Why do cities pay more?
An empirical examination of some competing theories of the urban wage premium ☆

Jeffrey J. Yankow

Department of Economics, Furman University, 3300 Poinsett Highway, Greenville, SC 29613

Abstract

OLS regression identifies a 19 percent wage advantage for workers in large urban areas. Fixed-effects estimation suggests that two-thirds of this premium can be explained by cities attracting workers of higher unmeasured skills and ability. The remaining wage premium is shown to consist of both level and growth elements. The wage level effect is consistent with a productivity advantage for firms located in cities. The wage growth effect is shown to relate in part to a cumulative advantage in the returns to job mobility for urban workers. This last finding highlights the importance of coordination in urban labor markets.
© 2006 Elsevier Inc. All rights reserved.

JEL classification: R23; J24; J31

1. Introduction

In a recent study, Glaeser and Maré [14] report that workers in dense metropolitan areas earn 25 percent more than their non-urban counterparts when controlling for basic observable characteristics. Although previous studies of wage determination have noted such large earnings differentials, few researchers have focused on explaining this difference. The importance of obtaining a clear understanding of the workings of urban labor markets is underscored by the fact that more than 80 percent of the US population resides in and nearly 85 percent of all jobs are located in metropolitan areas (Census [7]). Moreover, the past few decades have shown both

population and employment growing faster inside metropolitan areas than outside of them (Carlino [6]). Indeed, these recent trends have inspired some prominent researchers to herald an era of "new urbanism," in which households demonstrate a renewed demand for dense, pedestrian cities (Katz [20], Glaeser and Shapiro [15]).

Why might the wages of characteristically similar workers differ between urban and non-urban areas? As Glaeser and Maré [14] make clear, this question must be addressed from the perspectives of both labor supply and demand. On the supply side, if the nominal wages paid in cities are higher than those paid in non-urban areas, what prevents all mobile workers from moving to a city? The most obvious answer is that nominal wages will differ to the extent that prices and rents vary across urban and non-urban areas. That is, the nominal wage premium collected by urban workers may simply reflect a higher cost of living in cities. If so, real wages may not differ materially across urban and non-urban areas. However, if higher urban prices do not fully account for the higher nominal wages and real wage differences exist, then it must be true that workers in cities are more productive (Glaeser and Maré [14]).[1] On the labor demand side, even if higher urban rents and prices inhibit the mass migration of workers to cities, what prevents firms from fleeing high wage urban areas? Applying a similar logic, if firms are to remain in high wage areas they must either hire more productive workers or have lower production costs (Glaeser and Maré [13]).[2] Hence long-run equilibrium real wage differentials among similar workers can arise to the extent that there are differences in worker skills and/or productivity between urban and non-urban areas.

The literature offers a multitude of competing explanations purporting to explain the source of an urban productivity advantage. Among the oldest is that cities pay more because they attract the most able workers (Fuchs [10]). If so, a significant portion of the urban wage premium is then likely to be a return to unobserved skill. An alternative hypothesis holds that workers are more productive in urban settings due to "agglomeration economies" (Kim [21], Ciccone and Hall [8], Glaeser [11]). These are efficiency gains and cost savings that result from proximity to customers, suppliers, workers, and even competitors. Still other explanations posit urban externalities in learning and human capital production (Rauch [26], Glaeser [12], Moretti [24]) or efficiencies in job search and matching (Kim [22], Helsey and Strange [17], Sato [27]).

Surprisingly, few studies have attempted to sort through these competing hypotheses empirically. In the most comprehensive study to date, Glaeser and Maré [14] analyze urban wage premia using multiple data sources and a variety of empirical methodology. Their analysis reveals a 25 percent wage advantage for workers living in large cities when controlling for basic observational characteristics such as race, schooling, experience, and job tenure. Controlling for unobserved differences with a fixed-effects estimator reduces this wage difference significantly, the urban premium falling to somewhere between 4.5 and 11 percent, depending upon the sample analyzed.[3] These authors also examine the wage growth of workers who migrate into urban areas and find that a significant fraction of the wage premium accrues only with time spent in the

---

[1] Higher wages could also be necessary to compensate city residents for various urban disamenities such as crime, pollution, and congestion. Of course, cities also offer residents many positive amenities. However, if the net amenities were positive (beyond higher prices), then the real wages of equally able workers would need to be lower in cities in order to equalize utility levels across space.

[2] It is also possible that firms in high wage areas are able to charge higher prices for their products, but this requires an assumption of non-competitive product markets for such firms.

[3] Glaeser and Maré [14] analyze data on employed workers drawn from the 1990 Census 1 percent Integrated Public Use Microdata Series (IPUMS), Panel Survey of Income Dynamics (PSID), National Longitudinal Survey of Youth 1979 (NLSY), and the Current Population Survey (CPS).

city. This time-dependent return to urban residency leads them to conclude that the urban wage difference is not simply the result of higher ability workers living in cities but that cities actually make workers more productive. Though not discounting the importance of coordination, they favor faster rates of human capital accumulation for urban residents as the most likely explanation for this productivity advantage.

Although DuMond, Hirsch, and MacPherson [9] do not explore the urban wage premium per se, their analysis does provide some useful benchmark estimates of real wage differences across cities of different size. They find that workers in areas with populations between 200 and 500 thousand garner roughly a 5 percent wage advantage, those in areas with populations between 1/2 to 2 million a 7 percent advantage, and those in areas 2 million and over an approximate 10 percent advantage—as compared to workers in the smallest urban areas (those with populations less than 200,000)—when applying a regression-based cost-of-living adjustment. Although these authors do not go beyond simple OLS regression analysis, this result is important because it establishes a real urban wage premium that remains after accounting for variation in the cost of living.[4] Moreover, they demonstrate the importance of controlling for cost of living in the study of inter-area wage differences.

Taking Glaeser and Maré [14] as a point of departure, I extend their analysis by providing new evidence on the sources of the urban wage premium. Using an extensive panel of data drawn from the National Longitudinal Survey of Youth 1979 (NLSY79), I examine the wages of urban and non-urban workers across three distinct modes of analysis: wage levels, year-to-year wage growth, and between-job wage growth. The research begins with wage level analysis in order to establish the magnitude of the urban wage premium for the sample under consideration. A credible estimate is obtained by first stripping away the effect of both observed and unobserved heterogeneity as well as variation in the cost of living across urban areas. The panel aspects of the data are exploited to their fullest extent by estimating wage equations that control for both time-invariant fixed effects and time-dependent unobserved differences in life-cycle wage profiles ("experience effects"). Once the real wage premium has been identified, the patterns of job change and wage growth—both within and across urban and non-urban areas—are examined in an effort to identify which of the multitude of competing hypotheses best explains it.

The rest of the paper is organized as follows. The next section considers each of the theoretical explanations of the urban wage premium in greater detail with special attention paid to the testable implications emerging therein. Section 3 discusses the data and sample construction. Section 4 explains the empirical strategy and ensuing results. Section 5 concludes.

2. Competing hypotheses

Cost of living differences

If a high cost of living in cities is the primary explanation for the observed wage difference between urban and non-urban workers, then the urban wage premium is a purely nominal phenomenon. If so, assuming that adequate price data were available, wage equations that control for the local cost of living should identify no real wage difference between urban and non-urban workers of similar skill and ability. However, any real wage difference that remained after

---

[4] Johnson [19] also examines wage dispersion across urban areas using a sample of 33 SMSAs over the period 1973–1976 and finds real wage differences following adjustment for the cost of living. DuMond et al., however, criticize Johnson's method of correcting for cost of living.

controlling for inter-area price variation would then most likely be explained by a real skill or productivity difference between urban and non-urban areas.

Ability-sorting hypothesis

A plausible explanation for the urban wage premium is that cities demand, attract, and retain higher-quality workers than do employers in non-urban areas, and these skills are not reflected fully in measured variables (Fuchs [10]). Highly skilled and motivated workers may be particularly attracted to cities, where their skills can be most advantageously employed and rewarded (Borjas, Bronars, and Trejo [3]). The outcome of such labor market sorting is an equilibrium in which urban workers realize higher wages than observationally similar workers located in non-urban areas.

The ability-sorting hypothesis offers several testable implications. First, accurate individual data on human capital and other productivity-related attributes should lower empirical estimates of the urban wage premium significantly. Second, although differences in worker quality are generally observable to employers, they are largely unobservable to the econometrician and unmeasured in standard data sources. If these unmeasured personal characteristics are the primary explanation for the urban wage premium, fixed-effects estimation should identify no real wage difference between urban and non-urban workers. Finally, workers who move to cities should be paid higher wages prior to moving when compared to observationally equivalent workers that do not migrate from non-urban areas.

Firm-level productivity hypothesis

The firm-level productivity hypothesis asserts that workers are more productive in firms located within cities due to economies of agglomeration. That is, the marginal product of labor is higher for city-based firms due to the production and consumption benefits of urban density (Ciccone and Hall [8]). This may result from a lower cost of transporting goods to the product market, a lower cost of acquiring inputs from local suppliers, or because of intellectual spillovers across firms. This hypothesis offers a very clear empirical prediction: Workers moving into cities should receive sizeable, immediate wage gains, whereas workers leaving cities should experience immediate (and symmetrical) wage reductions.

