indicator for college interacted with an indicator for year 1980, an indicator for college interacted with an indicator for year 1990, an indicator for college interacted with an indicator for year 2000, years dummies, a cubic in potential experience, and dummies for gender and race. Estimates in columns 5 to 8 are from models that also include MSA fixed effects. Entries are the coefficients on the interactions of college and year and represent the conditional wage difference for the relevant year. The sample includes all US born wage and salary workers aged 25-60 who have worked at least 48 weeks in the previous year.[5]

My estimates in columns 1 to 4 indicate that the conditional nominal wage difference between workers with a high school degree and workers with college or more has increased significantly. The difference is 40% in 1980 and rises to 60% by 2000. Column 4 indicates that this increase amounts to 20 percentage points. This estimate is generally consistent with the previous literature (see, for example, Table 3 in Katz and Autor, 1999).

Models 2 and 3 in Table 3 show the conditional <u>real</u> wage differences between workers with a high school degree and workers with college or more. To quantify this difference, I estimate models that are similar to Model 1, where the dependent variable is the nominal wage divided by Local CPI 1 (in Model 2) or by Local CPI 2 (in Model 3). Two features are noteworthy. First, the level of the conditional college premium is lower in real terms than

[5]The sample includes both men and women. This may be a concern, since in a recent paper by Black et al. (2010) shows that female labor force participation is different in different cities. At the end of this subsection, I discuss a number of alternative specifications, including one when I estimate the college premium for men and women separetely. Estimates by gender are similar to those obtained from the pooled sample.

in nominal terms in each year. For example, in 2000 the conditional difference between the wage for college graduates and high school graduates is .60 in nominal terms and only .53 in real terms when Local CPI 1 is used as deflator. The difference is smaller—.51 percentage points—when Local CPI 2 is used as deflator. Second, the increase between 1980 and 2000 in college premium is significantly smaller in real terms than in nominal terms. For example, using Local CPI 1, the 1980-2000 increase in the conditional real wage difference between college graduates and high school graduates is 15 percentage points. In other words, cost of living differences as measured by Local CPI 1 account for 25% of the increase in conditional inequality between college and high school graduates between 1980 and 2000 (column 4).

The effect of cost of living differences is even more pronounced when the cost of living is measured by Local CPI 2. In this case, the increase in the conditional real wage difference between college graduates and high school graduates is 14 percentage points. This implies that cost of living differences as measured by Local CPI 2 account for 30% of the increase in conditional wage inequality between college and high school graduates between 1980 and 2000 (column 4).

When I control for fixed effects for metropolitan areas in columns 5-8, the nominal college premium is slightly smaller, but the real college premium is generally similar. The increase in the college premium is 18 percentage points when measured in nominal terms, and 14-15 percentage points when measured in real terms, depending on whether CPI 1 or CPI 2 is used as deflator. After conditioning on MSA fixed effects, cost of living differences account 22% of the increase in conditional inequality between college and high school graduates between 1980 and 2000 when CPI 2 is used as a deflator (column 8).

5 A Simple Framework

The interpretation of differences in real income is not straightforward. If amenities differ across cities, differences in real income do not necessarily equal differences in well-being. In this Section, I use a simple general equilibrium model to investigate the implications of my empirical findings for well-being disparities. The implications are different depending on the reasons for the increase in the share of college graduates in expensive cities. I consider two alternative explanations for such an increase.

1. First, it is possible that skilled workers move to expensive cities because the <u>relative demand</u> of skilled labor increases in expensive cities, as firms located in these cities increasingly seek to hire skilled labor. This can be due to localized skill-biased technical change or positive shocks to the demand faced by industries that employ skilled workers and are located in expensive cities (for example, high tech, finance, etc.). In this case, the increase in utility disparity between skilled and unskilled workers is smaller than the increase in nominal wage disparity, because the higher nominal wage of skilled workers is in part off-sets by higher cost of living in the cities where skilled jobs are located.

2. Alternatively, it is possible that skilled workers move to expensive cities because the relative supply of skilled labor increases in expensive cities, as skilled workers are increasingly attracted by amenities located in those cities. In this case, a higher cost of housing reflects consumption of desirable local amenities. Since this consumption arguably generates utility, it is possible to have large increases in utility disparities even when increases in real wage disparities are limited.

To formalize these two alternative hypotheses, and what they imply for inequality in utility and wages, I consider a simple general equilibrium model of the labor and housing market. The model is a generalization of the Roback (1982, 1988) model and has two types of workers, skilled workers (type H) and unskilled workers (type L). Like in Roback, workers and firms are mobile and choose the location that maximizes utility or profits. But unlike Roback, the elasticity of local labor supply is not infinite, so that productivity and amenity shocks are not always fully capitalized into land prices. This allows shocks to the relative demand and relative supply of skilled workers to have different effects on the utility of skilled and unskilled workers.

For simplicity of exposition, I model the two explanations as mutually exclusive. In the empirical tests that seek to distinguish between the two explanations (Section 6), I allow for the possibility that both demand and supply forces are at play at the same time.

5.1 Assumptions and Equilibrium

I assume that each city is a competitive economy that produces a single output good y which is traded on the international market, so that its price is the same everywhere and set equal to 1. Like in Roback, I abstract from labor supply decisions and I assume that each worker provides one unit of labor, so that local labor supply is only determined by workers' location decisions. The indirect utility of skilled workers in city c is assumed to be

$$U_{Hic} = w_{Hc} - r_c + A_{Hc} + e_{Hic} \tag{3}$$

where w_{Hc} is the nominal wage in the city; r_c is the cost of housing; A_{Hc} is a measure of local amenities. The random term e_{Hic} represents worker i idiosyncratic preferences for location c. A larger e_{Hic} means that worker i is particularly attached to city c, holding constant real wage and amenities. For example, being born in city c or having family in city c may make city c more attractive to a worker. Similarly, the indirect utility of unskilled workers is

$$U_{Lic} = w_{Lc} - r_c + A_{Lc} + e_{Lic} \tag{4}$$

In equations 3 and 4, skilled and unskilled workers in a city compete for housing in the same housing market and therefore face the same price of housing. This allows a shock to one group to be transmitted to the other group through its effect on housing prices.[6] While they have access to the same local amenities, different skill groups do not need to value these amenities equally: A_{Hc} and A_{Lc} represent the skill-specific value of local amenities.

Assume that there are two cities—Mobile (city a) and San Francisco (city b)—and a fixed number of workers is divided between the two cities. Tastes for location can vary by skill group. Specifically, skilled workers' and unskilled workers' relative preferences for city a over city b are, respectively

$$e_{Hia} - e_{Hib} \sim U[-s_H, s_H] \tag{5}$$

and

$$e_{Lia} - e_{Lib} \sim U[-s_L, s_L] \tag{6}$$

The parameters s_H and s_L characterize the importance of idiosyncratic preferences for location and therefore the degree of labor mobility. If s_H is large, for example, it means that preferences for location are important for skilled workers and therefore their willingness to move to arbitrage away real wage differences or amenity differences is limited. On the other hand, if s_H is small, preferences for location are not very important and therefore skilled workers are more willing to move in response to differences in real wages or amenities. In the extreme, if $s_H = 0$ skilled workers' mobility is perfect.

A worker chooses city a if and only if $e_{ia} - e_{ib} > (w_b - r_b) - (w_a - r_a) + (A_b - A_a)$. In equilibrium, the marginal worker needs to be indifferent between living in Mobile and San Francisco. This implies that skilled workers' labor supply is upward sloping, with the slope that depends on s. For example, the supply of skilled workers in San Francisco is:

$$w_{Hb} = w_{Ha} + (r_b - r_a) + (A_a - A_b) + s_H(\frac{N_{Hb} - N_{Ha}}{N}) \tag{7}$$

where N_{Hb} is the log of the number of skilled workers hired in San Francisco and $N = N_{Ha} + N_{Hb}$. If idiosyncratic preferences for location are not very important (s_H is small), then workers are very mobile and the supply curve is relatively flat. If idiosyncratic preferences for location are very important (s_H is large), then workers are rather immobile and the supply curve is relatively steep. Moreover, an increase is the real wage in Mobile, or an improvement in the relative amenities shifts back the labor supply curve in San Francisco.[7]

For simplicity, I focus on the case where skilled and unskilled workers in the same city work in different firms. This amounts to assuming away imperfect substitution between skilled and unskilled workers. This assumption simplifies the analysis, and it is not crucial

[6] It is easy to relax this assumption by assuming some residential segregation by skill level within a city.

[7] An important difference between the Rosen-Roback setting and this setting is that in Rosen-Roback, all workers are identical, and always indifferent across locations. In this setting, workers differ in their preferences for location. While the marginal worker is indifferent between locations, here there are inframarginal workers who enjoy economic rents. These rents are larger the smaller the elasticity of local labor supply.

(Moretti, 2010). The production function for firms in city c that use skilled labor is Cobb-Douglas with constant returns to scale: $\ln y_{Hc} = X_{Hc} + hN_{Hc} + (1-h)K_{Hc}$, where K_{Hc} is the log of capital and X_{Hc} is a skill and city-specific productivity shifter. Firms are assumed to be perfectly mobile. If firms are price takers and labor is paid its marginal product, labor demand for skilled labor in city c is

$$w_{Hc} = X_{Hc} - (1-h)N_{Hc} + (1-h)K_{Hc} + \ln h \tag{8}$$

The labor market for unskilled workers is similar. I assume that there is an international capital market, and that capital is infinitely supplied at a given price i.[8]

Each worker consumes one unit of housing, so that demand for housing is determined by the number of skilled and unskilled workers in a city. Specifically, the the local demand for housing is the sum the demand of skilled workers and the demand of unskilled workers. For example, in city b:

$$r_b = \frac{(2s_H s_L)}{(s_H + s_L)} - \frac{(2s_H s_L)(N_{Hb} + N_{Lb})}{N(s_H + s_L)} - \frac{s_L(w_{Ha} - w_{Hb} - r_a)}{(s_L + s_H)} - \frac{s_H(w_{La} - w_{Lb} - r_a)}{(s_L + s_H)} \tag{9}$$

To close the model, I assume that the supply of housing is

$$r_c = z + k_c N_c \tag{10}$$

where $N_c = N_{Hc} + N_{Lc}$ is the number of housing units in city c, which is the same as the number of workers. The parameter k_c characterizes the elasticity of the supply of housing. I assume that this parameter is exogenously determined by geography and local land regulations. In cities where geography and regulations make it is easy to build new housing, k_c is small. In the extreme case where there are no constraints to building new houses, the supply curve is horizontal, and k_c is zero. In cities where geography and regulations make it difficult to build new housing, k_c is large. In the extreme case where it is impossible to build new houses, the supply curve is vertical, and k_c is infinite.[9]

In period 1, the two cities are assumed to be identical. Equilibrium in the labor market is obtained by equating equations 7 and 8 for each city. Equilibrium in the housing market is obtained by equating equations 9 and 10. I consider two scenarios for period 2. In the first scenario, the relative demand of skilled workers increases in one of the two cities (Section

[8] In equilibrium demand for capital is equal to its supply and marginal product of capital is the same for firms that use skill labor and those that use unskilled labor: $X_{Hc} - hK_{Hc} + hN_{Hc} + \ln(1-h) = \ln i X_{Lc} - hK_{Lc} + hN_{Lc} + \ln(1-h) = \ln i$.

