City Air and City Markets:
Worker productivity gains across city sizes.

Douglas J. Krupka[1]; Douglas S. Noonan[2]

[1]Institute for Research on Labor, University of Michigan, Ann Arbor, MI, USA

[2]School of Public Policy, Georgia Institute of Technology, Atlanta, GA, USA

Abstract: Higher nominal wages in urban areas are a well-documented phenomenon which implies higher productivity of urban workers. Yankow (2006) and Wheeler (2006) show that these gains come through a variety of sources including static agglomeration economies, and dynamic learning and matching efficiencies in cities. Yet earlier papers offer little evidence of how the effects of learning and matching on urban wage differentials vary by city size. This paper allows for the relative importance of these productivity advantages to differ according to the size of the city, and finds significant differences between small, medium and large cities. I find that learning efficiencies are most important in medium-sized cities, while a mix of learning and matching efficiencies are important in the largest and smallest cities.

I. Introduction

An old German saying goes "*Stadtluft macht frei*."[1] Although the original saying concerned feudal obligations, the idea that city air is somehow different persists both in the broader culture and – more recently – in the urban economic literature. Glaeser (1999) and Glaeser and Maré (2001) offer theory and evidence suggesting that urban areas increase wages through a process of learning. However, Glaeser (1999) acknowledges that the pattern of evidence found in Glaeser and Maré (2001) is also consistent with the potential of better labor market matches in dense urban labor markets, which are realized gradually through more intensive or productive job search. Indeed, Wheeler (2006) and Yankow (2006) find evidence supporting this hypothesis.

In terms of our understanding of cities and their economic function the two theories offer very different views. While Glaeser (1999) makes a persuasive defense of his assumptions, the learning externality essentially assumes that there is something different about cities in the Marshallian or Jacobsian way: that the urban atmosphere somehow imparts knowledge to those who breathe it. On the other hand, the matching mechanism for enhanced productivity of urban workers requires only that city labor markets be thicker on both the supply and demand side, which is something we already know to be true. However, the relative importance of these urban wage growth effects need not be uniform across all locations. Most prior work, like Glaeser and Maré (2001), Wheeler (2006) and Yankow (2006), offer little evidence on how these effects on growth differ by city size. This paper extends these analyses to examine how the relative importance of agglomeration, learning and matching economies varies with city size.

---

[1] "City air liberates," in German.

1

The rest of the paper is organized as follows. Part II briefly discusses several explanations for urban wage premia and differential urban wage growth. Sections III and IV discuss the empirical model and describe the data. Section V presents the results, which are consistent with the earlier work by Wheeler (2006) and Yankow (2006), but which expose significant differences in the sources of urban wage premia across city sizes. These results suggest that the functions of cities vary with their size. Middle-sized cities seem to generate the largest learning effects, while small and large cities are characterized by a mix of matching and learning effects. Section VI comments on these findings and Section VII concludes.

II. Background

Wages are higher in urban areas, and higher still in larger urban areas. From the perspective of workers, it is not difficult to see why this would be the case. To attract labor supply to large cities, employers must compensate them for the high cost of living and possibly also for congestion externalities associated with these areas. Looking at the question from the labor demand side does not yield as obvious an explanation. Why should profit-maximizing firms be willing to pay workers more in large cities than in smaller cities or rural areas? Since the costs of doing business *besides* wages (land costs, congestion costs, regulatory costs and taxes, for instance) are also higher in urbanized areas, it must be the case that workers in large cities are more productive than workers in smaller cities. Why?

There are several possible answers. Following Glaeser and Resseger (2010), I categorize these theories as knowledge-based and not knowledge-based. The former include the two theories of interest here – city air (learning) and city markets (matching) – discussed below. The other theories include many intuitive

explanations. It could be the case that city workers are more able than rural workers. Sorting of more productive workers into larger cities could explain the wage premium, but would require explanation itself. If urban residents earn more because of their talents, and not because of their residence, then they should be able to move to the country or a smaller city (with lower costs of living) and have a higher material quality of life. For the ability-sorting explanation to hold up in locational equilibrium, high-skill individuals would have to have an unobserved taste for urban areas. Heuermann et al. (2010) summarize the several studies measuring the extent to which sorting accounts for the urban wage premium, concluding generally that worker sorting accounts for much but not all of the premium. Glaeser and Maré (2001) and Yankow (2006) find controlling for worker fixed effects reduces the premium by over half. Sorting on observable skill and unobservables continue to receive attention in the literature (see, e.g., Adamson et al. 2004, Gould 2007, Baum-Snow and Pavan forthcoming).