Learning hypothesis

The learning hypothesis postulates that cities speed the rate of human capital accumulation. Glaeser [12] contends that urban density accelerates the rate of interaction between people and that when people learn through interactions, human capital accumulation is accelerated. He argues that such externalities tend to emerge when individuals have increased contact with particularly knowledgeable and highly skilled people, which cities tend to have in abundance. This interpretation is supported by the empirical work of Rauch [26], who finds a positive correlation between individual earnings and the average level of human capital in cities. Because workers only become more productive with time in the city, the learning hypothesis implies that workers moving into cities will not experience immediate real wage gains. Likewise, workers leaving cities might not experience wage losses (except through that part of the wage related to cost-of-living differences).

Coordination hypothesis

The coordination hypothesis holds that urban density facilitates the matching of workers and firms. With knowledge of the distribution of wages in a labor market (but uncertainty regarding a particular wage offer), searchers maximize their expected return to search by calculating a reservation wage (or minimum acceptance wage) by equating the marginal benefit and cost of sampling one additional wage offer (Lippman and McCall [23]). The marginal benefit of search is primarily a function of the distribution of wage offers and the arrival rate of offers. The cost of search is composed of both out-of-pocket expenses and opportunity costs including time and travel (Mortenson [25]). This suggests at least two advantages to searching in urban rather than non-urban labor markets. First, a greater number of job openings in any time period allow more jobs to be sampled more quickly and increases the probability of receiving an acceptable offer (Burdette [5]). Second, the greater density of employment opportunities and more efficient spatial configuration in urban areas reduces search time and travel costs per job vacancy. The likely outcome for workers searching in urban labor markets is then both more productive job turnover and greater wage growth when changing employers.

As with the learning hypothesis, the coordination hypothesis implies that workers find their most productive job matches only with time spent in the city.[5] Hence workers moving into cities are unlikely to experience substantial immediate wage gains and workers migrating out of cities are unlikely to face significant wage losses. Further, it may be possible to distinguish between the learning and coordination hypotheses by examining wage growth and turnover patterns in cities. The crux of the analysis would involve determining the relative importance of within-job versus between-job wage growth amongst urban dwellers. An exceptional contribution from between-job wage growth to overall wage development would tend to support the coordination hypothesis.[6]

3. Data

The NLSY79 provides a comprehensive panel data set well suited for exploring the source of urban wage premia. The original NLSY79 cohort contains 12,686 men and women between the ages of 14 and 22 in the initial year of the survey. This age group is attractive because Glaeser and Maré [14] suggest that it is young workers that learn the fastest in dense urban settings. Further, the NLSY79 allow for both wage level and wage growth analysis, and include a rich set of personal and job characteristics including the Armed Forces Qualifying Test (AFQT). Finally, the sample covers roughly the same time periods examined by Glaeser and Maré [14] and DuMond et al. [9], facilitating comparison of results.[7]

In order to generate a sample suitable for empirical analysis, several selection criteria are introduced. The sample is first restricted to young men in order to avoid issues of labor force par-

---

[5] In a related line of inquiry, cities may promote coordination that allows individuals to specialize in specific tasks, which also may lead to higher wages over time (Becker and Murphy [1]).

[6] The learning and coordination hypotheses are not mutually exclusive. Indeed, social interactions of the type described by Glaeser [11] likely relate to the more efficient operation of urban labor markets. For example, social networks that arise among people residing in the same location have been shown to be an important source of information about job opportunities (Holzer [18], Granovetter [16]). Recent empirical work by Topa [28] and Weinberg, Reagan, and Yankow [30] suggest that social interactions can be a significant factor in the determination of labor market outcomes.

[7] One limitation of the NLSY79 data set is that it does not provide information on place of work but rather location of residence. I know of no way of overcoming this difficulty.

ticipation that may be correlated with urban residency. The original male cohort of the NLSY79 consists of 6403 young men. I next delete the 824 individuals in the military sub-sample and another 928 men observed at some point in the military. Wage observations are then limited to employed periods subsequent to the final exit from formal schooling. The date of schooling exit cannot be determined for 46 individuals, while another 11 have indeterminate schooling levels. Fixed-effects estimation requires at least two valid wage observations. For a wage observation to be considered valid, respondents must be working full time and identified as either residing within or outside of metropolitan areas. If residing within a metropolitan area, the MSA code must be known. All relevant personal and job-related information must also be available, and reported hourly wages less than $2 or greater than $100 are treated as outliers. Some 1104 individuals fail to report at least two valid wage observations, a valid MSA code for each wage observation, or other pertinent labor market information subsequent to the date of schooling exit. These restrictions leave a base sample of 3490 young men contributing 23,956 wage observations over the years 1979–1994 for empirical analysis.

Throughout this paper, the terms "city" and "urban area" are used interchangeably. By "city" I mean a dense metropolitan area. In practice, I use MSAs with over 250,000 inhabitants to identify workers living in cities. These urban areas are then partitioned into "big cities" and "small cities." Big cities are those MSAs with populations of more than 1 million whereas small cities are those with populations between 250,000 and 1 million; "non-urban" constitutes all metropolitan and rural areas with populations less than 250,000.[8]

Table 1 demonstrates how urban and non-urban residents differ in regards to their personal and job-related characteristics. For fixed characteristics like race and schooling, the averages are taken over persons; for time-varying traits, the averages are calculated over person-years. Whites (non-black, non-Hispanic) contribute a greater relative percentage to the non-urban sample, while blacks and Hispanics are more likely to be located in cities. Urban residents tend to be more highly educated in terms of highest grade attended (*HGA*) and highest grade completed (*HGC*), and are less likely to be married. On the whole, urban residents tend also to have less experience and job tenure, but are more likely to be unionized. Both average nominal and CPI-deflated wages are larger for urban residents: average (non-standardized) nominal wages are 21 percent larger while CPI-deflated wages are 18 percent larger. Overall, the pooled sample includes 18,875 observations contributed by urban residents and 5099 observations contributed by non-urban residents.

4. Empirical strategy and results

This section explores the urban wage premium in a series of panel regressions. I analyze both wage level and growth patterns, including the wage growth occurring in job transitions. The

---

[8] This distinction differs somewhat from the definitions applied in previous research. Glaeser and Maré [14] define a city to be any MSA regardless of population level (with big cities having populations greater than 500,000). Their non-urban control sample thus consists of anyone living outside of an MSA. DuMond et al. [9] make no formal distinction between urban and non-urban areas. Limiting their analysis to MSAs, they use respondents living in MSAs with populations below 200,000 as their control sample. However, if one thinks of these smallest MSAs as essentially non-urban, then a "city" is implicitly defined to be any MSA with a population greater than 200,000. I believe both of these definitions are overly restrictive. On the one hand, very small metropolitan areas, like Altoona (PA) or Anniston (AL), hardly strike me as "cities" in the conventional use of the term. On the other, I see no valid reason to delete workers living outside of MSAs from the analysis. Hence the definition employed in this study is meant to capture the more attractive aspects of each approach.

Table 1
Sample statistics

Variable	Total		Urban area		Nonurban area	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Age	27.0	4.31	27.1	4.29	26.91	4.40
Black	0.239	0.427	0.255	0.436	0.179	0.384
Hispanic	0.208	0.406	0.224	0.417	0.151	0.358
HGA	12.65	2.45	12.73	2.51	12.32	2.20
HGC	12.32	2.54	12.40	2.59	12.00	2.32
Experience	8.41	4.49	8.34	4.46	8.69	4.60
Job tenure	3.22	3.38	3.19	3.34	3.34	3.52
Avg. hours	44.5	8.37	44.3	8.18	45.1	8.99
Married	0.447	0.497	0.427	0.495	0.522	0.500
Amenity	0.987	0.115	0.990	0.120	0.978	0.093
AFQT	40.00	29.3	39.95	29.60	40.16	28.2
AFQT miss.	0.048	0.214	0.047	0.211	0.053	0.225
Union[a]	0.214	0.410	0.223	0.416	0.181	0.385
Firm size[b]	2146.3	12899.7	2396.0	13623.6	1202.8	9631.1
Public emp.	0.145	0.352	0.145	0.352	0.146	0.353
Nom. wage	9.92	7.25	10.29	7.43	8.52	6.35
Real wage	8.00	5.31	8.27	5.34	6.99	5.06
Observations	23,956		18,857		5099	

[a] Total obs. = 21,682.
[b] Total obs. = 15,425.

specific methodology employed in each section is discussed in turn along with the estimation results.

4.1. Wage level analysis

I first examine wage levels across three distinct modes of analysis: Ordinary least squares (OLS) models, in which the intercept term is constrained to be equal for all individuals; fixed-effects models, in which the intercept is allowed to vary across individuals but not over time; and fixed-effects models that allow the effects of experience to vary across individuals ("experience effects"). The latter estimates are based on individual deviations from the typical experience profile where the strength of the common experience profile is allowed to vary across individuals. To gauge the impact of individual heterogeneity on estimates of the urban wage premium and to enable the reader to assess appropriate controls, each mode of analysis begins with a basic specification that is gradually augmented with more thorough control variables.