[9] A limitation of equation 10 is housing production does not involve the use of any local input. Roback (1982) and Glaeser (2008), among others, discuss spatial equilibrium in the case where housing production involves the use of local labor and other local inputs. Moreover, equation 10 ignores the durability of housing. Glaeser and Gyourko (2001) point out that once built, the housing stock does not depreciate quickly and this introduces an asymmetry between positive and negative demand shocks. In particular, when demand declines, the quantity of housing cannot decline, at least in the short run.

5.2). In the second scenario, the relative supply of skilled workers increases in one of the two cities (Section 5.3). The implications of the two scenarios for the empirical analysis are summarized in Section 5.4.

5.2 Increase in the Relative Demand of Skilled Labor

Here I consider the case where the productivity of skilled workers increases relative to the productivity of unskilled workers in San Francisco. Nothing happens to the productivity of unskilled workers in San Francisco and the productivity of skilled and unskilled workers in Mobile. In other words, the relative demand for skilled labor increases in San Francisco. The amenities in the two cities are identical and fixed. Formally, I assume that in period 2, the productivity shifter for skilled workers in San Francisco is higher than in period 1: $X_{Hb2} = X_{Hb1} + \Delta$, where $\Delta > 0$ represents a positive, localized, skill-biased productivity shock. I have added subscripts 1 and 2 to denote periods 1 and 2. The dot-com boom experienced by the San Francisco Bay Area is arguably an example of such a localized skill biased shock. Driven by the advent of the Internet and the agglomeration of high tech firms in the area, the demand for skilled workers increased significantly (relative to the demand for unskilled workers) in San Francisco in the second half of the 1990s.[10]

Because skilled workers in San Francisco have become more productive, their nominal wage increases by an amount Δ/h, proportional to the productivity increase. Attracted by this higher productivity, some skilled workers leave Mobile and move to San Francisco. Following this inflow of skilled workers, the cost of housing in San Francisco increases by

$$r_{b2} - r_{b1} = \frac{s_L N k_b \Delta}{h(k_a N s_H + 2 s_H s_L + k_a N s_L + k_b N s_H + k_b N s_L)} \geq 0 \tag{11}$$

In Mobile, the cost of housing declines by the same amount because of out-migration. In San Francisco, real wages of skilled workers increase by

$$(w_{Hb2} - r_{b2}) - (w_{Hb1} - r_{b1}) = \frac{k_a N s_H + k_b N s_H + k_a N s_L + 2 s_H s_L}{h(k_a N s_H + 2 s_H s_L + k_a N s_L + k_b N s_H + k_b N s_L)} \Delta \geq 0 \tag{12}$$

It is easy to see that, because of the increased cost of housing, the increase in real wages is smaller than the increase in nominal wages Δ/h. Moreover, this increase in the real wage of skilled workers is larger the more elastic is housing supply in San Francisco (small k_b). Intuitively, a more elastic housing supply implies a smaller increase in housing prices in San Francisco, and therefore a larger increase in real wage, for a given increase in nominal wage. The increase in the real wage of skilled workers is also larger the smaller the elasticity of

[10]Beaudry, Doms and Lewis (2008) argue that over the past 30 years, technological change resulted in increases in the productivity of skilled workers in cities that already had many skilled workers. These cities also happen to be cities with a higher than average initial share of college graduates and cost of housing. See also Berry and Glaeser (2005).

local labor supply of skilled workers (large s_H). Intuitively, lower elasticity of labor supply implies less mobility. With less mobility, a larger fraction of the benefit of the productivity shocks is capitalized in real wages. In the extreme case of no mobility, ($s_H = \infty$), the entire productivity shock is capitalized in the real wage of skilled workers. The increase in the real wage of skilled workers is larger the larger the elasticity of local labor supply of skilled workers (small s_L). A higher elasticity of labor supply of unskilled workers implies that a larger number of unskilled workers move out in response to the inflow of skilled workers, so that the increase in housing costs is more limited.

In Mobile nominal wages don't change and housing costs decline, so that real wages for skilled workers increase by

$$(w_{Ha2} - r_{a2}) - (w_{Ha1} - r_{a1}) = \frac{s_L k_a N}{h(k_a N s_H + 2 s_H s_L + k_a N s_L + k_b N s_H + k_b N s_L)} \Delta \geq 0 \quad (13)$$

Although the shock has increased productivity only in one city, the equilibrium real wages of skilled workers increase in both cities because of mobility. By comparing equation 12 with 13, it is easy to see that the increase in real wages in the city directly affected by the productivity shock (San Francisco) is larger than the increase in real wages in the city not affected by the productivity shock (Mobile): $(w_{Hb2} - r_{b2}) - (w_{Hb1} - r_{b1}) \geq (w_{Ha2} - r_{a2}) - (w_{Ha1} - r_{a1})$. This is not surprising. While labor mobility causes real wages to increase in Mobile following a shock in San Francisco, real wages are not fully equalized because mobility is not perfect and only the marginal worker is indifferent between the two cities in equilibrium. With perfect mobility ($s_H = 0$), real wages are completely equalized.

What happens to the wage of unskilled workers? Because their productivity is fixed, their nominal wage does not change. However, housing costs increase in San Francisco and decline in Mobile. As a consequence, the real wage of unskilled workers in San Francisco decreases by

$$(w_{Lb2} - r_{b2}) - (w_{Lb1} - r_{b1}) = -\frac{s_L N k_b}{h(k_a N s_H + 2 s_H s_L + k_a N s_L + k_b N s_H + k_b N s_L)} \Delta \leq 0 \quad (14)$$

Effectively, unskilled workers compete for scarce housing with skilled workers, and the inflow of new skilled workers in San Francisco hurts inframarginal unskilled workers through higher housing costs. Marginal unskilled workers leave San Francisco, since their real wage is higher in Mobile. Inframarginal unskilled workers (those who have a strong preference for San Francisco over Mobile) opt to stay in San Francisco, even if their real wage is lower. For the same reason, the real wage and utility of inframarginal unskilled workers in Mobile increases:

$$(w_{La2} - r_{a2}) - (w_{La1} - r_{a1}) = \frac{s_L N k_a}{h(k_a N s_H + 2 s_H s_L + k_a N s_L + k_b N s_H + k_b N s_L)} \Delta \geq 0 \quad (15)$$

The equilibrium number of skilled workers increases in San Francisco, while the equilibrium number of unskilled workers decreases. Changes in employment in Mobile are exactly specular, by assumption. On net, the overall population of San Francisco increases because the number of skilled workers who move in is larger than the number of unskilled workers who leave.[11]

The productivity shock creates winners and losers. Skilled workers in both cities and landowners in San Francisco benefit from the productivity increase. Inframarginal unskilled workers in San Francisco are negatively affected, and inframarginal unskilled workers in Mobile are positively affected.[12] The exact magnitude of the changes in utility for skilled and unskilled workers and for landowners crucially depends on which of the three factors—skilled labor, unskilled labor or land—is supplied more elastically at the local level. Specifically, the incidence of the shock depends on the elasticities of labor supply of the two groups (which are governed by the preference parameters s_H and s_L) and the elasticities of housing supply in the two cities (which are governed by the parameters k_a and k_b). Moretti (forthcoming) provides detailed discussion of the incidence and welfare consequences of relative demand shocks.

The model also illustrates that a non-degenerate equilibrium is possible. After a shock that makes one group more productive, both groups are still represented in both cities. This conclusion hinges upon the assumption of a less than infinite elasticity of local labor supply.[13] Firms are indifferent between cities because they make the same profits in both cities. While labor is now more expensive in San Francisco, it is also more productive there. Because firms produce a good that is internationally traded, if skilled workers weren't more productive, employers would leave San Francisco and relocate to Mobile.[14]

[11] In particular, the number of skilled workers in San Francisco increases by

$$N_{Hb2} - N_{Hb1} = \frac{\Delta N((k_a + k_b)N + 2s_L)}{2h(k_a N(s_H + s_L) + k_b N(s_H + s_L) + 2s_H s_L)} \geq 0 \tag{16}$$

The number of unskilled workers declines by

$$N_{Lb2} - N_{Lb1} = -\frac{N^2(k_a + k_b)}{2h(k_a N s_H + 2s_H s_L + k_a N s_L + k_b N s_H + k_b N s_L)}\Delta \leq 0 \tag{17}$$

San Francisco population increases by

$$(N_{Hb2} + N_{Lb2}) - (N_{Hb1} + N_{Lb1}) = \frac{\Delta N s_L}{h(k_a N(s_H + s_L) + k_b N(s_H + s_L) + 2s_H s_L)} \geq 0 \tag{18}$$

[12] Although inframarginal unskilled workers in San Francisco are made worse off by the decline in their real wage, they are still better off in San Francisco than in Mobile because of their preference for San Francisco.