It is also possible that urban wages are higher because of the large capital stocks that are at workers' disposal in urban employment, relative to small towns or rural areas. Under normal assumptions, larger capital stocks increase the marginal product of labor, and thus generate higher wages in short-run equilibrium. However, in the long run the capital stock is endogenous. If capital is mobile, it is hard to see why employers would pay workers more in cities when lower wages could be paid to rural workers had the capital been invested in the rural area instead. The empirical implications of a capital-deepening explanation of urban wage premia are that migrants to urban areas should experience immediate wage gains, and urban-rural migrants should experience immediate wage losses. Glaeser and Maré (2001) find instead that urban migrants appear to increase their wage gradually, while the wage

losses of urban-rural migrants are not persistent. Furthermore, they find that the urban wage premium is largest for workers with the most experience, an effect that does not seem consistent with a straight capital-deepening story. Glaeser and Resseger (2010) explore this possibility further, finding no evidence to support this hypothesis.

A related explanation for the urban wage premium is that urban firms benefit from large stocks of unpriced inputs like public infrastructure. These roads, schools, ports and utilities could increase the productivity of firms that have access to them. If there are economies of scale in the provision of such inputs, then dense urban areas could retain their advantage over rural areas and the urban wage premium could constitute a locational equilibrium. Similar effects could derive from un-produced productive amenities such as coastal location or access to natural resources. However, the impact of public capital on business productivity is in dispute, with some papers finding no net effect of public infrastructure expenditures (Holtz-Eakin and Lovely 1996). Even if public infrastructure is productive, this source of the urban wage premium would have empirical implications identical to those of the private capital story above (immediate wage gains for urban in-migrants, immediate wage losses to urban out-migrants), so the evidence cited above about gradual wage gains also undermines a public-infrastructure- or productive-amenity-based explanation for urban wage premia.

The remaining explanations for the urban wage premium are all forms of agglomeration economies.[2] Fujita et al.(1999) describe the three major agglomeration economies, which they attribute to Marshall, as access to customers, access to specialized inputs and human capital spillovers. Krugman (1991) models the first of these types of agglomeration economies. While Krugman (1991) is at the head of a

---

[2] From this point forward, all the models derive results that stand up in locational equilibrium.

4

very exciting literature,[3] the implications for the urban wage premium are similar to the public and private capital interpretations, above. In these models, transport costs plus economies of scale make location near consumers valuable. Workers also find it attractive to locate near the production centers because that lowers their cost of living.[4] In all of this, the agglomeration forces rest completely within the firm. It is the firm's access to consumers and scale economies which give rise to the agglomeration tendencies. Mobile workers are homogenous. If this is the case, then any mobile worker can move to the site of agglomeration, take a job and earn a wage as high as any long-term resident of the agglomerated area. It is not the workers, but the firm and the agglomeration that raise productivity. Thus, the predictions of immediate wage gains and losses persist.[5]

The knowledge-based theories of agglomeration economies often start with human capital spillovers, as Glaeser (1999) models. He includes the oft-quoted passage from Marshall that in cities the "mysteries of the trade become no mystery: but are as it were, in the air." It is not simply the air's saturation with trade secrets that increases productivity in cities, however. In Glaeser's model learning is achieved through interactions with more skilled individuals. Cities foster more learning than rural areas because interactions are more frequent in cities. Given certain simplifying assumptions, Glaeser derives privately and socially optimal city sizes and skill distributions. Although the theory is in the context of long-run equilibrium, Glaeser draws out the dynamic implications for the urban wage premium. If urban workers