4.1.1. OLS baseline estimates
First, a standard log wage equation of the following form is estimated:

$$\ln W_{it} = \sum_{h=1}^{H} \beta_h X_{iht} + \sum_{k=1}^{2} \gamma_k CITY_{ikt} + \sum_{m=1}^{M-1} \tau_m YEAR_{imt} + \varepsilon_{it} \tag{1}$$

where $\ln W_{it}$ is the log of the regional CPI-deflated hourly wage for worker i at time t. The vector X contains $H-1$ personal, job, and labor market characteristics, and β contains the corresponding coefficients ($X_1 = 1$ and β_1 is the intercept). *CITY* contains two dummy variables

designating residence in urban areas of different population size: residence in an urban area with population greater than one million (*Big City*) and residence in an urban area with population between 250,000 and one million (*Small City*). The coefficients in γ are the adjusted log earnings differences by urban residence status relative to non-urban residents. *YEAR* includes dummy variables for the years 1979–1993 (1994 omitted). For now, ε_{it} is assumed to be a well-behaved error term consisting of white noise.

The first six columns of Table 2 contain OLS estimates of the urban wage advantage based on Eq. (1). Column (1) pertains to a simple specification that includes only indicators of urban residence, race dummies, potential experience (and its square), and year dummies. The point estimate for *Big City* is measured at 0.220 with a standard error of 0.017, suggesting that workers in large cities receive almost a 25 percent wage premium over workers residing outside of urban areas. The coefficient on *Small City* is measured at 0.095 with a standard error of 0.019. Although less than half the magnitude of the *Big City* coefficient, it nevertheless suggests more than a 10 percent wage premium for small city residents relative to non-urban workers. This naïve specification, however, explains only 16 percent of the variation in log wages.

Columns (2)–(4) augment the baseline specification by including observable measures of human capital. The specification displayed in column (2) includes the worker's highest grade attended (*HGA*). As the ability-sorting hypothesis would suggest, the city coefficients falls—but only slightly. The point estimates for *Big City* and *Small City* are measured at 0.190 and 0.069, respectively. Column (3) presents estimates from a specification that adds the respondent's age-adjusted AFQT score. Including this proxy for aptitude and skill lowers the *Big City* coefficient estimate further to 0.187. Interestingly, the *Small City* coefficient estimate rises marginally to 0.075. The specification in column (4) adds controls for marital status, job tenure, employment in the public sector, and one-digit industry and occupation controls. Although this model now explains 41 percent of the variation in log wages, there is no discernible reduction in the *Big City* coefficient estimate. The point estimate for *Small City* rises slightly to 0.082.

One possible explanation for the remaining premium is that urban workers are more likely to be unionized. Column (5) shows the results from a specification that includes a control for wages set by collective bargaining agreements (*Union*). The point estimates for the city coefficients do fall , but only to 0.176 and 0.078, respectively. Another possibility is that the urban wage premium is really a firm-size premium. Brown and Medoff [4] show that large firms tend to pay higher wages than smaller firms other things equal, and large firms tend to locate in metropolitan areas. Including controls for the size of the firm in which the respondent is employed lowers both city coefficients slightly to 0.176 and 0.078, respectively.[9] In sum, after controlling for an impressive array of observable characteristics, OLS regression identifies about a 19 percent wage advantage for large city residents and an 8 percent advantage for workers in small cites.

4.1.2. Fixed effects

A much discussed explanation for the urban wage premium is that cities pay more because they attract the most highly skilled and able workers. If worker skills or ability are not measured adequately by the control variables but are correlated with urban residence, then OLS estimates of the urban wage premium will be biased. In this section, I attempt to determine the extent of such bias by accounting for unobserved worker skill and ability in estimation.

---

[9] Firm size measures are only available in the NLSY79 post-1985. In order to maintain sample size, I set missing observations to the sample mean and include a dummy indicating that firm size was missing.

Table 2
Effect of urban residence on wages

Variable	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)	OLS (6)	Fixed (7)	Fixed (8)	Exper. (9)	Exper. (10)
Big City	0.220[a]	0.190[a]	0.187[a]	0.187[a]	0.176[a]	0.175[a]	0.050[b]	0.054[a]	0.049[b]	0.052[a]
	(0.017)	(0.015)	(0.015)	(0.012)	(0.012)	(0.012)	(0.022)	(0.020)	(0.022)	(0.020)
Small City	0.095[a]	0.069[a]	0.075[a]	0.082[a]	0.078[a]	0.076[a]	0.033	0.038[c]	0.026	0.032
	(0.019)	(0.018)	(0.017)	(0.014)	(0.014)	(0.014)	(0.023)	(0.021)	(0.023)	(0.021)
Black	−0.300[a]	−0.253[a]	−0.163[a]	−0.087[a]	−0.098[a]	−0.101[a]				
	(0.016)	(0.014)	(0.016)	(0.013)	(0.013)	(0.013)				
Hispanic	−0.161[a]	−0.064[a]	−0.019	−0.001	−0.005	−0.006				
	(0.017)	(0.017)	(0.017)	(0.014)	(0.014)	(0.014)				
Exper.	0.012[a]	0.042[a]	0.042[a]	0.020[a]	0.019[a]	0.020[a]	0.065[a]	0.044[a]		
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)		
(Exper. sq.)/10	−0.012[a]	−0.012[a]	−0.013[a]	−0.007[a]	−0.007[a]	−0.007[a]	−0.022[a]	−0.016[a]		
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)		
HGA		0.085[a]	0.058[a]	0.044[a]	0.042[a]	0.041[a]				
		(0.003)	(0.004)	(0.003)	(0.003)	(0.003)				
AFQT			0.004[a]	0.003[a]	0.003[a]	0.003[a]				
			(0.000)	(0.000)	(0.000)	(0.000)				
AFQT miss.			0.008	0.017	0.017	0.019				
			(0.030)	(0.026)	(0.025)	(0.025)				
Married					0.121[a]	0.117[a]	0.115[a]		0.053[a]		0.054[a]
					(0.009)	(0.009)	(0.009)		(0.008)		(0.008)
Tenure					0.060[a]	0.056[a]	0.054[a]		0.039[a]		0.025[a]
					(0.003)	(0.003)	(0.003)		(0.003)		(0.001)
(Tenure sq.)/10					−0.027[a]	−0.026[a]	−0.025[a]		−0.020[a]		−0.029[a]
					(0.002)	(0.002)	(0.002)		(0.002)		(0.003)
Govt.					−0.023[b]	−0.050[a]	−0.050[a]		−0.000		0.007
					(0.012)	(0.011)	(0.011)		(0.009)		(0.009)
Union						0.184[a]	0.175[a]		0.119[a]		0.118[a]
						(0.010)	(0.010)		(0.009)		(0.009)
Union miss.						0.032[b]	0.039[a]		0.034[a]		0.032[a]
						(0.015)	(0.015)		(0.012)		(0.012)
Industry	No	No	No	Yes	Yes	Yes	No	Yes	No	Yes
Occupation	No	No	No	Yes	Yes	Yes	No	Yes	No	Yes
Firm size	No	No	No	No	No	Yes	No	No	No	No
R^2	0.16	0.26	0.28	0.41	0.43	0.43	0.13	0.20	0.07	0.14

Notes. Standard errors are in parentheses. Dependent variable is log of regional CPI-deflated hourly wage. *Big City* refers to residence in an MSA with population > 1 million residents, *Small City* refers to residence in an MSA with population > 250,000 but less than 1 million residents. All regressions include year dummies. Standard errors are corrected for heteroskedasticity and within-person correlation.
[a] Significant at the 1% level.
[b] Idem, 5%.
[c] Idem, 10%.

Equation (1) is modified to account for unobserved personal differences fixed over time. The wage equation can be written as

$$\ln W_{it} = \sum_{h=1}^{H} \beta_h X_{iht} + \sum_{k=1}^{2} \gamma_k CITY_{ikt} + \sum_{m=1}^{M-1} \tau_m YEAR_{imt} + \alpha_i + \varepsilon_{it}. \tag{2}$$

The error term is now decomposed into a time-invariant, person-specific quality component (α_i) and a purely random element (ε_{it}). Assuming the fixed effect α_i is positively correlated with urban residence, OLS estimation of Eq. (2) will result in estimates of the urban wage premium that are biased upward. Accordingly, estimation is performed using standard fixed-effects techniques.[10]

Columns (7) and (8) of Table 2 present estimates of the urban wage premium obtained from the fixed-effects model. The specification displayed in column (7) includes only potential experience and urban residency (time-invariant characteristics such race and schooling fall out of the model). Controlling for person-specific fixed effects reduces the urban wage premium significantly, the point estimates falling to 0.050 for *Big City* and a statistically insignificant 0.033 for *Small City*. Column 8 shows the results from a model that also includes marital status, job tenure (and its square), public sector employment, and industry and occupation controls. The point estimates rise to 0.054 for *Big City* and 0.038 for *Small City*, the latter estimate now significant at a 10 percent level. These fixed-effects results indicate that, when the impact of time-invariant unobserved heterogeneity is removed, large cities pay a wage premium of about 6 percent while smaller cities pay a premium no larger than 4 percent. These results further suggest that time-invariant unobserved heterogeneity can explain between one-half to two-thirds of the urban wage premium identified in the baseline OLS model.