[13] In the absence of individual preferences for location, no unskilled worker would remain in San Francisco and the equilibrium would be characterized by complete geographic segregation of workers by skill level. This is not realistic, since in reality we never observe cities that are populated by workers of only one type.

[14] An assumption of this model is that skilled and unskilled workers are employed by different firms, so that the labor market is segregated by skill within a city. This assumption effectively rules out imperfect substitutability between skilled and unskilled labor. In a more general setting, skilled and unskilled workers

5.3 Increase in the Relative Supply of Skilled Labor

In the case of demand pull described above, the number of skilled workers in San Francisco increases because the relative demand of skilled workers increases. I now turn to the opposite case, where the number of skilled workers in San Francisco increases because the relative supply of skilled workers in San Francisco increases.

Specifically, I consider what happens when San Francisco becomes relatively more desirable for skilled workers compared to Mobile. I assume that in period 2, the amenity level increases for skilled workers in San Francisco: $A_{Hb2} = A_{Hb1} + \Delta'$, where $\Delta' > 0$ represents the improvement in the amenity. I assume that the productivity of both skilled and unskilled workers, as well as the amenity level in Mobile, do not change.[15]

Unlike the case of demand, here the nominal wage of skilled workers in San Francisco and Mobile remains unchanged.[16] Attracted by the better amenity, some skilled workers move from Mobile to San Francisco and some unskilled workers leave San Francisco to Mobile.[17] On net, the population in San Francisco increases by

$$(N_{Hb2} + N_{Lb2}) - (N_{Hb1} + N_{Lb1}) = \frac{\Delta' N s_L}{h(k_a N(s_H + s_L) + k_b N(s_H + s_L) + 2 s_H s_L)} \geq 0 \quad (19)$$

As a consequence, housing costs in San Francisco increase by

$$r_{b2} - r_{b1} = \frac{s_L N k_b \Delta'}{h(k_a N s_H + 2 s_H s_L + k_a N s_L + k_b N s_H + k_b N s_L)} \geq 0 \quad (20)$$

work in the same firm. The qualitative results generalize, but the equilibrium depends on the degree of imperfect substitution between skilled and unskilled labor. Specifically, complementarity between skilled and unskilled workers implies that the marginal product of unskilled workers increases in the number of skilled workers in the same firm. Thus, the inflow of skilled workers in city b caused by the increase in their productivity endogenously raises the productivity of unskilled workers in city b. As a consequence, the real wage of unskilled workers declines less than in the case described above. This mitigates the negative effect on the welfare of unskilled workers in city b and it reduces the number of unskilled workers who leave the city.

[15]For simplicity, I have assumed that supply shocks are driven by increases in amenities for given tastes. Glaeser and Tobio (2007) have a model that makes a similar assumption. Alternatively I could assume that (i) amenities are fixed, but the taste for those amenities increase; or (ii) both amenities and tastes are fixed, but amenities are a normal good so that college graduates consume more of them than high school graduates (Gyourko, Mayer, and Sinai, 2006).

[16]This may be surprising at first. While one might expect wage *increases* in response to demand increases (indeed, this is what happens in subsection 5.2), one might expect wage *decreases* in response to supply increases. Why nominal wages do not decline in San Francisco? The reason is that in a model with capital, nominal wages do not move in San Francisco because capital flows to San Francisco and leaves Mobile, offsetting the changes in labor supply in the two cities. (In a model without capital nominal wages do decline.)

[17]Specifically, the number of skilled workers who move to San Francisco is equal to $\frac{\Delta' N((k_a+k_b)N+2s_L)}{2h(k_a N(s_H+s_L)+k_b N(s_H+s_L)+2s_H s_L)} \geq 0$. The number of unskilled workers who move to Mobile is equal to $\frac{\Delta' N^2(k_a+k_b)}{2h(k_a N(s_H+s_L)+k_b N(s_H+s_L)+2s_H s_L)} \geq 0$.

and decline in Mobile by

$$r_{a2} - r_{a1} = -\frac{s_L N k_a \Delta'}{h(k_a N s_H + 2 s_H s_L + k_a N s_L + k_b N s_H + k_b N s_L)} \leq 0 \tag{21}$$

Real wages of skilled workers in San Francisco decline by an amount equal to equation 20 (with a minus sign in front). This reflects the compensating differential for the better amenity in San Francisco. Real wages of skilled workers in Mobile increase by an amount equal to equation 21 (with a minus sign in front).

Similarly, the real wage for unskilled workers in San Francisco declines by

$$(w_{Lb2} - r_{b2}) - (w_{Lb1} - r_{b1}) = -\frac{s_L N k_b}{h(k_a N s_H + 2 s_H s_L + k_a N s_L + k_b N s_H + k_b N s_L)} \Delta' \leq 0 \tag{22}$$

and it increases in Mobile.

Like for the case of demand shocks, a supply shock generates winners and losers. Here inframarginal skilled workers benefit from the improvement in amenities. While the utility gain is larger for inframarginal skilled workers in San Francisco, inframarginal skilled workers in Mobile are also made better off, even if there is no change in amenity there. On the other hand, inframarginal unskilled workers in San Francisco are made worse off by the increase in housing prices. Similarly, inframarginal unskilled workers in Mobile are made better off by the decline in local housing prices.

5.4 Implications for Inequality in Wages and Utility

The model has three implications that are useful in guiding the interpretation of the empirical findings.

(A) First, the model clarifies the relationship between changes in relative wages and changes in relative utility in the two scenarios. The analysis in Sections 5.2 and 5.3 suggests that *for a given nation-wide increase in the nominal wage gap between skilled and unskilled workers*, the demand pull hypothesis implies a more limited increase in utility inequality, while the supply push hypothesis implies a larger increase in utility inequality.

More specifically, in the demand pull scenario the nominal wage difference between skilled and unskilled workers averaged across the two cities increases.[18] The utility difference between skilled and unskilled workers averaged across the two cities also increases, but by an amount smaller than the increase in the nominal wage gap. It is possible to show that the larger is the increase in housing costs experienced by skilled workers relative to unskilled workers, the smaller is the increase in average utility experienced by skilled workers relative to unskilled workers.[19]

[18] This average is a weighted average reflecting the size of the two cities.

[19] To formally see this, consider the population-weighted average across the two cities of the change in the skilled-unskilled nominal wage difference and compare it with the population-weighted average across the

The intuition is simple. The benefits of a higher nominal wage for skilled workers are in part eroded by the higher cost of housing in the cities where the new skilled jobs are created. Thus, the relative utility of skilled workers does not increase as much as their relative nominal wage. Put differently, if college graduates move to expensive cities like San Francisco and New York because of increases in the relative demand for college graduates in these cities—and not because they particularly like living in San Francisco and New York—then part of the benefit of higher nominal wage is offset by the higher cost of living. In this case, the increase in their real wage and utility level is smaller than the increase in their nominal wage.

By contrast, in the supply push scenario, the utility difference between skilled and unskilled workers averaged across the two cities increases more than the nominal and real wage difference between skilled and unskilled workers averaged across the two cities. Intuitively, if college graduates move to expensive cities like San Francisco and New York because improvements in amenities raise the relative supply of college graduates there—and not because of labor demand—then there may still be a significant increase in utility inequality even if the increase in real wage inequality is limited. In this case, increases in the cost of living in these cities simply reflect the increased attractiveness of these cities to skilled workers and represent the price to pay for the consumption of desirable amenities.[20]

(B) Second, the equilibrium described in subsections 5.2 and 5.3 suggests a simple empirical test to distinguish between the two cases. If relative demand shifts are responsible for the geographical reallocation of labor, we should see that in equilibrium, cities that experience large increases in the relative number of skilled workers (in the model: San Francisco)

two cities of the change in the skilled-unskilled utility difference. In the simple case where $k_a = k_b = k$, the difference between the two is

$$\frac{Nk\Delta^2 s_L(s_L + 2kN)}{2h^2(kNs_H + s_H s_L + kNs_L)^2} \geq 0 \tag{23}$$

which is non-negative, indicating that the relative nominal wage of skilled workers grows more than their relative utility. In the more general case where $k_a \neq k_b$, the difference between the two remains positive as long as the elasticity of housing supply in the city affected by the demand shock is not too large compared with the elasticity of housing supply in the city not directly affected by the demand shock.

[20] To formally see this, note that the simple case where $k_a = k_b = k$, the population-weighted average change in the skilled-unskilled nominal wage difference minus the population-weighted change in the skilled-unskilled utility difference is equal to

$$\frac{-\Delta'(-kNs_L\Delta' + 2kNs_H s_L + 2kNs_H s_L + s_H s_L + s_L\Delta' s_H + k^2N^2s_L^2 + k^2N^2s_H^2 + k^2N^2s_H\Delta' + 2kNs_H^2 s_L + 2Nk\Delta' s_H s_L}{(2(kNs_H + s_H s_L + kNs_L)^2)} \tag{24}$$

which is non-positive unless the elasticity of local labor supply of skilled workers is too small compared with the elasticity of local labor supply of unskilled workers. In the more general case where $k_a \neq k_b$, the expression is considerably more complicated, but the difference remains non-positive unless the elasticity of local labor supply of skilled workers is too small.

also experience increases in the relative nominal wage of skilled workers, compared to cities that experience small increases (or declines) in the relative number of skilled workers (in the model: Mobile). By contrast, if relative supply shifts are responsible for the geographical reallocation of labor, we should see that in equilibrium, cities that experience an increase in the relative number of skilled workers experience no change in the relative nominal wage of skilled workers.

One might have expected that an increase in the relative supply of factor of production in a city should cause a *decline* in its equilibrium relative price. Why in the model the nominal wage of skilled workers in San Francisco remains constant following an increase in the relative supply of skilled workers? As discussed in Section 5.3, this is due to the endogenous reaction of capital. Because capital is supplied with infinite elasticity at a fixed interest rate, nominal wages do not move in San Francisco because capital flows to San Francisco and leaves Mobile, thus offsetting the effect of changes in labor supply in the two cities. In a model without capital, nominal wages of skilled workers decline in San Francisco following an increase in their supply.