---

[3] Ottaviano and Thisse (2004) survey the theoretical literature which has followed the Krugman (1991) paper, while Head and Mayer (2004) survey the empirical literature.
[4] These models assume away costs to urban congestion. The model has been extended and modified to allow for such agglomeration costs by Tabuchi (1998) and Ottaviano et al. (2002), with no major change in the qualitative behavior of the model.
[5] There are other microsources of agglomeration economies besides proximity to consumers. Duranton and Puga (2004) discuss these from a theoretical view point, and Rosenthal and Strange (2004) review the empirical literature. For the purposes of this paper, many of these theories will be observationally equivalent. This paper focuses on two of these micro-foundations which are discussed in some detail below.

earn more because they have learned and are learning through interactions with peers, two implications are obvious. First, the wage premium should develop gradually as migrants to cities learn their skills from their seniors. Second, the urban wage premium should increase with experience.[6] Glaeser and Maré (2001) take these predictions to several datasets, and find evidence that broadly supports the learning hypothesis. Results found by other authors support this interpretation as well.[7] Thus, while Krugman (1991) and the consequent new economic geography literature model Marshall's first agglomeration economy, Glaeser (1999) models his third.

What Glaeser (1999) acknowledges is that these same patterns could be explained by Marshall's *second* source of agglomeration economies: thicker markets for specialized inputs. Specifically, if firms require certain kinds of labor, then having a larger labor market should improve the quality of employment matches. Helsley and Strange (1990) model this kind of labor-matching, finding that privately optimal cities would be too large (have too many firms). For simplicity, Helsley and Strange assume a kind of perfect information: once a person moves to a city, it is costless for him to find the employer that will provide the highest quality match.[8] This is probably a reasonably accurate description of a worker's situation in the long run. However, in the short-run, workers probably take some time to search out and discover this best match. This near-term search will be characterized by the kind of

---

[6] Knowledge spillovers that occur between firms, such as those in the nursery cities model of Duranton and Puga (2001), would not cause the kinds of effects predicted by Glaeser (1999). However, such spillovers, which make it optimal for firms to undertake research and development activities in agglomerations and then move mass-production activities to rural areas, might be expected to increase the returns to education in large cities. I consider this possibility in the empirical section below.

[7] Charlot and Duranton (2004) model communication and wage simultaneously and find that communication affects wages, at least among the French. Rauch (1993) finds that the average education in an SMSA increases wages and rents about equiproportionately, which is consistent with metropolitan human capital increasing productivity, but not being an amenity. Glaeser and Resseger (2010) go further in exploring this connection. Gould (2007) and Baum-Snow and Pavan (forthcoming) relate their findings to learning among white-collar workers and large-city residents, respectively.

[8] Teulings and Gautier (2003) model search friction more explicitly in a search model with exogenous urban scale, but do not draw out the dynamics of the model with job changes.

rapid job turnover documented in Topel and Ward (1992) and in the urban context by Wheeler (2008) and Finney and Kohlhase (2008). This period of searching for better matches and accruing gradual wage increases through job mobility could take some years (indeed, an optimist might believe that there is always a better match somewhere out there). Thus, the Helsley and Strange (1990) model of agglomeration arising from thick city input markets or better matching would give rise to wage patterns very similar to the learning model in Glaeser (1999): wage premia should accrue gradually and should be increasing with time spent in the city.

This paper addresses the mechanisms by which these wage premia gradually accrue to city dwellers. The observed urban wage premium may arise from wage level effects or via wage growth rate effects, or both. The learning and matching hypotheses both predict urban wage gains over time, while other theories suggest immediate wage differentials. I am particularly interested in the growth aspect, getting closer to the microfoundations of how the growth differential comes about in cities (learning vs. matching), and whether the roles of "city air" and "city markets" in generating this differential vary with the scale of the metropolitan area. Most prior work that evaluates the learning and the matching hypotheses either restrict cities' effect on wage growth (via learning or matching) to be the same regardless of city size (e.g., Glaeser and Maré 2001, Gould 2007) or impose a monotonic relationship between city size and these wage dynamics (e.g., Wheeler 2006, Glaeser and Resseger 2010). The estimation approach taken here improves on this by including flexible city-size interactions.