Before moving on, it is also worthwhile mentioning how these estimates compare to those of Glaeser and Maré [14], who perform a similar analysis with NLSY79 data. Employing a fixed-effects estimation procedure, these authors measure urban wage premia of 11 percent for large cities and 7 percent for small urban areas.[11] The fact that both studies identify a significant wage premium for workers in urban areas after controlling for time-invariant, unobserved heterogeneity highlights the need to go beyond the simple ability-sorting hypothesis to develop a more complete explanation of the urban wage advantage.

4.1.3. Experience effects

Another difficulty arises if individuals experience different growth rates in some unmeasured productive attributes. If a person's unobserved skill or ability changes over time, then correcting solely for a fixed effect will be inadequate. Indeed, the empirical work of Glaeser and Maré [14] suggest that a significant portion of the urban wage premium accrues to workers over time as they

---

[10] The fixed-effects model provides an unbiased measure of the urban wage premium under the assumptions that residential switching is exogenous and ability is equally valued at the margin by employers in urban and non-urban areas. Unfortunately, urban migration is endogenous, causing a potential selection bias to occur via correlation between the transitory component of the error term and residential mobility. Such issues are most likely to arise if workers are poorly matched with jobs in their location of origin. Because the city coefficients are identified off of workers moving both into and out of cities, the direction of this bias is uncertain. Another bias can occur if employers in cities value "ability" more at the margin than employers in non-urban areas. In this case, fixed-effects estimates will overstate the wage gain that would result from relocating a randomly selected person. One solution would be to find a measure that predicts urban residency but is orthogonal to ability, and use it as an instrument. Unfortunately, I am unaware of any such variables and so do not instrument the city indicator variables. Glaeser and Maré [14] experiment with the urbanization of the states in which the respondent's parents were born. They report, however, that this potential instrument fails standard specification tests, as it is positively correlated with current wages even when controlling for current urban residency. Because the magnitude of such biases is difficult to ascertain, readers should interpret the fixed-effects estimates with some caution.

[11] These estimates, which can be found in Glaeser and Maré [14, Table 3], are somewhat larger than those reported in this study. The most likely explanation for this empirical disparity is that they use respondents living outside of MSAs as their control group. In contrast, I use respondents living in areas with populations below 250,000 as the "non-urban" control sample, which includes many small MSAs.

gain experience. To contend with this possibility, I next estimate fixed-effects models that allow the impact of experience to vary uniquely across individuals ("experience effects"). Parameter estimates were produced according to the following simple two-stage procedure. First, a typical experience profile was generated for each variable (including the wage) by regressing it on a quadratic in experience using all person-year observations in the sample. Letting z_{it} denote an arbitrary variable used in the analysis and e_{it} denote years of potential experience, I estimate:

$$z_{it} = \phi_1^z + \phi_2^z e_{it} + \phi_3^z e_{it}^2 + \upsilon_{it}^z. \tag{3}$$

The residuals from each regression were then retained:

$$\hat{z}_{it} = \hat{\phi}_2^z e_{it} + \hat{\phi}_3^z e_{it}^2. \tag{4}$$

In the second stage, the residuals of the dependent variable were regressed on the residuals of the independent variables using a fixed-effects estimator. Because this procedure uses deviations in the variables from individual-specific experience-trends, it controls for the impact of unobserved interpersonal differences in the strength of experience on the variables as well as time-invariant individual fixed effects.

The final two columns of Table 2 report estimates obtained from the procedure described above. Column (9) pertains to a specification with only the city-size indicators. Accounting for time-dependent unobserved heterogeneity reduces the coefficient estimates for *Big City* to 0.049 and *Small City* to a statistically insignificant 0.026. The specification depicted in column (10) includes marital status, job tenure (and its square), public employment status, union status, industry, and occupation controls. Including these controls raises the coefficient estimates slightly. The point estimate for *Big City* is measured at 0.052 and highly significant. This estimate suggests that workers in large metropolitan areas still receive a wage premium of just over 5 percent once unobserved fixed and experience-dependent effects are taken into account. The statistically insignificant estimate of 0.032 for *Small City* implies that workers in small urban earn wages that are no different from similar non-urban workers.

4.1.4. Cost of living

The preceding analysis offers suggestive evidence that large cities pay workers a wage premium. Large urban areas, however, have a higher cost of living than less densely populated areas. Because the fixed-effects procedure does not account for this fact, it is still possible that the results are driven in part by inter-area price differences. Consequently, some attempt must be made to deal with cost of living in the empirical analysis. Ideally, one would like to control for price variation across both urban and non-urban areas. While several indices provide cost of living information for metropolitan areas, none that I know of cover non-urban areas. Thus, controlling for cost of living across the entire sample of wage observations is not feasible. Recall, however, that much of the theoretical discussion of the urban wage premium centered on the productive benefits of urban density. According to the various urban productivity hypotheses, larger metropolitan areas are expected to pay higher wages than less-populated areas due to the efficiencies created by urban density *even after controlling for cost of living*. Exploiting this notion, I focus the analysis on a subsample of wage observations obtained from MSAs of different population size for which cost of living information is available. Cost of living data for 185 MSAs are obtained from DuMond et al. [9], who develop a price index from quarterly data provided by the American Chamber of Commerce Researchers Association (ACCRA).[12] This in-

---

[12] The complete index can be found in DuMond et al. [9, Appendix Table A.1].

dex is particularly attractive because it covers the period 1985:Q4 through 1995:Q2, a time span roughly coincident with the period from which I draw the majority of wage observations from the NLSY79.[13] The metropolitan areas covered by the index contain roughly 70 percent of the US labor force. Limiting the sample to these MSAs leaves 21,285 wage observations (contributed by 3109 individuals) for empirical analysis.

In the ensuing analysis, cost of living is controlled for in two different ways. The first is a "full" cost of living (COL) adjustment whereby wages are deflated directly using the ACCRA index. DuMond et al. [9], however, argue that because this index is calculated using a fixed bundle of consumer goods and services, it systematically overstates the true cost of living in more expensive areas. They contend the relationship between wages and prices is nonlinear because consumers can and do substitute goods and location amenities across locations. Thus, in order to keep utility constant as one moves from less costly to more costly areas, nominal wages must increase, but only at a decreasing rate. Consequently, these authors endorse a "partial" COL adjustment method that uses the ACCRA index measure as an explanatory variable rather than deflating the dependent variable directly. The partial COL-adjusted estimates of the urban wage premium are the preferred estimates in this study.

Table 3 presents OLS, fixed-effects, and experience-effects estimates of the urban wage premium corrected for cost of living. To be clear, the unadjusted wage regressions (introduced for comparison purposes) simply use $\ln W$ as the dependent variable; "full" COL-adjusted regressions use $\ln(W/P)$ as the dependent variable; "partial" COL-adjusted regressions use $\ln W$ as the dependent variable and include $\ln P$ as an explanatory variable, where W is the CPI-deflated hourly wage and P is the ACCRA cost of living index. Note that the control sample against which urban workers are now measured consists exclusively of workers in MSAs with populations below 250,000. Looking first at the OLS estimate with no cost of living adjustment in column (1), we see that workers living in big cities collect a wage premium of more than 15 percent whereas those living in small cities receive only a 4 percent advantage. Fully adjusting for cost of living reduces the big city wage premium substantially—the point estimate for *Big City* falling to 0.044. Interestingly, the coefficient estimate for *Small City* (0.051) now actually exceeds that of *Big City* by a small amount. The partial COL-adjusted estimates return to the more familiar pattern, the point estimate for *Big City* is measured at an intermediate 0.121 and the estimate for *Small City* measured at 0.042.[14]

Columns (4)–(9) of Table 3 contain both the fixed-effects and experience-effects estimation results. Because the pattern of results is virtually identical across estimation procedures, except for the experience-effects estimates being slightly smaller in magnitude, I focus discussion on the results produced by the experience-effects model only. The experience-effects estimates in column 7 are unadjusted for cost of living. The point estimate for *Big City* is measured at 0.047 and is statistically significant at the 5 percent level. The estimate for *Small City* of 0.023 is statistically insignificant. Implementing the full COL adjustment lowers the estimate for *Big City* to a

---

[13] Because each ACCRA quarterly report is a unique comparison of area prices at a single point in time, it cannot be used to measure price inflation. Making the reasonable assumption that *relative* inter-area prices changed little over the observation period, DuMond et al. [9] condense the quarterly data into a single index number for each of the 185 MSAs by averaging across the time series. For specifics regarding the construction of the condensed index see [9, pp. 582–583].