(C) Finally, it is important to point out that, while the focus of the paper is on inequality related to labor market outcomes, the broader welfare consequences of the demand and supply shocks depend not just on changes in relative wages, but also on which of the two education groups originally owns the land in the cities that benefit from the demand and supply shocks. In the model, some landowners benefit from the demand and supply shocks (namely those in San Francisco), while other are hurt (namely those in Mobile). The relevant empirical question in this respect is which of the two skill groups owns more of the land in the neighborhoods that whose land prices are raised by the inflow of new residents in cities that experience positive skill-biased shocks and the neighborhoods that are abandoned by the outflow of residents is cities that experience negative shocks. This is an important but complicated question. A full empirical treatment of this issue is beyond the scope of this paper and is left for future research.

6 Interpreting the Evidence: Demand Pull or Supply Push?

I now present empirical evidence that seeks to determine whether relative demand or relative supply shifts—or a combination of the two—drive changes in the geographical location of different skill groups. The analysis in subsections 5.2 and 5.3 suggests that the demand pull and the supply push hypotheses have similar predictions for equilibrium housing costs: under both hypotheses, cities that experience large increases in the share of college graduates should also experience large increases in housing costs.

But the demand pull and supply push hypotheses have different predictions for wage changes. Under the demand pull hypothesis, cities that experience large increases in the share of college graduates should experience large increases in the equilibrium relative wage of college graduates. By contrast, under the supply push hypothesis, there should be no positive relationship between increases in the share of college graduates and changes in the equilibrium relative nominal wages. (See Section 5.4, part B.) Intuitively, increases in the relative demand of a factor of production in a city should result in increases in its equilibrium relative price there. Increases in the relative supply of factor of production in a city can not cause an increase in its equilibrium relative price. A similar idea is used in Katz and Murphy (1992) to explain *nationwide* changes in relative wages.[21]

It is important to highlight that the two hypotheses are not mutually exclusive since it is possible that cities experience both demand and supply shocks. It is also possible that relative demand shifts endogenously generate relative supply shifts, and vice versa. For example, an increase in the relative demand for skilled labor in a city may result in an increase in the number of college educated residents in that city and this in turns may result in increases in the local amenities that are attractive to college graduates, such as good schools, good theaters, good restaurants, etc. Alternatively, an increase in the supply of skilled workers in a city may generate agglomeration spillovers that lead to increases in the productivity of firms and workers in that city (Moretti 2004a, 2004b).

I present two pieces of empirical evidence. First, I look at the OLS relationship between changes in the college share and changes in the college premium across US metropolitan areas. The finding of a positive coefficient indicates that relative demand shifts are important, but does not rule out the existence of relative supply shifts. Second, to shed more light on whether relative supply shifts are important, I use an instrumental variable strategy.

(1) First, in Figure 2, I show the empirical relationship between the equilibrium college share and the equilibrium college premium across US metropolitan areas, both in the 2000 cross-section and in changes between 1980 and 2000. Demand pull would predict a positive slope, while supply push would predict zero slope. Note that that the relationship in the Figure is *not causal*. Rather, it is an *equilibrium* relationship between relative number of college graduates and their relative wage. This is in contrast with earlier work, including my own, that seeks to establish the *causal* effect of increases in college share on wages and therefore estimate different specifications.[22]

[21] In the literature, there are several existing measures of quality-of-life differences across metropolitan areas (see, for example, Chen and Rosenthal, 2008; and Albouy, 2009). One may be tempted to use these measures to provide an additional empirical test of my conclusions by estimating the relationship between these existing measures of local amenities and the share of college graduates across cities. However, this type of test would be difficult to implement in practice. As explained above, what matters in my framework is the (change over time in) the difference in amenities that attract skilled and unskilled workers. By contrast, the existing measures of local amenities are typically city-wide, and do not differentiate between skilled and unskilled workers.

[22] For example, in Moretti (2004), I try to establish the causal effect of increases in college share on wages.

The Figure shows a positive association between the college share and the college premium across US metropolitan areas, both in levels as well as in changes. Columns 1 and 2 in Table 4 quantify the corresponding regression coefficients. The level of observation is the metropolitan area. The dependent variable is the city-specific college premium, defined as the city-specific difference in the log of hourly wage for college graduates and high school graduates conditional on all the controls used in the regressions (a cubic in potential experience, year effects, gender and race). Models are weighted by city size. The coefficient for the specification in column 2 is positive and statistically significant: .388 (.057).

This evidence is consistent with demand factors playing a significant role in driving variation in college share across cities. This conclusion is consistent with Berry and Glaeser (2005), who argue that demand factors play a more important role than supply factors in explaining the sorting of skilled workers across US metropolitan areas.

(2) The evidence in Figure 2 and Table 4 suggests that demand factors are important, but does not rule out that supply factors are also present. As a second piece of evidence that may shed more light on whether relative supply factors play any role in driving variation in college share across cities, I use observable shocks to the relative demand of skilled labor as an instrumental variable for college share.

This IV estimate isolates the effect on the college premium of changes in the college share that are driven exclusively by changes in relative demand. Put differently, the instrumental variable estimate establishes what happens to the college premium in a city when the city experiences an increase in the number of college graduates that is driven purely by an increase in the relative demand for college graduates. By contrast, the OLS estimate above establishes what happens to the college premium in a city when the city experiences an increase in the number of college graduates that may be driven by either demand or supply shocks. The comparison of the two estimates is therefore informative about the relative importance of demand and supply shocks.

To isolate relative demand shocks, I use as an instrument the weighted average of nationwide relative employment growth by industry, with weights reflecting the city-specific employment share in those industries:

$$\text{Change in Relative Demand in City c} = \sum_s \eta_{sc}(\Delta E_{Hs} - \Delta E_{Ls}) \tag{25}$$

where η_{sc} is the share of jobs in industry s in city c in 1980; ΔE_{Hs} is the nationwide change

The econometric specification adopted here differs from the specification there, because in Moretti (2004) the econometric model seeks to *control for* shocks to the relative demand of skilled labor. To this end, I include in the regressions as controls several variables in order to absorb changes in the relative demand for college graduates. I also use instrumental variables to further control for relative demand shocks. By contrast, in this paper, I engage in a completely different exercise. I do not seek to hold constant demand shocks. Instead, I am interested in establishing the role played by demand shocks in affecting changes in college share across cities. What I am measuring in Figure 2 and Table 4 is the relationship between the wage gap and the college share, inclusive of any human capital spillover.

between 1980 and 2000 in the log of number of jobs for college graduates in industry s (excluding city c); ΔE_{Ls} is a similar change for high school graduates. If relative employment of skilled workers in a given industry increases (decreases) nationally, cities where that industry employs a significant share of the labor force will experience a positive (negative) relative shock to the labor demand of skilled workers (Katz and Murphy, 1992).

The first stage relationship between demand shocks and changes in college share is shown graphically in Figure 3. The figure shows that in cities that experience an increase in the relative demand of college graduates the share of college graduates increases and the relationship appears fairly tight. The regression coefficient is .42(.02), with R^2 of .44. This is interesting because it means that this measure of demand shocks alone accounts for almost half of the variation in the geographical location of different skill groups. Since there are other demand shocks that are not captured by the instrument, this lends indirect support to the notion that demand shocks play an important role.

The instrumental variable estimate, in column 3 of Table 4, is .371 (.106) and is remarkably close to the OLS estimate. The similarity between the OLS and the IV estimates suggests that the increase in the college premium in a city caused by a demand shock (IV estimate in column 3) is not very different from the empirical correlation between the college share and the college premium that is observed in the data (OLS estimate in column 2). In other words, most of the empirical correlation between the college share and the college premium that is observed in the data seems to be driven by demand shocks.

7 Conclusions

Because of their different geographical distribution, US households are exposed to different levels and trends in cost of living. One contribution of this paper is to document that, as a consequence, the conditional difference between the wage of skilled workers and of unskilled workers is significantly lower in real terms than in nominal terms and has grown less. In 2000, the level of the college premium is 60% in nominal terms and only 51% in real terms. More importantly, the *increase* in the college premium between 1980 and 2000 in real terms is significantly smaller than the increase in nominal terms. Specifically, at least 22% of the documented increase in the college premium between 1980 and 2000 is accounted for by differences in the cost of living.

The implications of this empirical finding for disparities in well-being depend on the reasons for the increase in the share of college graduates in expensive cities. Using a simple general equilibrium model of the labor and housing markets, I consider two broad classes of explanations. Under a demand pull hypothesis, the relative demand of college graduates increases in expensive cities because of localized skill-biased technical change or other demand shocks. In this case, college graduates move to expensive cities because the jobs for college graduates are increasingly located in those cities, and not because they particularly like living in those cities. The increase in their utility level is smaller than the increase in their nominal wage due to higher cost of living. Under a supply push hypothesis, the relative supply of college graduates increases in expensive cities because college graduates are increasingly attracted by amenities located in those cities. The increase in the cost of living in those cities reflects the attractiveness of the cities to skilled workers and is the price for the consumption of desirable amenities. In this case, there may still be a significant increase in utility inequality even if the increase in real wage inequality is limited. Of course, the two hypotheses are not mutually exclusive and it is possible that cities experience both demand and supply shocks.

To determine whether the variation in the relative number of college graduates across cities is driven by relative demand or relative supply shocks, I analyze the equilibrium relationship between changes in college premium and changes in the share of college graduates across metropolitan areas. Consistent with demand shocks playing an important role, I find a positive association between changes in college premium and changes in college share: cities that experience large increases in the fraction of college graduates also experience large increases in the relative wage of college graduates. I also present an instrumental variable estimate obtained by instrumenting changes in college share with a measure of arguably exogenous relative demand shocks.

The weight of the evidence seems consistent with the notion that changes in the geographical location of different skill groups are mostly driven by changes in their relative demand. I conclude that the increase in well-being disparities between 1980 and 2000 is significantly smaller than we previously thought based on the existing literature.

8 Bibliography

Aguiar, Mark and Erik Hurst "The Increase in Leisure Inequality", mimeo (2007a).

Aguiar, Mark and Erik Hurst, "Measuring Trends in Leisure: The Allocation of Time Over Five Decades", Quarterly Journal of Economics (2007b).