The prior literature differentiating between the learning and matching hypotheses across city size is quite limited. Empirical analyses distinguishes between the two hypotheses by separating 'within-job' wage growth from 'between-job' wage

growth. Although imperfectly, within-job raises reflect learning while between-job wage gains support the matching or coordination hypothesis. Wheeler (2006) and Yankow (2006) examine urban wage premia and urban wage growth effects and find evidence for both the learning and the matching efficiency arguments: urban residents earn more because their wages grow faster, and their wages grow faster in part because job changes are more frequent and more rewarding financially.[9] Yet the accumulated evidence on learning vs. matching across urban scales remains cloudy. For between-job wage gains, Wheeler (2006) finds significant urban effects but Yankow (2006) does not. Moreover, some of Wheeler's models and all of his fixed-effects models fail to find significant effects of urbanicity on between-job wage gains. For within-job wage growth, some studies find significantly higher rates in cities (Glaeser and Maré 2001, Glaeser and Resseger 2010), but others show no association between growth rates and city population (Glaeser and Resseger 2010, Wheeler 2006). The different measures of urban size used by these researchers suggest that the relationship between city size and learning and matching may be rather complicated.

In this paper, I reproduce these earlier results and show how the urban wage premia – in particular, that part deriving from differential wage growth rates – depend on the place of the city within the urban hierarchy. This focus on the interactions between city size and work experience and job mobility contrasts with the emphases of other recent studies of the urban wage premium, which focus on other themes like wage distributions (Matano and Naticchioni 2011), sorting (Gould 2007, Baum-Snow and Pavan forthcoming), and education (Abramson et al. 2004). The results here give a relatively direct test of the learning and matching hypotheses behind the urban wage

---

[9] Other recent papers offer some evidence on the learning and matching theories such as the transferability of city learning back to rural areas (Gould 2007), the frequency of changing jobs over a career (Wheeler 2008) and how age-earnings profiles are steeper in highly educated cities (Glaeser and Resseger 2010).

premium. Incorporating both in the same model also offers a measure of their relative contributions, and how their importance varies with city size. The estimation approach and larger dataset used here yield more precise estimates than previous studies. These new results provide further evidence for this important research area which has thus far produced a good deal of insignificant and inconsistent results. This inquiry improves our understanding of the dynamics and implications of urban growth. This could allow for economic development and employment policies to be differentiated by city size. How cities' roles in worker productivity – as a forum for matching (city markets) or learning (city air) – vary with size can guide urban policy to foster the sorts of economic environments most suited a city's scale.

III. Empirical Framework

The presence of an urban wage premium can be confirmed quite simply by regressing individual i's log wage at time t on his contemporaneous urban residence, as in equation (1):

1) $\quad \ln(Wage_{it}) = \beta_0 + \beta_{Urb}Urban_{it} + \varepsilon_{1it}.$

While $\beta_{Urb}>0$ confirms the existence of an urban wage premium, it does nothing to explain it. The possibility, discussed above, that this premium is an artifact of observably more able workers sorting into cities could be tested by including individual characteristics in the equation:

2) $\quad \ln(Wage_{it}) = \beta_0' + \beta_{Urb}'Urban_{it} + \alpha X_{it} + \gamma Y_i + \varepsilon_{2it},$

where X is a vector of time-varying individual characteristics, and Y is a vector of time-invariant characteristics. If individual characteristics completely account for the urban wage premium, the addition of these characteristics to the regression equation should drive β_{Urb}' to insignificance. Conversely, selection on observables may

explain part, but not the entire urban wage premium, in which case β'_{Urb} would be smaller than β_{Urb}, but still significantly greater than zero.

The addition of individual fixed effects (μ_i) to the equation allows for the possibility of selection into cities based on unobservable characteristics such as skill, ambition, work ethic or other wage- or productivity-enhancing characteristics unobserved by the econometrician.