[14] Using individual data drawn from the Current Population Survey Outgoing Rotation Group (CPS-ORG) monthly earnings files 1985–1995, DuMond et al. [9] produce estimates that are quite comparable in magnitude to these figures. Applying the partial COL adjustment, they find that workers in areas with populations greater than 1 million collect about a 9 percent wage premium, while those in areas with populations between 200,000 and 1 million garner roughly a 5.5 percent wage advantage.

Table 3
Effect of urban residence on wages adjusted for inter-MSA cost-of-living differences

Variable	OLS (1)	OLS (2)	OLS (3)	Fixed (4)	Fixed (5)	Fixed (6)	Exper. (7)	Exper. (8)	Exper. (9)
Big City	0.148[a]	0.044[a]	0.121[a]	0.052[b]	−0.011	0.062[b]	0.047[b]	−0.016	0.057[b]
	(0.016)	(0.016)	(0.016)	(0.025)	(0.026)	(0.025)	(0.025)	(0.026)	(0.025)
Small City	0.039[b]	0.051[a]	0.042[a]	0.031	0.052[c]	0.028	0.023	0.045[c]	0.019
	(0.017)	(0.017)	(0.017)	(0.026)	(0.027)	(0.026)	(0.026)	(0.027)	(0.026)
Black	−0.100[a]	−0.123[a]	−0.106[a]						
	(0.013)	(0.014)	(0.013)						
Hispanic	−0.009	−0.039[a]	−0.017						
	(0.014)	(0.015)	(0.014)						
HGA	0.041[a]	0.042[a]	0.041[a]						
	(0.004)	(0.004)	(0.004)						
Exper.	0.019[a]	0.022[a]	0.019[a]	0.044[a]	0.044[a]	0.044[a]			
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)			
(Exper. sq.)/10	−0.006[a]	−0.007[a]	−0.006[a]	−0.016[a]	−0.016[a]	−0.016[a]			
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)			
AFQT	0.003[a]	0.003[a]	0.003[a]						
	(0.000)	(0.000)	(0.000)						
AFQT miss.	0.020	0.005	0.016						
	(0.027)	(0.027)	(0.027)						
Married	0.117[a]	0.133[a]	0.121[a]	0.052[a]	0.055[a]	0.051[a]	0.052[a]	0.055[a]	0.052[a]
	(0.009)	(0.010)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Tenure	0.058[a]	0.057[a]	0.058[a]	0.040[a]	0.040[a]	0.040[a]	0.025[a]	0.025[a]	0.025[a]
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
(Tenure sq.)/10	−0.027[a]	−0.026[a]	−0.027[a]	−0.021[a]	−0.020[a]	−0.021[a]	−0.031[a]	−0.030[a]	−0.031[a]
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Govt.	−0.042[a]	−0.043[a]	−0.042[a]	0.002	0.001	0.002	0.008	0.006	0.008
	(0.012)	(0.012)	(0.011)	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.009)
Union	0.183[a]	0.164[a]	0.178[a]	0.118[a]	0.118[a]	0.118[a]	0.117[a]	0.117[a]	0.117[a]
	(0.011)	(0.011)	(0.011)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Union miss.	0.046[a]	0.031[b]	0.042[a]	0.035[a]	0.035[a]	0.034[a]	0.032[a]	0.032[a]	0.032[a]
	(0.015)	(0.015)	(0.015)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
COL adjustment	None	Full	Partial	None	Full	Partial	None	Full	Partial
R^2	0.44	0.43	0.44	0.21	0.21	0.21	0.09	0.08	0.09

Notes. Standard errors are in parentheses. Dependent variable is log of unadjusted CPI-deflated hourly wage, log of full cost-of-living-adjusted CPI-deflated hourly wage or log of partial cost-of-living-adjusted CPI-deflated hourly wage. *Big City* refers to residence in an MSA with population > 1 million residents, *Small City* refers to residence in an MSA with population > 250,000 but less than 1 million residents. All regressions include industry, occupation, and year dummies. Standard errors are corrected for heteroskedasticity and within-person correlation.

[a] Significant at the 1% level.
[b] Idem, 5%.
[c] Idem, 10%.

statistically insignificant −0.016, while raising the estimate for *Small City* to a marginally significant 0.045. This result deserves some comment. Why might workers in the largest urban areas receive cost-of-living-adjusted wages that are below the level paid to workers in smaller cities? One possibility is that residents of the biggest and most vibrant cities implicitly "pay" for the amenities typically found in large urban centers (more and better cultural opportunities, restaurants, night life, etc.) through lower real wages. Although plausible, the more likely explanation relates to the use of the "full" cost-of-living correction described earlier. Recall that the ACCRA index systematically overstates the cost of living in more expensive areas. Consequently, wages

in the largest cities are "overcorrected" with the full adjustment method. The final column of Table 3 shows the preferred experience-effects results using the partial COL adjustment method. The point estimate for *Big City* is now measured at 0.057 and significant at the 5 percent level. The point estimate for *Small City* is a statistically insignificant 0.019. These estimates suggest about a 6 percent wage advantage for workers living in large metropolitan areas, but no difference in the wages paid by small cities when compared to the least densely populated MSAs. These results are also very much in line with the comparable estimates reported in Table 2. In sum, even after controlling for variation in the cost of living across urban areas, it still possible to identify a modest and statistically significant wage premium for workers in large cities.

4.2. Wage growth analysis

The results of the previous section demonstrate that wage level estimates of the urban wage premium that fail to account for unmeasured worker skill and ability are upward biased considerably. However, even after controlling for both time-invariant and experience-dependent unobserved heterogeneity, workers in large metropolitan areas are still shown to receive a wage premium of more than 5 percent. Glaeser and Maré [14] argue that much of this premium is likely to be a wage growth effect, varying with the duration of residency in a city. This section explores the patterns of wage growth experienced by city residents in an effort to explain this time-dependent effect.

4.2.1. Annual wage growth

In this section attention is focused on the annual wage growth experienced by workers in the sample. To obtain a suitable estimating equation, I simply difference wage level Eq. (2) across time periods. Letting Δ represent changes between survey years, the wage change model takes the form (dropping the individual subscript i)

$$\Delta \ln W_d = \sum_{h=1}^{H} \beta_h \Delta X_{hd} + \sum_{k=2}^{K} \gamma_k \Delta CITY_{kd} + \sum_{d=2}^{D} \phi_d PERIOD_d + \varepsilon_d \qquad (5)$$

where d indexes the time periods over which values are differenced and $PERIOD_d$ are dummies for the periods 1980–1981 through 1993–1994 (with 1979–1980 as the reference period). The vector $\Delta CITY$ now contains (up to) K dummy variables designating various types of residential movement into (or out of) urban areas, with the coefficients in γ meant to capture the return to urban migration and residency. All independent variables X are differenced across time periods, and so time-invariant measures fall out of the estimating equation. The dependent variable is the difference in reported log wages across survey dates. In the ensuing analysis, I examine both "short" and "long" differences in wages. A short difference in wages ($\Delta \ln W(t+1)$) measures wage growth across two consecutive survey periods (t and $t+1$), or roughly wage growth over a one-year period in time; a long difference in wages ($\Delta \ln W(t+4)$) measures wage growth occurring over the period spanning periods t to $t+4$, or wage growth over a four-year time span.[15]

---

[15] Notice how differencing the data causes any time-invariant error components to fall out of the estimating equation. Hence Eq. (5) is a fixed-effects model (with all of the caveats of the previous section). When the data are differenced over pairs of contiguous years, the estimates obtained from Eq. (5) will differ in an important way from a fixed-effects wage level model. Since this wage growth model only requires the individual-specific effect to be fixed over a time period of one year, it is more akin to an "experience-effects" model. To the extent that both unobserved ability and the urban wage premium vary over time or with labor market experience, this is a highly desirable feature of the differenced wage model.

The short difference model allows for a direct test of the firm-level productivity hypothesis. Recall that this hypothesis asserts that cities pay more because urban firms raise the productivity of workers (either due to greater demand or economies of agglomeration). Therefore, if workers become more productive simply by working in urban firms, then estimates of γ will be positive since a productivity/wage increase will be experienced *immediately upon arrival in the city*. Moreover, these wage gains will be symmetric to the wage reductions experienced by workers moving from cities to non-urban areas. On the other hand, if the urban wage premium is due entirely to cities attracting workers with higher levels of unobserved skill and ability or raising productivity over time through coordination and/or learning externalities, then marginal products will be unchanged (in the near term) when moving between urban and non-urban areas. If so, estimates of γ will be close to zero in a model of year-to-year wage growth.

The first two columns of Table 4 display the results from models using the annual change in log wages as the dependent variable ($\Delta \ln W(t+1)$). Column 1 refers to a specification that includes only a single variable that identifies workers moving into and out of cities ($\Delta City$) and a standard set of controls. The coefficient estimate is measured at 0.059 and is significant at the five percent level. This implies that workers moving into cities from non-urban areas experience 6 percent higher wage growth during the year of the move, when compared to workers either living in a city for both years or living outside of an urban area for both years. This point estimate can also be compared to both the fixed and experience-effects estimates in Table 2, which are very similar in magnitude. It also implies that workers living in cities in year t, but moving to non-urban areas in year $t+1$, see their wages fall by 6 percent relative to the comparison group of non-movers. However, because this specification restricts estimates of γ to be symmetrical both for workers moving into cities and out of cities by construction, it cannot be used to test the validity of the productivity hypothesis.