Bayer, Patrick, Robert McMillan, and Kim Rueben "An Equilibrium Model of Sorting in an Urban Housing Market" NBER Working Paper 10865

Bayer, Patrick, Shakeeb Khan, Christopher Timmins "Nonparametric Identification and Estimation in a Generalized Roy Model", NBER Working Paper 13949

Bayer, Patrick, and Christopher Timmins "Estimating Equilibrium Models of Sorting Across Locations" Economic Journal, March 2007

Bayer, Patrick, and Robert McMillan "Racial Sorting and Neighborhood Quality," NBER Working Paper 11813

Basker, Emek "Selling a Cheaper Mousetrap: Wal-Mart's Effect on Retail Prices", JUE 2005, which uses the average prices of drugstore items in selected cities for 1982-2002 and

Baum-Snow Nathaniel and Ronni Pavan Inequality and City Size Brown University, mimeo, 2009.

Beaudry, Paul, Mark Doms and Ethan Lewis "Should the PC be Considered a Technological Revolution? City Level Evidence from 1980-2000", mimeo (2008)

Beeson, Patricia E., 1991. "Amenities and regional differences in returns to worker characteristics," Journal of Urban Economics, Elsevier, vol. 30(2), pages 224-241, September.

Black, Dan, Natalia Kolesnikova, and Lowell J. Taylor, "Earnings Functions When Wages and Prices Vary by Location", Journal of Labor Economics, 2009.

Black, Dan, Natalia Kolesnikova, and Lowell J. Taylor, Why do so Few Women Work in New York (and so many in Minneapolis)? Labor Supply of Married Women across U.S. Cities, mimeo, University of Chicago (2010)

Black, Dan, Natalia Kolesnikova, and Lowell J. Taylor, Local Price Variation and Labor Supply Regional Economic Development 4(1) October 2008 2-14.

Black, Dan, Natalia Kolesnikova, and Lowell J. Taylor, S. Sanders, and M. Wessel A Divergent View on Black-White Earnings Convergence, Mimeo, University of Chicago (2009).

Blanchard Olivier Jean and Lawrence F. Katz, 1992. "Regional Evolutions," Brookings Papers on Economic Activity, Economic Studies Program, The Brookings Institution, vol. 23(1992-1), pages 1-76.

Blundell, R, L. Pistaferri and I. Preston, "Consumption inequality and partial insurance" American Economic Review 2008

Bound, John and Holzer, Harry J, 2000. “Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s,” Journal of Labor Economics, University of Chicago Press, vol. 18(1), pages 20-54, January.

Buera, Francisco J. and Joseph P. Kaboski ”The Rise of the Service Economy”, NBER Working Paper No. 14822 (2009)

Bureau of Labor Statistics, “Handbook of Methods”, June 2007.

Chen, Tong and Stuart S. Rosenthal “Local Amenities and Migration Over the Life Cycle: Do People Move for Jobs or Fun?” Journal of Urban Economics (2008).

Dora L. Costa and Matthew E. Kahn, 2000. ”Power Couples: Changes In The Locational Choice Of The College Educated, 1940-1990,” The Quarterly Journal of Economics, MIT Press, vol. 115(4), pages 1287-1315, November.

Dahl, Gordon “Mobility and the Return to Education: Testing a Roy Model with Multiple Markets,” Econometrica, Vol. 70, No. 6, pp. 2367-2420, (2002).

Duranton, Gilles ”Spatial wage disparities: Sorting matters!” Journal of Urban Economics, 2008, 63(2), 723-742

Duranton, Gilles ” Mind the Gaps: The Evolution of Regional Inequalities in the U.K. 1982-1997” Journal of Regional Science, 2002, 42(2), 219-256

Duranton, Gilles ”Urban evolutions” American Economic Review, 2007, 97 (1), 197-221

Duranton, Gilles, Kristian Behrens and Frdric Robert-Nicoud ”Productive cities: Sorting, selection, and agglomeration”, mimeo, University of Toronto (2010).

Glaeser, E, Gyourko, J. and R. Saks (2005) Why Have Housing Prices Gone Up? American Economic Review, forthcoming

Glaeser, E. and C. Berry, “The Divergence of Human Capital Levels across Cities” Harvard Institute of Economic Research Discussion Paper Number 2091 (2005)

Glaeser, Edward L and Mare, David C, 2001. ”Cities and Skills,” Journal of Labor Economics, University of Chicago Press, vol. 19(2), pages 316-42, April.

Goldin Claudia and Lawrence F. Katz, 2007. “Long-Run Changes in the U.S. Wage Structure: Narrowing, Widening, Polarizing,” NBER Working Papers 13568, National Bureau of Economic Research, Inc

Gordon, Robert and Ian Dew-Beker “Controversies About the Rise of American Inequality: A Survey”, NBER WP 13982, 2008.

Gordon, Robert and Ian Dew-Beker ”Selected Issues in American Inequality”. Brookings Papers on Economic Activity, 2007, no. 2, pp. 169-92.

Greenstone M., R. Hornbeck and E. Moretti ”Identifying Agglomeration Spillovers: Evidence from Million Dollar Plants",Journal of Political Economy (forthcoming).

Hamermesh, Dan "Demand for Labor," in International Encyclopedia of the Social and Beh avioral Sciences, Pergamon Press, 2001.

Hamermesh, Dan "Changing Inequality in Markets for Workplace Amenities," Quarterly Journal of Economics, November 1999.

Hamermesh, Dan "A General Model of Dynamic Labor Demand," Review of Economics and Stat istics, November 1992.

Heckman James, Lance Lochner and Christopher Taber, 1998. "Explaining Rising Wage Inequality: Explanations With A Dynamic General Equilibrium Model of Labor Earnings With Heterogeneous Agents," Review of Economic Dynamics, vol. 1(1), pages 1-58, January.

Jappelli, T. and L. Pistaferri, "Consumption and Income Inequality" (2008)

Kahn Matthew E., 1995. "A Revealed Preference Approach to Ranking City Quality of Life," Journal of Urban Economics, Elsevier, vol. 38(2), pages 221-235, September.

Kahn, Matthew E. 2005. "Estimating Housing Demand With an Application to Explaining Racial Segregation in Cities," Journal of Business and Economic Statistics, American Statistical Association, vol. 23, pages 20-33, January.

Kahn, Matthew E., 1999. "Climate consumption and climate pricing from 1940 to 1990," Regional Science and Urban Economics, Elsevier, vol. 29(4), pages 519-539, July.

Katz, Lawrence F and Murphy, Kevin M, 1992. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," The Quarterly Journal of Economics, MIT Press, vol. 107(1), pages 35-78, February.

Katz L. and D. Autor "Changes in the Wage Structure and Earnings Inequality", Handbook of Labor Economics, Ashenfelter and Crad, Eds, Elsevier (1999).

Krupka Douglas J. and Kwame Donaldson, "Wages, Rents and Heterogeneous Moving Costs," IZA Discussion Papers 3224, Institute for the Study of Labor (2007)

Lewbel, Arthur and Krishna Pendakur, 2008, "Tricks with Hicks: The EASI Implicit Marshallian Demand System for Unobserved Heterogeneity and Flexible Engel Curves.", American Economic Review, forthcoming.

Krueger, Perri, Pistaferri, Violante "Cross Sectional Facts for Macroeconomists" Review of Economic Dynamics 2010.

Malpezzi, Stephen, Gregory Chun and Richard Green (1998) "New Place-to-Place Housing Price Indexes for US Metro Areas", Real Estate Economics, 26.

Moretti, Enrico "Workers' Education, Spillovers and Productivity: Evidence from Plant-Level Production Functions" American Economic Review 94(3) (2004b)

Moretti, Enrico “Local Labor Markets” in Ashenfelter and Card Eds. ”Handbook of labor Economics (forthcoming)

Roback, Jennifer “Wages, rents and the quality of Life”, Journal of Political Economy, 90-6, pp. 1257-1278 (1982)

Roback, Jennifer ”Wages, rents, and amenities: differences among workers and regions” Economic Inquiry, (1988)

Pendakur, Krishna, 2002. ”Taking prices seriously in the measurement of inequality,” Journal of Public Economics, Elsevier, vol. 86(1), pages 47-69, October

Pendakur, Krishna, 1998. ”Changes in Canadian Family Income and Family Consumption Inequality between 1978 and 1992,” Review of Income and Wealth, Blackwell Publishing, vol. 44(2), pages 259-83, June.

Pendakur, Krishna, 2009. ”Testing and imposing Slutsky symmetry in nonparametric demand systems,” Journal of Econometrics, Elsevier, vol. 153(1), pages 33-50, November.

Pendakur, Krishna 2005. ”Semiparametric estimation of lifetime equivalence scales,” Journal of Applied Econometrics, John Wiley and Sons, Ltd., vol. 20(4), pages 487-507

Pierce, Brooks ”Compensation Inequality”, Quarterly Journal of Economics (2001)

Poole R., F. Ptacek and R. Verbrugge “The treatment of Owner-Occupied Housing in the CPI” (2006), Office of Price and Living Conditions, Bureau of Labor Statistics.