3) $\ln(Wage_{it}) = \beta''_{Urb}Urban_{it} + \alpha'X_{it} + \mu_i + \varepsilon_{3it}$

Let fixed effects μ_i represent time-invariant, individual-specific productivity components. Again, if sorting based on observable and unobservable wage-earning ability explains the urban wage premium, β''_{Urb} will be insignificant. If such selection describes some but not all of the premium, we would expect β''_{Urb} to be significant and positive, but smaller than β'_{Urb} or β_{Urb}. If the β''_{Urb} is positive and significant, it means that something about cities (not individuals) makes people earn more there.

Differentiating between the several theories for the urban wage premium is the primary focus of our discussion. I follow Glaeser and Maré (2001) by using interactions to identify the different avenues of the urban effect in the wage equation.

4) $\ln(Wage_{it}) = \beta''_{Urb}Urban_{it} + \beta_{UrbExp}Urban_{it} * Exp_{it}$

$$+\beta_{UrbTen}Urban_{it} * Ten_{it}$$

$$+\beta_{UrbEdu}Urban_{it} * Edu_{it}$$

$$+\beta_{Exp}Exp_{it} + \beta_{Ten}Ten_{it} + \beta_{Edu}Edu_{it} + \tilde{\alpha}\tilde{X}_{it} + \mu'_i + \varepsilon_{4it}.$$

In equation (4), $\tilde{X}$ is a vector of time-varying individual characteristics that does not include experience (*Exp*), tenure at an individual firm (*Ten*) or education (*Edu*). As before, μ'_i captures individual-specific, time-invariant ability. To run through the predictions of the various models: the capital deepening, public infrastructure, productive amenity and agglomeration (access to output markets) explanations all

predict a significant urban level effect, even after controlling for observables, unobservables and the interaction terms. This means that $\beta'''_{Urb} + \beta_{UrbEdu}Edu_{Avg.}$ should be significantly greater than zero, or that someone with no urban work experience receives an immediate wage bonus for working in a city. One interpretation of the Duranton and Puga (2001) model is that the returns to education should be higher in large cities, which implies that $\beta_{UrbEdu}{>}0$.

The learning effect of cities on productivity should be apparent by an increased return to work experience in cities, relative to rural areas. But since tenure and experience will increase equally when an employee does not switch positions, a pure Glaeserian learning effect would be represented by $\beta_{UrbExp} > 0$ and $\beta_{UrbTen} = 0$ urban work experience increases wages whether this experience is gained at one employer ($Exp_{it} = Ten_{it}$), or several ($Exp_{it} > Ten_{it}$). On the other hand, if the differential urban wage growth comes about because of improved matches, then working at the same urban firm will not generate differential wage gains. Only by changing employers and sampling the distribution of jobs (which is thicker in cities) will urban residents be able to gain differentially compared to rural workers. Such job switching will be identified through increases in experience while tenure decreases. If the entire urban "experience" effect is driven by matching instead of learning, we would expect $\beta_{UrbTen} < 0 < \beta_{UrbExp}$ and $\beta_{UrbTen} + \beta_{UrbExp} = 0$. That is, urban work experience at one employer (such that experience and tenure increase together) does not cause wages to increase faster than they would in rural areas. Urban wage growth accruing from matching means that employees must switch employers to realize these gains (experience must increase while tenure decreases).

The two hypotheses are not mutually exclusive. If matching and learning are both important, we would expect $\beta_{UrbTen} < 0 < \beta_{UrbExp}$, but $\beta_{UrbTen} + \beta_{UrbExp} >$

0. This implies that there are increases in wage growth due to faster learning in cities, but that staying at one employer (and thus foregoing the possibility of a better match) diminishes the wage growth relative to rural workers. When this is the case, a rough statistic measuring the importance of learning relative to matching can be generated from the regression output: $(\beta_{UrbTen} + \beta_{UrbExp})/\beta_{UrbExp}$. As can be seen by plugging the predicted coefficients under the different hypotheses into the formula, this statistic will be equal to one given the "pure learning" results, zero given the "pure matching" results and somewhere in between if both learning and matching are important. While the exposition above presented only one variable for urban residence and its interactions below I will present results where urban areas are characterized according to their relative size in three categories. This categorization sheds light on the nature of the agglomeration economies acting at different levels of urbanization.