In order to test empirically the assumption of symmetry in the wage changes, I augment the specification to include separate dummy variables indicating migration into an urban area (*In-City*), migration out of an urban area (*Out-City*), and urban residence in the first year (*City*(1)). With this modification, the control group now consists of those workers residing outside of urban areas in both years. The results of this model are presented in column 2. The coefficient estimate

Table 4
Effect of change in urban residency on wages

Variable	$\Delta \ln W(t+1)$	$\Delta \ln W(t+1)$	$\Delta \ln W(t+4)$	$\ln W_t$	$\ln W_{t+1}$	$\ln W_{t+4}$
ΔCity	0.059[b]					
	(0.026)					
City(1)		0.013[a]	0.007	0.187[a]	0.191[a]	0.188[a]
		(0.004)	(0.013)	(0.018)	(0.018)	(0.024)
In-City		0.061[c]	0.111[c]	0.039	0.088[b]	0.170[a]
		(0.037)	(0.058)	(0.044)	(0.041)	(0.059)
Out-City		−0.069[c]	0.009	−0.108[a]	−0.178[a]	−0.122[b]
		(0.037)	(0.048)	(0.040)	(0.040)	(0.057)
R^2	0.02	0.02	0.12	0.14	0.14	0.14
Observations	17,510	17,510	8249	17,510	17,510	8249

Notes. Standard errors are in parentheses. In addition to the variables indicated, wage growth regressions include the change in experience squared, change in public sector employment status, change in marital status, change in state amenity index, and year dummies. Wage level regressions also include race dummies.

[a] Significant at the 1% level.
[b] Idem, 5%.
[c] Idem, 10%.

of 0.013 for *City*(1) implies that workers living in a city in both periods experienced slightly faster wage growth than workers residing outside of urban areas. The *In-City* variable measures the change in wages of non-urban-to-urban migrants relative to workers remaining in non-urban areas over the observation period. The coefficient estimate is 0.061, and like the estimate reported in column 1, suggests about a six percent relative increase in wages for city migrants. The *Out-City* variable captures the wage change of a worker who moves from a city to a non-urban area. The coefficient estimate of -0.069 implies that workers moving out of cities experience about a 7 percent wage reduction relative to city stayers and a 6 percent reduction $(-0.069 + 0.013)$ compared to the control sample of non-urban workers.[16] What is striking is the clear symmetry in the wage changes experienced by workers moving into and out of cities, roughly six percent in both directions. This result is consistent with both a higher marginal product of labor and an increased cost of living in cities. In either case, it does strongly suggest that workers change locations in order to achieve higher utility, otherwise workers migrating out of cities would not be willing to accept such steep relative wage reductions.

One important question not yet examined is whether the urban wage premium includes a growth effect. I address this question by estimating a "long difference" model of wage growth. In this model, location decisions are still identified between periods t and $t + 1$, but wage growth is now measured as the difference in wages between periods t and $t + 4$.[17] By measuring wage growth over a longer time horizon, I hope to shed some light on the time-dependent impact of city residency. This mode of analysis is important because both the coordination and learning hypotheses hold that workers moving into cities only collect significant pecuniary benefits with time spent in the city. Unfortunately, because respondents are required to contribute three yearly observations at specific points in time, the sample size to falls to 8249 wage growth observations.

Column 3 of Table 4 presents parameter estimates obtained from a model using the four-year change in wages as the dependent variable $(\Delta \ln W(t + 4))$. The control group is again those workers living outside of urban areas in both years t and $t + 1$. The insignificant coefficient estimate for *City*(1) suggests similar growth rates in wages between city natives (those workers observed to reside in a city in both years t and $t + 1$) and the non-urban control group. The positive and significant estimate for *In-City* indicates that workers who moved into cities between years t and $t + 1$ experience a post-migration growth rate in wages that is greater than either native city workers or non-urban workers, a finding consistent with both urban coordination and learning externalities. The insignificant coefficient estimate of 0.009 for *Out-City* shows that the wages of urban-to-non-urban migrants also grew at the same rate as the control group. Because the short difference model of column 2 identified a significant initial reduction in wages over the first year, this estimate may actually imply a faster cumulative growth rate over the period from $t + 1$ to $t + 4$ relative to the control sample (though still lower than the growth rate enjoyed by city entrants).

Although the results presented in column 3 are suggestive, they are by no means conclusive. Moreover, it still remains to be seen whether the wage gains experienced by recent city entrants are commensurate with the premium collected by city natives. Columns 4–6 of Table 4 show

---

[16] The coefficient estimates for *In-City* and *Out-City* are only significant at the 10 percent level, demonstrating that although this specification is less restrictive, the gain from reduced bias is offset in part by the loss of precision from the smaller sample of migrants.

[17] For example, consider a worker with valid observations available for the survey years 1984, 1985, and 1988. This model uses the location decisions made between 1984 and 1985 but the change in wages occurring between the 1984 and 1988 surveys.

the results from wage level specifications that include dummies for the four residential groups. The results presented in column 4 clarify the relative positioning of wage levels in the year prior to a move ($\ln W_t$). The coefficient estimate for *City*(1), measured at 0.187, shows that workers living in cities earned wages that were nearly 20 percent larger than workers living in non-urban areas in period t. This is essentially the nominal wage difference between urban and non-urban workers first identified in Table 2. The statistically insignificant parameter estimate of 0.039 for *In-City* reveals that workers residing in a non-urban area in the first year, but who subsequently move to a city in the next year, earned wages roughly equivalent to other workers living in non-urban areas (who subsequently do not move into cities). This is interesting because it works against the ability-sorting hypothesis, which predicts these future movers would have earned superior wages in their location of origin. This result, however, most likely relates to the selective nature of migration. Even if (future) city migrants are drawn from the upper-tail of the ability distribution, they are also most likely the workers with the poorest wage draws in the location of origin. Finally, those workers living in cities in year t, but who subsequently move to non-urban areas in year $t + 1$, earn wages significantly below city natives who remain in cities in the second year, again demonstrating the selectivity of migration. The coefficient estimate for *Out-City* shows their wages to be -0.108 log points below city stayers (but still 0.079 log points above the control sample of non-urban stayers).

Column 5 of Table 4 shows how these wages evolve in the short-term, by identifying the relative positioning of wages in the following year ($\ln W_{t+1}$). In year $t + 1$, city residents still earn roughly 20 percent more than workers in non-urban areas. The point estimate for *In-City* is 0.088 and statistically significant, indicating that workers who moved from non-urban areas into cities now receive wages more than 9 percent higher than non-urban workers. Most interesting, however, is the finding that, in the year after moving, the wages of these city entrants are still nearly 10 percent *lower* than city natives. In other words, while these non-urban-to-urban migrants see a significant increase in their wages after moving to a city, they still do not collect the full urban wage premium enjoyed by city natives.

The results displayed in the final column of Table 4 illustrate how these wages evolve over the longer term. The dependent variable is now the wage earned four years after a move ($\ln W_{t+4}$); the reference group remains those workers who lived in a non-urban area in years t and $t + 1$. The coefficient estimate for *City*(1) tells the same story as the previous two columns: workers living in cities collect wages that are 20 percent higher than non-urban workers. The point estimate for *In-City* is now measured at 0.170 and highly significant. This clearly shows that the wages of city entrants have grown much faster over this four-year period relative to city natives, virtually closing the wage gap identified in the year after a move. Indeed, by year $t + 4$, there is only a 0.018 log point difference in wages between city entrants and natives. The implied wage growth rate for the city entrants is also consistent with the estimate identified in the wage change equation in column 3.

Taken as a whole, Table 4 offers a rather complicated picture of the pecuniary returns to urban residency. It is clear that city natives collect wages some 20 percent higher than workers in non-urban areas. However, they appear to receive little advantage in terms of wage growth. Workers migrating into cities do receive significant wage gains at the time of the move, suggesting that a portion of the urban wage premium is a wage level effect. Indeed, the symmetry identified in the wage changes for city entrants and city leavers is consistent with the firm-level productivity hypothesis. Yet if this were the only effect at play, then recent city entrants would collect the whole of the urban wage premium upon entry into the city. Instead, their wages are well below the level of city natives in the first year post migration. These recent migrants do experience a

superior growth rate in wages over the ensuing four years, nearly closing the gap in wages with city natives. This "catch-up" effect suggests that an important part of the urban wage premium is then also a wage growth effect. It is important to note that this wage trend for recent city entrants is similar to the type of catch-up effect identified in Borjas et al. [2] for interstate migrants. As such, this effect is not unique among urban migrants, though it may very well be strongest in cities. Finally, these results offer solid evidence in favor of something other than ability-sorting and firm-level productivity enhancement as explanations of the residual urban wage difference.[18]

4.2.2. Between-job wage growth

Given what appears to be an important growth effect in the wages of urban workers, especially for recent city entrants, it becomes necessary to ascertain whether this growth relates more strongly to job tenure (within-job wage growth) or job mobility (between-job wage growth). Although Glaeser [12] provides some empirical support for the importance of learning on the job for urban workers, the role of job mobility in the wage growth process has not been examined empirically. To do so, I compare the returns to job changing in cities against that in non-urban areas. Given Topel and Ward [29] report that one third of all wage growth in the first ten years of the working career occurs via job change, the importance of such analysis cannot be overstated.