Shapiro, Jesse, ”Smart cities: Quality of life, productivity, and the growth effects of human capital” Review of Economics and Statistics, May 2006

Table 1

Metropolitan Area	Nominal Income	Real Income	Percent Difference
San Jose, CA	54405.32	39187.88	-0.279705
Stamford, CT	79374.09	60490.73	-0.237903
San Francisco-Oakland-Vallejo, CA	48579.09	38189.25	-0.213875
Santa Cruz, CA	42374.27	34700.63	-0.181092
Danbury, CT	58414.7	48454.08	-0.170516
Santa Barbara-Santa Maria-Lompoc, CA	38030.82	32626.27	-0.14211
Ventura-Oxnard-Simi Valley, CA	40325.79	34930.42	-0.133795
Boston, MA-NH	46411.32	40253.07	-0.132689
Honolulu, HI	33805.31	29322.25	-0.132614
Santa Rosa-Petaluma, CA	40565.8	35323.53	-0.129229
Salinas-Sea Side-Monterey, CA	32324.44	28257.47	-0.125817
New York-Northeastern NJ	42989.14	37646.13	-0.124288
Washington, DC/MD/VA	48057.21	42142.48	-0.123077
Los Angeles-Long Beach, CA	35780.87	31614.55	-0.11644
Monmouth-Ocean, NJ	44945.88	40199.32	-0.105606
San Diego, CA	37271.78	33352.23	-0.105161
Seattle-Everett, WA	43504.58	39010.95	-0.103291
Nashua, NH	41359.91	37343.74	-0.097103
Trenton, NJ	47288.35	42817.15	-0.094552
Bridgeport, CT	47589.96	43161.88	-0.093046
Fort Lauderdale-Hollywood-Pompano Beach, FL	36237.65	32922.89	-0.091473
Austin, TX	41269.63	37975.6	-0.079817
Naples, FL	38590.7	35576.77	-0.0781
Anchorage, AK	40751.3	37590.96	-0.077552
Newburgh-Middletown, NY	38105.78	35186.28	-0.076616
Atlanta, GA	39956.47	36913.66	-0.076153
Barnstable-Yarmouth, MA	38719.36	35817.87	-0.074937
West Palm Beach-Boca Raton-Delray Beach, FL	39084.94	36189.78	-0.074074
San Luis Obispo-Atascad-P Robles, CA	36011.36	33478.59	-0.070333
Denver-Boulder, CO	40966.75	38117.18	-0.069558
Ann Arbor, MI	45702.14	42652.77	-0.066723
Miami-Hialeah, FL	29531.75	27644.46	-0.063907
Chicago, IL	41064.57	38514.93	-0.062088
New Haven-Meriden, CT	39954.43	37641.28	-0.057895
Santa Fe, NM	36554.24	34458.73	-0.057326
Las Vegas, NV	32834.18	30964.15	-0.056954
Dutchess Co., NY	39612.11	37361.38	-0.056819
Reno, NV	36923.75	34849.31	-0.056182
Orlando, FL	33899.67	32215.98	-0.049667
Fort Collins-Loveland, CO	38589.2	36674.33	-0.049622
Philadelphia, PA/NJ	39752.9	37826.65	-0.048456
Portland, OR-WA	37142.06	35433.87	-0.045991
Brockton, MA	35189.11	33642.55	-0.04395
Minneapolis-St. Paul, MN	42254	40526.07	-0.040894
Sarasota, FL	34270.81	32876.42	-0.040687
Atlantic City, NJ	33758.89	32426.39	-0.039471

Manchester, NH	32247.5	31020.46	-0.038051
Wilmington, DE/NJ/MD	41096.62	39570.11	-0.037144
Raleigh-Durham, NC	39050.45	37648.44	-0.035902
Dallas-Fort Worth, TX	38891.01	37567.47	-0.034032
Sacramento, CA	36571.38	35380.04	-0.032576
Hartford-Bristol-Middleton- New Britain, CT	42644.75	41292.27	-0.031715
Phoenix, AZ	35694.89	34648.93	-0.029303
Madison, WI	40291.4	39190.29	-0.027329
Colorado Springs, CO	35250.73	34354.23	-0.025432
Baltimore, MD	39244.79	38340.1	-0.023052
Yolo, CA	34478.04	33687.54	-0.022928
Riverside-San Bernadino, CA	30286.15	29643.71	-0.021212
Bremerton, WA	35006.92	34267.89	-0.021111
Portland, ME	38112.26	37377.45	-0.01928
Fort Myers-Cape Coral, FL	32143.15	31544.09	-0.018637
Bellingham, WA	31298.4	30852.01	-0.014262
Tampa-St. Petersburg-Clearwater, FL	33340.45	32867.79	-0.014177
Olympia, WA	34837.69	34405.98	-0.012392
Salt Lake City-Ogden, UT	34928.59	34562.71	-0.010475
Tacoma, WA	34070.5	33771.25	-0.008783
Fort Pierce, FL	32310.43	32037.19	-0.008457
Detroit, MI	39068.79	38747.12	-0.008234
Kenosha, WI	34721.79	34483.32	-0.006868
Albany-Schenectady-Troy, NY	35908.91	35691.44	-0.006056
Worcester, MA	35944.34	35751.28	-0.005371
Vineland-Milville-Bridgetown, NJ	28310.53	28169.08	-0.004997
Charlotte-Gastonia-Rock Hill, NC-SC	36711.14	36581.58	-0.003529
Jacksonville, FL	33719.22	33675.38	-0.0013
Stockton, CA	30235.71	30226.49	-0.000305
Houston-Brazoria, TX	36133.14	36124.92	-0.000228
Richmond-Petersburg, VA	37188.42	37242.79	0.001462
Savannah, GA	33286.53	33375.5	0.002673
Nashville, TN	35934.79	36039.38	0.002911
Charlottesville, VA	36498.48	36637.57	0.003811
Rochester, NY	34722.74	34873.8	0.00435
Galveston-Texas City, TX	35582.36	35737.96	0.004373
Charleston-N.Charleston,SC	32779.3	32924.33	0.004425
Modesto, CA	29215.99	29385.18	0.005791
Norfolk-VA Beach--Newport News, VA	32256.25	32504.74	0.007704
Waterbury, CT	27905.45	28152.26	0.008845
Milwaukee, WI	38241.19	38599.76	0.009377
Daytona Beach, FL	28399.17	28687.39	0.010149
Melbourne-Titusville-Cocoa-Palm Bay, FL	31699.74	32026.86	0.010319
Columbus, OH	37124.03	37563	0.011825
Provo-Orem, UT	32740.22	33157.5	0.012745
Punta Gorda, FL	29018.52	29425.38	0.014021
Des Moines, IA	37387.23	37915.51	0.01413
Iowa City, IA	40470.77	41111.8	0.015839
Eugene-Springfield, OR	29887.08	30362.39	0.015903
Wilmington, NC	31289.38	31789.55	0.015985
Rochester, MN	39944.34	40612.47	0.016727

Kansas City, MO-KS	36585.43	37217.44	0.017275
Indianapolis, IN	36740.44	37417.32	0.018423
Salem, OR	30279.69	30878.07	0.019762
Fort Walton Beach, FL	30826.94	31459.54	0.020521
Bloomington, IN	35287.22	36057.2	0.02182
Albuquerque, NM	31368.05	32077.96	0.022632
Racine, WI	37175.47	38017.38	0.022647
Akron, OH	34589.23	35381.11	0.022894
Medford, OR	29551.2	30228.2	0.022909
Hamilton-Middleton, OH	36560.79	37405.85	0.023114
Fitchburg-Leominster, MA	32646.05	33401.82	0.02315
Boise City, ID	32759.97	33570.17	0.024731
Lancaster, PA	32411.22	33222.16	0.025021
Grand Junction, CO	29795.67	30544.78	0.025142
Springfield-Holyoke-Chicopee, MA	32634.31	33461.07	0.025334
Tallahassee, FL	34320.47	35221.08	0.026241
Flagstaff, AZ-UT	28418.07	29186.23	0.027031
Allentown-Bethlehem-Easton, PA/NJ	33345.67	34297.5	0.028544
Chico, CA	29226.81	30095.82	0.029733
Bryan-College Station, TX	35135.84	36205.46	0.030443
Memphis, TN/AR/MS	35329.28	36406.64	0.030495
Providence-Fall River-Pawtucket, MA/RI	32869.16	33881.61	0.030803
Greeley, CO	31834.76	32825.89	0.031134
Lafayette-W. Lafayette, IN	33098.84	34171.1	0.032396
Myrtle Beach, SC	28886.44	29837.68	0.03293
Janesville-Beloit, WI	33542.3	34666.22	0.033508
Tucson, AZ	29832.98	30836.94	0.033653
Panama City, FL	28760.07	29741.86	0.034137
Cleveland, OH	35535.54	36750.79	0.034198
Grand Rapids, MI	34467.61	35661.8	0.034647
Lexington-Fayette, KY	37080.36	38399.98	0.035588
Reading, PA	32627.74	33793.17	0.035719
Harrisburg-Lebanon--Carlisle, PA	32895.91	34091.15	0.036334
Gainesville, FL	33813.2	35043.57	0.036387
Omaha, NE/IA	36312.59	37660.39	0.037117
Columbia, SC	33362.96	34614.38	0.037509
Richland-Kennewick-Pasco, WA	32015.18	33248.74	0.038531
San Antonio, TX	30728.81	31920.98	0.038796
Rockford, IL	33912.7	35248.54	0.03939
Bloomington-Normal, IL	37261.69	38733.46	0.039498
Corpus Christi, TX	30003.89	31212.33	0.040276
Lansing-E. Lansing, MI	35652.79	37116.3	0.041049
Lincoln, NE	35533.91	37015.4	0.041692
Champaign-Urbana-Rantoul, IL	34266.46	35727.89	0.042649
Green Bay, WI	35894.98	37429.34	0.042746
Dover, DE	29339.38	30600.08	0.04297
Redding, CA	29306.26	30606.95	0.044383
St. Louis, MO-IL	35693.13	37308.71	0.045263
Syracuse, NY	32085.34	33556.29	0.045845
Glens Falls, NY	28794.21	30168.43	0.047725
Kankakee, IL	30986.48	32470.51	0.047893