The fixed-effects estimator in equation (4) allows identification of the urban wage premium as both a wage level effect (β_{Urb}''') and a wage growth effect (β_{UrbExp}, β_{UrbTen}). Although the model is estimated for wage levels, the use of *Exp* and *Ten* in a fixed-effects model that identifies wage effects off of within-worker variation allows for an interpretation of those coefficients as 'how does an individual's wage change as time spent working increases.' I refer to these marginal effects of time spent working on wages as "wage growth" throughout the paper, even though the dependent variable is not wage changes as in Yankow (2006) and Wheeler (2006). Equation (4) is identified from migration in or out of cities of different sizes, changes in a worker's city size, more time spent working, changes in time worked at the current job, and variation in *Edu* and $\tilde{X}$. As Yankow (2006) notes, the fixed-effects model removes bias from correlated unobserved ability while the wage level approach has an added

advantage of retaining more variation in the data. Conversely, the counterfactual is less straightforward as identification comes from more than just migration.

Finally, one objection to the methodology used in this paper is that urban residence is not assigned randomly, and thus there is the possibility of endogeneity bias in these estimates. While the identification of the urban effects through within-observation variation addresses this to some extent, it is of course still likely that urban-rural and rural-urban migration is not exogenous, either. If all changes were from rural to urban, the effect of such endogeneity would be to bias all of the estimated urban and urban interaction coefficients away from zero: those who expect (or realize) a large urban wage premium will be more likely to move to cities. However, the flow of individuals from cities to rural areas will induce the opposite bias. These flows are of similar magnitude in our data (about 4,000 rural/urban moves versus about 3,300 urban/rural moves), so the bias is likely quite small. On the other hand, it is unlikely that experience, education, tenure, or any of the other time-varying variables are exogenous either. Ideally, one would find compelling instruments for all these variables. At least for the case of urban status, I do not feel such a compelling instrument exists in the NLSY data. The results to follow are thus based on simple Mincerian regressions, and best interpreted as a good, first cut at the question, hoping that most of the endogeneity in these variables is driven by time-invariant factors such as tastes, ability or family background.

IV. Data

I use the National Longitudinal Survey of Youth 1979 (NLSY79) restricted-use geocoded and work history files to estimate the above equations. The NLSY79 offers a long panel dataset of a nationally representative sample. The initial cohort of

12,686 people aged 14-22 interviewed in 1979 were surveyed annually through 1994 and then biennially after that. Unlike the many previous studies of the urban wage premium that use the NLSY79 survey for years 1979 – 1994, I extend the panel through 2004. My vector of control variables includes cumulative experience (in weeks), tenure with current employer (in weeks), age, sex, race, marital status (=1 for married individuals, =0 for all other individuals) and the percentile score on the AFQT. I also sometimes include a battery of occupation, industry and year indicators. I use data from the entire panel (1979-2004), with the requirement that a person-year work at least 35 hours per week, have valid geographic and occupational information and have a "reasonable" reported wage.[10] The industry codes come from either the 1970 SIC codes, or the revised 2000 codes, depending on the year of the survey (2002 and 2004 waves use the more current codes). I generated a hybrid code system grouping industries into 16 categories for all survey years.[11] Similarly, occupation codes are reported in 1970's codes for all survey years except 2002. Comparing the 1970 and 2000 census occupation codes, I made an educated guess about assigning 2000 codes into 1970's bins. The assignment rules for occupation and industry are available to those interested.

To ensure comparability across the sample, I only counted work experience gained after the age of 18.[12] The experience variable was generated by summing the reported number of weeks worked since the previous interview. The interaction

---

[10] Reasonability of the reported wage was determined in the following manner. For each individual, I computed their average reported log wage in years they reported working over the 25 years (21 panels) of data collection. I also computed the standard deviation of the reported log wage. From this, I computed a person-year specific z-score of the reported log of wage. Person-years whose z-score was greater in magnitude than 2 were dropped from the sample. This prevents the inclusion of observations from several individuals who (one time each) reported hourly wages in excess of $10,000. The results are not terribly sensitive to the inclusion of these observations.

[11] These industries are agriculture, mining, construction, manufacturing, transportation, utilities, information, wholesale, retail, FIRE, business services, health, education, personal services, other services and public administration.