The analysis requires measurement of the between-job wage growth accompanying a change of jobs. Consider a worker transitioning from job $j - 1$ to job j at time t (assuming a negligible period of unemployment between jobs). The between-job wage change can be calculated straightforwardly as the difference between the "stop wage" on job $j - 1$, denoted $w_{j-1,t}$, and the "start wage" on job j, denoted $w_{j,t}$, where w refers to the natural log of the real hourly wage. Unfortunately, while the NLSY79 does provide the true stop wage at time t, the survey instrument does not collect the starting wage for each job. What is available for job j is the current wage at the time of the next survey. Denote this first wage observation for job j as $w_{j,t+1}$. Although one could use the difference between $w_{j,t+1}$ and $w_{j-1,t}$ as an estimate of between-job wage growth, it would overstate the actual change in wages experienced across jobs. The reason is that $w_{j,t+1}$ includes any within-job wage growth occurring between the date of hire (time t) and the survey date in which the wage observation was reported (time $t + 1$).

Applying a variant of the estimator suggested by Topel and Ward [29], an approximate between-job wage change can be estimated as

$$\mathrm{E}(w_{j,t} - w_{j-1,t} | w_{j,t+1}, w_{j-1,t}) = w_{j,t+1} - w_{j-1,t} - \mathrm{E}(w_{j,t+1} - w_{j,t} | w_{j,t+1}, w_{j,T}). \quad (6)$$

The first term on the right-hand side is the difference between the two available wage observations: the first reported wage on job j ($w_{j,t+1}$) and the true stop wage on job $j - 1$ ($w_{j-1,t}$). The last term is the expected wage growth occurring on job j between the actual start date (time t) and the date at which the first wage observation is recorded (time $t + 1$). Because this latter term

---

[18] Glaeser and Maré [14] also examine wage growth. Rather than using a differenced wage model, they estimate wage level models with dummy variables pegging observations to the year of a move. The advantage of the wage level approach is that it preserves significantly more variation, allowing for more precise estimates. The disadvantage is a less clear-cut counterfactual in fixed-effects estimation, since workers always living in cities cannot be distinguished from workers always residing outside of urban areas. They find that rural-to-urban migrants experience wage gains of 15 percent over the first few years post migration. However, after more than five years they still earn wages below the level of city natives. In contrast, after starting with wages that were significantly below urban stayers, urban-to-rural migrants see their relative earnings drop by between 1 and 5 percent. Interestingly, in the few years after the move, they show little in the way of relative wage improvements.

Table 5
Estimates of between-job wage growth at job transitions

Residence	Total	Experience interval (years)			Schooling level	
		0–4	4–8	> 8	HGC > 12	HGC ⩽ 12
Big City	0.090	0.126	0.086	0.056	0.121	0.079
	(0.008)	(0.013)	(0.013)	(0.015)	(0.016)	(0.009)
Small City	0.082	0.108	0.078	0.038	0.098	0.078
	(0.011)	(0.020)	(0.019)	(0.018)	(0.024)	(0.013)
Non-urban	0.079	0.121	0.063	0.030	0.102	0.074
	(0.013)	(0.020)	(0.024)	(0.023)	(0.033)	(0.014)
Rural	0.066	0.085	0.067	0.038	0.096	0.060
	(0.018)	(0.028)	(0.036)	(0.027)	(0.048)	(0.019)

Notes. Standard errors are in parentheses. Estimates are log difference in wages across jobs adjusted for within-job wage growth as defined in Eq. (6). *Big City* refers to residence in an MSA with population > 1 million residents; *Small City* refers to residence in an MSA with population > 250,000 but less than 1 million residents; *Non-urban* indicates residence in an area with less than 250,000 inhabitants; *Rural* refers to residence in an area with less than 100,000 in population. *HGC* refers to highest grade of schooling completed.

is not observed, it must first be estimated and then subtracted out in order to achieve the "true" between-job wage change.[19]

Table 5 presents the average between-job wage growth experienced by workers changing jobs within large MSAs (*Big City*), small MSAs (*Small City*), non-urban areas (*Non-urban*), and non-urban areas with populations less than 100,000 (*Rural*). Column 1 shows that individuals changing jobs within large cities experience the largest average wage gain, around 9 percent. Workers changing jobs within small cities receive gains of 8.2 percent. This can be contrasted against the returns to job changing in non-urban areas. For all non-urban workers, the typical wage growth is only 7.9 percent; for workers in rural areas, the wage gains are only 6.6 percent on average. The next three columns of the table explore how between-job wage growth varies with years of labor market experience. For workers with fewer than four years of experience, the average gain from job changing is 12.6 percent for workers living in big cities. In contrast, workers in rural areas with similar experience collect wage gains of only 8.5 percent. For workers with experience levels between 4 and 8 years, those changing employers within big cities receive wage gains of 8.6 percent, whereas non-urban job changers collect average returns of only 6.3 percent. Workers with experience levels of more than eight years receive wage gains of 5.6 percent in big cities but only 3 percent in non-urban areas. The final two columns of the table show the breakdown for workers with completed schooling levels of more than 12 years (*HGC* > 12) and twelve or less years (*HGC* ⩽ 12). Workers in the high schooling category experience wage improvements of 12.1 percent when changing jobs within big cities but only 10.2 percent in non-urban areas; workers in the low schooling category experience between-job wage growth of 7.9 percent in big cities and 7.4 percent in non-urban areas.

---

[19] Before Eq. (6) can be used to construct measures of between-job wage growth, I must first obtain an estimate of $E(w_{j,t+1} - w_{j,t} | w_{j,t+1}, w_{j,T})$. To do this, I estimate a simple model of within-job wage growth by regressing the difference between the first ($w_{j,t+1}$) and last wage observation ($w_{j,T}$) for each job j on differences in experience and tenure (higher-order terms only). Retaining the coefficients, I use them to construct estimates of the within-job wage growth that would have occurred on job j between the start date and the date of the first wage observation. These predicted values are then used to adjust the measured difference in wages across jobs $j - 1$ and j to arrive at a credible estimate of between-job wage growth.

Table 6
Frequency of job change by residence and location experience

Residence	N	Total	Consecutive years in location				
			1	2	3	4	5
Big City	2592	0.268	0.309	0.285	0.273	0.247	0.224
Small City	1074	0.220	0.326	0.199	0.202	0.169	0.182
Non-urban	875	0.202	0.291	0.195	0.175	0.130	0.145
Rural	411	0.180	0.295	0.146	0.182	0.117	0.120

Notes. Estimate refers to the percentage of respondents changing employers between interview dates. *Big City* refers to residence in an MSA with population > 1 million residents; *Small City* refers to residence in an MSA with population > 250,000 but less than 1 million residents; *Non-urban* indicates residence in an area with less than 250,000 inhabitants; *Rural* refers to residence in an area with less than 100,000 in population.

Although Table 5 appears to provide evidence supporting the notion that workers changing employers within cities (especially large cities) receive superior between-job wage growth, I found no statistical difference between urban and non-urban workers in the average wage gain associated with a single job change.[20] However, it is still possible for the cumulative effect to be significant if workers in cities change jobs more often than workers in non-urban areas. I test for this possibility by examining the frequency of job change for urban and non-urban workers at similar points in the working life-cycle.[21] Table 6 shows the frequency of job change by residency status and cumulative location experience. Over all experience categories (the column labeled *Total*), workers residing in big cities changed jobs between survey dates 26.8 percent of the time. In contrast, workers residing outside of urban areas change jobs in 20.2 percent of the observations, and only 18 percent in rural areas. Looking across cumulative years of location experience, an interesting pattern emerges. For workers with one year of location experience, the frequency of job change between urban and non-urban workers is roughly the same (31.5 percent to 29.1 percent). As one might expect, mobility declines with experience for both residential categories. However, the decline is much steeper for workers in non-urban areas. Workers living in big cities for two consecutive years change jobs 28.5 percent of the time whereas workers living in non-urban areas for two years have mobility rates of only 19.5 percent—a 9 percentage point difference. This difference expands to an 11.7 percent difference for workers with four consecutive years of location experience (with big city workers changing jobs 24.7 percent of the time and non-urban workers changing only 13.0 percent). The gap closes somewhat for workers with five consecutive years of location experience, but the overall trend is clear.