Asheville, NC	29905.74	31390.12	0.049635
New Orleans, LA	29611.95	31093.11	0.050019
Cincinnati-Hamilton, OH/KY/IN	37125.43	38984.99	0.050089
York, PA	31886.46	33490.59	0.050308
Spokane, WA	31396.95	32977.11	0.050328
Elkhart-Goshen, IN	32975.71	34646.94	0.050681
South Bend-Mishawaka, IN	32496	34174.89	0.051664
Fayetteville, NC	27727.22	29169.54	0.052018
Greensboro-Winston Salem-High Point, NC	32968.43	34685.62	0.052086
Dayton-Springfield, OH	33525.07	35305.95	0.053121
Fresno, CA	26400.66	27816.07	0.053613
Appleton-Oskosh-Neenah, WI	33979.91	35829.42	0.05443
Birmingham, AL	34560.63	36505.23	0.056267
Yakima, WA	25919.84	27388.5	0.056662
Sioux Falls, SD	33423.83	35330.05	0.057032
State College, PA	31111.42	32903.18	0.057592
Buffalo-Niagara Falls, NY	32703.02	34617.96	0.058556
Athens, GA	31827.26	33709.53	0.05914
Little Rock--North Little Rock, AR	31263.36	33112.97	0.059162
Bakersfield, CA	26545.53	28122.33	0.0594
Lakeland-Winterhaven, FL	28077.17	29766.25	0.060159
Kokomo, IN	34404.86	36498.88	0.060864
Springfield, IL	34403.17	36518.55	0.061488
Louisville, KY/IN	35785.21	38015.79	0.062333
Pittsburgh, PA	32515.76	34555.45	0.062729
Wichita, KS	32812.72	34871.91	0.062756
Ocala, FL	25824.02	27451.35	0.063016
Kileen-Temple, TX	28723.57	30544.92	0.063409
Jackson, MS	31900.38	33931.7	0.063677
Cedar Rapids, IA	35101.52	37342.66	0.063848
Kalamazoo-Portage, MI	32821.18	34920.95	0.063976
Jackson, MI	32275.01	34351.12	0.064326
Biloxi-Gulfport, MS	28555.5	30425.64	0.065491
Tyler, TX	30786.03	32843.83	0.066842
Clarksville- Hopkinsville, TN/KY	27947.97	29822.21	0.067062
Sioux City, IA/NE	29643.03	31633.48	0.067148
Benton Harbor, MI	31828.24	33971.68	0.067344
Fayetteville-Springdale, AR	28158.83	30096.75	0.068821
Tulsa, OK	32452.94	34698.32	0.069189
Merced, CA	24285.04	25973.69	0.069534
Pueblo, CO	25764.25	27569.99	0.070087
Greenville-Spartanburg-Anderson SC	31161.92	33396.48	0.071708
Jacksonville, NC	25722.61	27569.11	0.071785
Eau Claire, WI	29371.44	31493.31	0.072243
Yuba City, CA	26722.65	28677.49	0.073153
Pensacola, FL	28796.34	30926.57	0.073976
Sheboygan, WI	34787.58	37418.89	0.075639
Fort Wayne, IN	33094.55	35612.56	0.076085
Toledo, OH/MI	32716.15	35213.5	0.076334
Visalia-Tulare-Porterville, CA	23146.36	24919.9	0.076623
Peoria, IL	33346	35915.83	0.077066

New Bedford, MA	31271.67	33702.01	0.077717
Waco, TX	29352.35	31654.29	0.078424
Oklahoma City, OK	30862.54	33289.41	0.078635
Topeka, KS	31721.66	34262.81	0.080108
Fargo-Morehead, ND/MN	33548.76	36243.2	0.080314
Chattanooga, TN/GA	32081.63	34659.54	0.080355
Wausau, WI	32489.85	35136.48	0.08146
Evansville, IN/KY	32559.77	35213.92	0.081516
Baton Rouge, LA	31183.34	33727.48	0.081587
LaCrosse, WI	34183.65	36979.45	0.081788
Mobile, AL	28130.21	30431.66	0.081814
Saginaw-Bay City-Midland, MI	31386.66	33974.64	0.082455
Jackson, TN	30507.69	33027.63	0.0826
Montgomery, AL	31185.29	33765.91	0.082751
Canton, OH	30242.19	32755.95	0.083121
Hickory-Morgantown, NC	28170.19	30522.96	0.08352
Columbia, MO	33155.13	35941.18	0.084031
Tuscaloosa, AL	30058.36	32594.16	0.084363
Lubbock, TX	31212.24	33856.98	0.084734
Knoxville, TN	30461.5	33084.84	0.08612
Columbus, GA/AL	29505.53	32055.37	0.086419
Davenport, IA-Rock Island -Moline, IL	31681.31	34429.66	0.08675
Flint, MI	28358.23	30818.55	0.086759
Greenville, NC	30216.54	32850.9	0.087183
Wichita Falls, TX	28037.92	30484.03	0.087243
St. Cloud, MN	31913.62	34732.34	0.088323
Auburn-Opekika, AL	30647.94	33356.53	0.088378
Yuma, AZ	22278.95	24251.76	0.088551
Huntsville, AL	34406.28	37491.35	0.089666
Muncie, IN	29209.46	31871.53	0.091137
Beaumont-Port Arthur-Orange,TX	28093.2	30685.27	0.092267
Roanoke, VA	31772.61	34716.59	0.092658
Macon-Warner Robins, GA	29467.5	32201.09	0.092766
Billings, MT	30643.27	33494.16	0.093035
Hagerstown, MD	29961.08	32771.91	0.093816
Springfield, MO	28736.26	31453.17	0.094546
Augusta-Aiken, GA-SC	30384.05	33264.34	0.094796
Goldsboro, NC	27127.99	29750.23	0.096662
Amarillo, TX	29573.3	32437.88	0.096864
Decatur, IL	31643.43	34791.23	0.099477
Lake Charles, LA	29620.86	32614.64	0.10107
Monroe, LA	27984.93	30828.65	0.101616
Utica-Rome, NY	27671.59	30505.45	0.102411
Erie, PA	28233.96	31133.81	0.102708
Longview-Marshall, TX	28698.81	31655.51	0.103025
Mansfield, OH	28688.1	31644.43	0.103051
Abilene, TX	27249.56	30057.68	0.103052
Binghamton, NY	30542.42	33696.98	0.103284
Odessa, TX	28770.84	31756.73	0.103782
El Paso, TX	22778.71	25148.91	0.104053
Rocky Mount, NC	26769.07	29555.19	0.10408

Laredo, TX	18744.58	20710.54	0.104882
Albany, GA	28972.03	32014.88	0.105027
Lafayette, LA	29112.28	32195.21	0.105898
Scranton-Wilkes-Barre, PA	29253.43	32409.53	0.107888
Jamestown-Dunkirk, NY	26639.57	29576.42	0.110244
Shreveport, LA	28185.75	31380.35	0.113341
Waterloo-Cedar Falls, IA	30618.42	34106.91	0.113934
Youngstown-Warren, OH-PA	28299.33	31533.73	0.114293
Lima, OH	29920.39	33359.2	0.114932
Terre Haute, IN	28164.12	31409.31	0.115225
Houma-Thibodoux, LA	25445.98	28455.34	0.118265
Sharon, PA	25922	29019.32	0.119486
St. Joseph, MO	28889.18	32359.76	0.120134
Lynchburg, VA	28828.97	32316.65	0.120978
Las Cruces, NM	22696.38	25466.04	0.122031
Williamsport, PA	27282.75	30626.77	0.122569
Joplin, MO	25723.08	28897.53	0.123409
Brownsville-Harlingen-San Benito, TX	20188.49	22689.15	0.123866
Duluth-Superior, MN/WI	30147.57	33907.75	0.124726
Fort Smith, AR/OK	25907.08	29178.6	0.126279
Johnson City-Kingsport--Bristol, TN/VA	26775.52	30171.94	0.126848
Sumter, SC	24520.5	27750.28	0.131717
Altoona, PA	25819.88	29271.94	0.133698
Alexandria, LA	25603.93	29047.51	0.134494
Hattiesburg, MS	27421.94	31294.43	0.141219
McAllen-Edinburg-Pharr-Mission, TX	17514.82	20035.72	0.14393
Florence, AL	26761.77	30614.56	0.143966
Decatur, AL	28988.35	33281.12	0.148086
Danville, VA	24925.41	28645.14	0.149234
Dothan, AL	25950.78	29873.63	0.151165
Gadsden, AL	25311.26	29179.26	0.152818
Anniston, AL	25777.07	29927.03	0.160994
Johnstown, PA	23383.64	27495.83	0.175858

Table 2: Changes in the Cost of Living, by Education Group

	1980	1990	2000	Percent Increase 1980-2000
	(1)	(2)	(3)	(4)
Official CPI				
High-School	1	1.53	2.02	102%
College	1	1.53	2.02	102%
Percent Difference	0	0	0	
Monthly Rent				
High-School	247	432	563	127%
College	259	491	642	147%
Percent Difference	4%	11%	14%	
Local CPI 1				
High-School	0.99	1.49	1.95	96%
College	1.01	1.58	2.07	105%
Percent Difference	2%	4%	6%	
Local CPI 2				
High-School	0.98	1.57	2.04	108%
College	1.01	1.71	2.22	119%
Percent Difference	3%	7%	9%	

Notes: Monthly rent refers to the rent paid for a two or three bedroom apartment. Local CPI 1 allows for local variation only in the cost of housing. Local CPI 2 allows for local variation both in the cost of housing and the cost of non-housing goods and services.