These results are confirmed in probit models of the likelihood of a job change. Table 7 presents the estimated coefficients, standard errors, and marginal effects for the indicators of urban resi-

---

[20] I first conducted simple *t*-tests of the difference in mean between-job wage growth and found no statistically significant difference. Next, I regressed between-job wage growth on indicators of urban residency and a standard set of controls across a variety of specifications. Regardless of specification, urban residency was never found to be a statistically significant determinant of between-job wage growth.

[21] To do so, I construct a sample consisting of those respondents with valid location information for the five consecutive surveys subsequent to their final exit from formal schooling. This allows me to calculate cumulative location experience for all respondents in the sample over the first five years of their career. Job changes are then identified between annual survey dates and frequencies tabulated for each location-experience category. Implementing these selection criteria generates a sample of 4541 person-year observations contributed by 1105 young men.

Table 7
Coefficients and marginal effects for residency indicators in probit models of the likelihood of job change

Residence	(1)	(2)	(3)
City	0.190[a]		
	[0.055]		
	(0.056)		
Big City		0.235[a]	
		[0.070]	
		(0.059)	
Small City		0.081	
		[0.025]	
		(0.068)	
Rural			−0.281[a]
			[−0.078]
			(0.080)
N	4541	4541	4541

Notes. Marginal effects are in square brackets. Standard errors are in parentheses. Standard errors are corrected for heteroskedasticity and within-person correlation. A job change means that the respondent changed employers between interview dates. *City* refers to residence in an MSA with more than 250,000 inhabitants; *Big City* refers to residence in an MSA with population > 1 million residents; *Small City* refers to residence in an MSA with population > 250,000 but less than 1 million residents; *Rural* refers to residence in an area with less than 100,000 in population. Probit models also include race, highest grade attended, years of potential experience, current job tenure, public sector employment, marital status, union or collective bargaining status, AFQT, and year dummies.

[a] Significant at the 1% level.

dency.[22] Column (1) shows the impact of urban residency (*City*) on job change. The estimate is positive and significant at the one percent significance level, with the marginal effect implying that urban residents are 5.6 percent more likely than non-urban residents to change jobs between interview dates. Column (2) shows the effect of city size on the likelihood of job change using indicators that distinguish between residency in big and small cities. The coefficient estimate for the *Big City* indicator is found to be positive and highly significant. The marginal effect indicates that residents in large urban areas are 7 percent more likely to change jobs. The insignificant coefficient estimate for *Small City*, however, suggests that residents of smaller cities are no more likely to change jobs between interview dates than non-urban workers. Column (3) shows the likelihood of job change for workers living in rural areas relative to all other workers. Interestingly, residents in rural areas are significantly less likely to change jobs than workers living in MSAs of any size.

In sum, although job-to-job wage growth is no larger in cities, the more-frequent changes that occur may lead to more rapid wage growth over time for urban workers. Not only is this pattern of job changing consistent with the coordination hypothesis, but it also fits the empirical finding of a lagged pecuniary return to urban residence. As discussed earlier, formal search theory suggests that drawing from a superior wage offer distribution, realizing a higher arrival rate of job offers, and/or having lower search costs increases the expected net benefits of search, making both search and job change more likely. Each is more likely to be true for workers searching in

---

[22] The probit models include controls for race, highest grade of schooling attended, years of potential work experience, current job tenure, public sector employment, marital status, union or collective bargaining status, AFQT, and year dummies. Because individuals contribute multiple observations to the sample, the standard errors are corrected for both heteroskedasticity and within-person correlation.

dense urban labor markets. While the evidence on job turnover presented above is by no means conclusive, it does suggest a reasonable explanation for the remaining (time-dependent) urban wage premium.

5. Conclusions

The object of this analysis has been to explore the sources of the significant wage premium realized by urban workers. Ordinary least squares regressions that control for measurable worker characteristics reveal a 19 percent wage difference between workers in large urban areas and non-urban residents. Fixed-effects panel estimates indicate that about two-thirds of this wage premium can be explained by cities attracting workers of higher unmeasured skills and ability. This result remained robust even when controlling for inter-personal differences in the strength of experience and urban cost of living.

Wage growth analysis revealed a distinct symmetry in the wage changes between in-city and out-city migrants, a find consistent with the "productivity" hypothesis. It was also shown that recent migrants to cities do not receive the bulk of the urban wage premium upon arrival but only after time spent in the city. The two leading explanations of this time-dependent urban wage premium are the "learning" and "coordination" hypotheses. Although I found no statistical difference in the between-job wage growth experienced by urban and non-urban workers as predicted by the coordination hypothesis, the pattern of job change identified in this study is suggestive of a cumulative advantage in the pecuniary return to job change for urban workers. If so, coordination efficiencies in dense urban settings have a prominent role to play in any comprehensive explanation of the urban wage premium.

Acknowledgments

References

[1] G.S. Becker, K.M. Murphy, The division of labor, coordination costs, and knowledge, Quarterly Journal of Economics 107 (1992) 1137–1160.

[2] G. Borjas, S.G. Bronars, S.J. Trejo, Assimilation and the earnings of young internal migrants, Review of Economics and Statistics (1992) 170–175.

[3] G. Borjas, S.G. Bronars, S.J. Trejo, Self-selection and internal migration in the United States, Journal of Urban Economics 32 (1992) 159–185.

[4] C. Brown, J. Medoff, The employer size–wage effect, Journal of Political Economy 97 (1989) 1027–1059.

[5] K. Burdett, A theory of employed job search and quit rates, American Economic Review 68 (1978) 212–220.

[6] G.A. Carlino, Trends in metropolitan employment growth, Federal Reserve Bank of Philadelphia Business Review (1998) 13–22.

[7] Census 2000 Summary File 1 (SF 1), GCT-P1 urban/rural and metropolitan/nonmetropolitan population, US Census Bureau, 2000.

[8] A. Ciccone, R.E. Hall, Productivity and the density of economic activity, American Economic Review 86 (1996) 54–70.

[9] J.M. DuMond, B.T. Hirsch, D.A. MacPherson, Wage differentials across labor markets and workers: Does cost of living matter? Economic Inquiry 37 (1999) 577–598.

[10] V. Fuchs, Differentials in hourly earnings by region and city size, 1959, Occasional paper 101, NBER, New York, 1967.

[11] E.L. Glaeser, Are cities dying, Journal of Economic Perspectives 12 (1998) 139–160.

[12] E.L. Glaeser, Learning in cities, Journal of Urban Economics 46 (1999) 254–277.

[13] E.L. Glaeser, D.C. Maré, Cities and Skills, Working paper No. 4728, NBER, 1994.

[14] E.L. Glaeser, D.C. Maré, Cities and skills, Journal of Labor Economics 19 (2001) 316–342.

[15] E.L. Glaeser, J. Shapiro, Is there a new urbanism? The growth of US cities in the 1990s, Discussion paper No 1925, Harvard Institute of Economic Research, 2001.

[16] M. Granovetter, Getting a Job: A Study of Contacts and Careers, second ed., Univ. of Chicago Press, Chicago, 1995.

[17] R.W. Helsey, W.C. Strange, Matching and agglomeration economies in a system of cities, Regional Science and Urban Economics 28 (1990) 189–212.

[18] H. Holzer, Search method use by unemployed youth, Journal of Labor Economics 6 (1988) 1–20.

[19] G.E. Johnson, Intermetropolitan wage differentials in the United States, in: J.E. Triplett (Ed.), The Measurement of Labor Cost, Univ. of Chicago Press, Chicago and London, 1983, pp. 309–330.

[20] P. Katz (Ed.), The New Urbanism: Toward an Architecture of Community, McGraw–Hill, New York, 1994.

[21] S. Kim, Diversity in urban labor markets and agglomeration economies, Papers of the Regional Science Association 62 (1987) 57–70.

[22] S. Kim, Labor heterogeneity, wage bargaining, and agglomeration economies, Journal of Urban Economics 28 (1990) 160–177.

[23] S.A. Lippman, J.J. McCall, The economics of job search: A survey: Part I, Economic Inquiry 14 (1976) 155–189.

[24] E. Moretti, Workers' education, spillovers and productivity: Evidence from plant-level production functions, American Economic Review 94 (2004) 656–690.

[25] D.T. Mortensen, Job search and labor market analysis, in: O. Ashenfelter, R. Layard (Eds.), Handbook of Labor Economics, vol. 2, North-Holland, Amsterdam, 1986, pp. 849–920.

[26] J.E. Rauch, Productivity gains from geographic concentration of human capital: Evidence from the cities, Journal of Urban Economics 34 (1993) 380–400.

[27] Y. Sato, Labor heterogeneity in an urban labor market, Journal of Urban Economics 50 (2001) 313–337.

[28] G. Topa, Social interactions, local spillovers and unemployment, Review of Economic Studies 68 (2001) 261–295.

[29] R.H. Topel, M.P. Ward, Job mobility and the careers of young men, Quarterly Journal of Economics 108 (1992) 439–479.

[30] B. Weinberg, P. Reagan, J. Yankow, Do neighborhoods affect hours worked? Evidence from longitudinal data, Journal of Labor Economics 22 (2004) 891–924.