Table 3: Nominal and Real Conditional Wage Difference Between Workers with a High School Degree and Workers With College or More, by Year - Baseline Estimates

	1980	1990	2000	1980-2000 Increase	1980	1990	2000	1980-2000 Increase
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Model 1								
Nominal Wage Difference	.40	.53	.60	.20	.35	.47	.53	.18
	(.011)	(.012)	(.013)		(.007)	(.006)	(.007)	
Model 2								
Real Wage Difference - Local CPI 1	.38	.48	.53	.15	.37	.46	.52	.15
	(.010)	(.008)	(.008)		(.008)	(.006)	(.007)	
Percent of Nominal Increase Accounted for by Cost of Living				25%				17%
Model 3								
Real Wage Difference - Local CPI 2	.37	.45	.51	.14	.37	.46	.51	.14
	(.009)	(.008)	(.008)		(.008)	(.006)	(.007)	
Percent of Nominal Increase Accounted for by Cost of Living				30%				22%
MSA Fixed Effects	No	No	No		Yes	Yes	Yes	

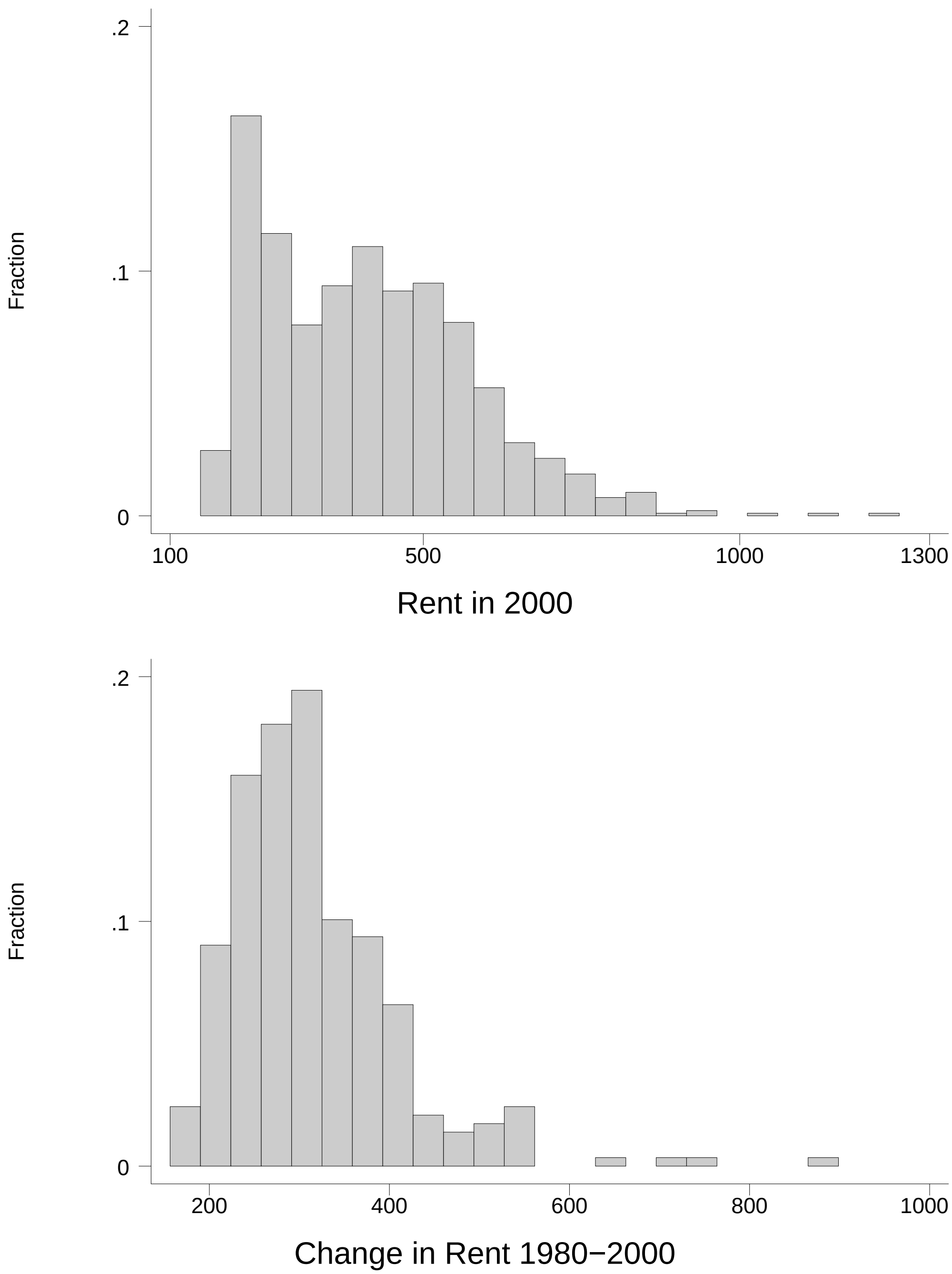
Notes: Standard errors clustered by metropolitan area in parentheses. The dependent variable in Model 1 is the log of nominal hourly wage. The dependent variable in Model 2 is the log of real hourly wage, where real hourly wage is the ratio of nominal wage and Local CPI 1. The dependent variable in Model 3 is the log of real hourly wage, where real hourly wage is the ratio of nominal wage and Local CPI 2. All models include dummies for gender and race, a cubic in potential experience, and year effects. Models in columns 5 to 8 also include MSA fixed effects. Sample size is 5,024,221.

Table 4: The Relation between Share of College Graduates and College Premium

	2000 Cross-section	1980-2000 Change	
	OLS	OLS	IV
	(1)	(2)	(3)
College Share	.375	.388	.371
	(.031)	(.070)	(.106)
R^2	.30	.10	

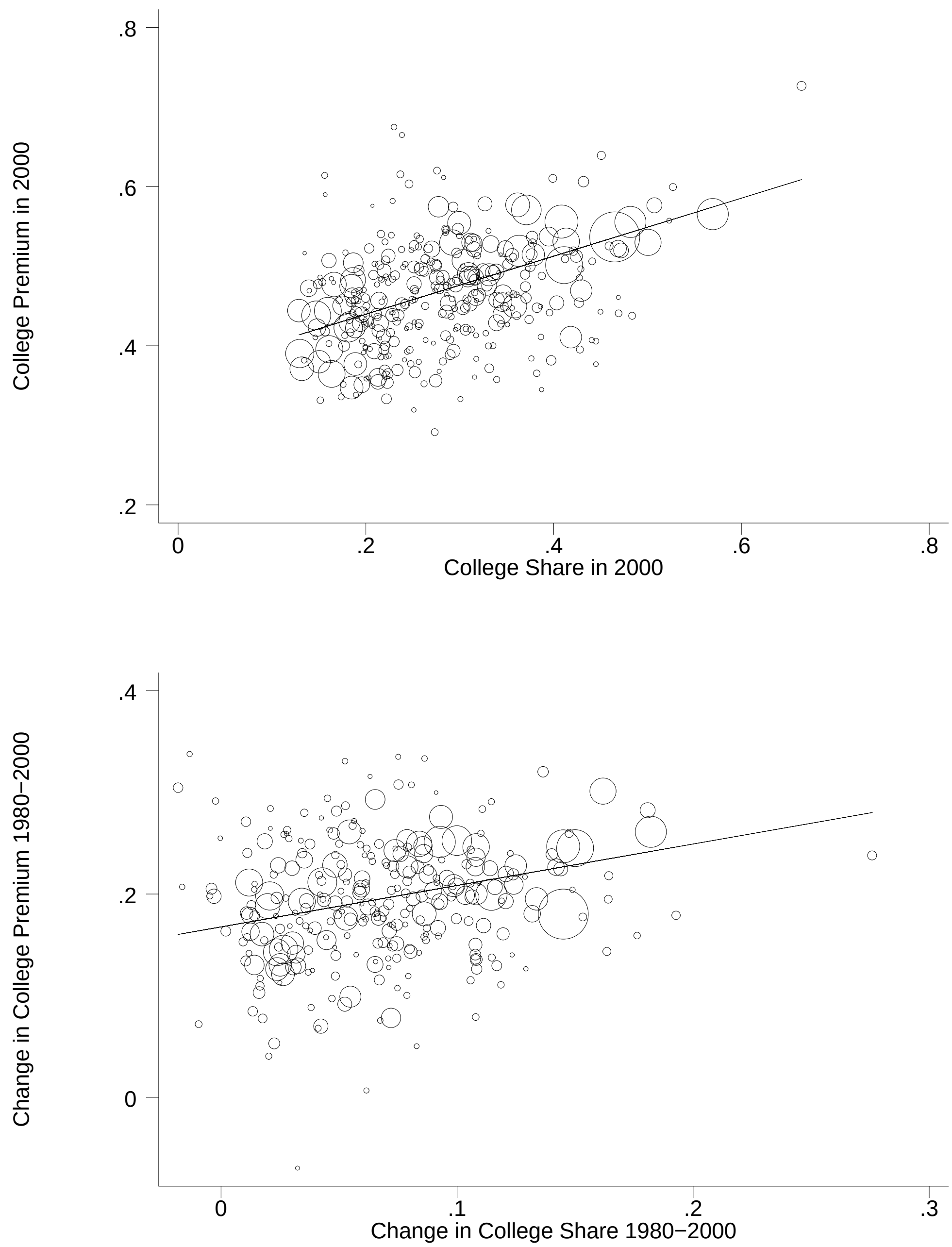
Notes: Standard errors in parentheses. The dependent variable in column 1 is the city-specific college premium, defined as the city-specific difference in the log of hourly wage for college graduates and high school graduates conditional on gender, a cubic in potential experience, race and year. The dependent variable in columns 2 and 3 is the change in the city-specific college premium. Entries are the coefficient on college share in column 1 and change in college share in columns 2 and 3. All models are weighted by city size.

Figure 1: The Distribution of Average Rental Costs Across Metropolitan Areas: 2000 Cross-Section and 1980-2000 Change



Notes: The top panel shows the distribution of the average cost of renting a 2 or a 3 bedroom apartment in year 2000. The bottom panel shows the distribution of the changes between 1980 and 2000 in the average cost of renting a 2 or a 3 bedroom apartment.

Figure 2: Share of College Graduates and College Premium, by City



Notes: The top panel plots estimates of the city-specific college premium in 2000 against the share of college graduates in 2000. The bottom panel plots the 1980-2000 change in college premium against the 1980-2000 change in the share of college graduates.

Figure 3: Share of College Graduates and Relative Demand Shocks, by City

Notes: The panel plots changes in the share of college graduates 1980-2000 on the y-axis against 1980-2000 shocks to the relative demand of college graduates due to 1980 differences in industry mix on the x-axis. Shocks to the relative demand are defined in equation 25.

Appendix Table 1: Relative Importance of the Main Aggregate Components in the BLS Consumer Price Index (CPI-U)

Housing	42.7%	
Shelter		32.8%
Fuels and Utilities		5.3%
Other Housing		4.6%
Transportation	17.2%	
Food and Beverages	14.9%	
Medical Care	6.2%	
Education and Communication	6.0%	
Recreation	5.5%	
Apparel	3.7%	
Other Goods and Services	3.5%	

Notes: Entries are the share of the main aggregate components of the CPI-U. For more disaggregated categories see Appendix 4 in Chapter 17 of the Bureau of Labor Statistics's "Handbook of Methods" (2007).