[12] Age is measured approximately: birth year was subtracted from interview year. So, this measure is slightly noisy.

between urban status and experience was done during the data processing: it is not simply the product of contemporaneous urban status and cumulative work experience. Rather, this interaction was generated by re-computing the cumulative experience variable assigning zero to any weeks where work was done outside an urban area.[13] The interactions of tenure and education with urban status, on the other hand, were computed in the standard way.

Urban status of county of residence was derived from the reported county of residence, in conjunction with the USDA ERS's rural/urban continuum scale.[14] These data are convenient because they provide measures of urbanity derived from the 1970-2000 censuses. For the years 1979-1985, 1980 county urbanity was used. For 1986-1995, 1990 county urbanity codes were used. For the remainder of the data, urbanity codes from the 2000 census were used. The 2000 codes are presented in Figure 1. These codes allow for three gradations of urban status which I use throughout the rest of the paper. The most restrictive, *Big Urban*, includes only metropolitan areas with populations of more than one million (black shapes in Figure 1). The next category, *Medium & Big Urban*, includes all those large cities, plus counties in cities with populations greater than 250,000, but less than one million (dark grey shapes in Figure 1). The most inclusive, *Urban*, includes all metropolitan areas with populations exceeding 50,000. Table 1 presents the sample means for the sample of worker-years in the NLSY from 1979 through 2004, as well as some comparisons across sizes of urban areas. Note that the urban wage premium is apparent in the group averages: big cities residents make 18 log points more than medium city

---

[13] This process was necessarily an approximation. If a respondent worked 50 weeks in a year, 25 of them in a rural area, but other 25 (including the interview) in an urban area, the entire 50 weeks would be counted as "urban" experience. While this is unfortunate, it represents a large improvement over simply multiplying the entire cumulative experience by one if the last interview occurred in an urban area, which could wrongly count several years of rural work as urban.

[14] These data are available at http://www.ers.usda.gov/Data/RuralUrbanContinuumCodes/ .

residents, 23 log points more than small city residents and 34 log points more than

rural residents in this sample. It should be noted that the urban indicators are

constructed different than typical, mutually exclusive dummy variables. As Table 1

makes clear, big cities take a value of unity for all three urban scale variables while

small cities have only the *Urban* value equal to one. This construction makes it easy

to detect significant incremental effects of city size, this paper's emphasis, in the

results shown below.

V. Results

The existence of the urban wage premium is confirmed in the first column of Table 2,

which takes equation (1) to the data, with the three-fold distinction amongst

metropolitan areas.[15] The results show that, over the 1979-2004 period, full-time

workers in urban areas earn around a 12.5 percent wage premium,[16] while workers in

cities with populations greater than 250,000 earn about an additional 4.5 percent over

small cities, and workers in the largest cities (metropolitan areas with populations

over one million) earn an additional 20 percent over medium cities. While these

coefficients are highly significant and economically substantial, the predictive power

of this basic equation is very low. Columns 2 and 3 of Table 2 report estimates from

equation 2, which control for a variety of individual characteristics (column 2) and a

set of occupation, industry and year fixed effects (column 3). These characteristics

increase the predictive power substantially, have expected signs and are highly

---

[15] Measuring $Urban_{it}$ with three categorical variables for city size is just one specification out of many possibilities. Population, density, and industrial diversity can be found in the literature. Using three categories follows from Yankow (2006) and Baum-Snow and Pavan (forthcoming), who observe three size categories is an improvement over some previous work and population and density measures are highly correlated. Moreover, it allows for nonmonotonic relationships, which Glaeser and Resseger's (2010) findings suggest might be worth exploring.

[16] 12.5 percent is the percentage increase represented by the coefficient of 0.1176. It is computed through the formula $e^{0.1176}$ -1. All percentage effects for the urban dummy variables are reported using a parallel formula.

significant. Note that while the measures of experience (*Exp*) and tenure (*Ten*) are measured in weeks, the coefficients in tables 2 – 4 have been scaled to represent a year's worth of experience in general or at a single firm.[17] We see that the inclusion of these observable characteristics reduce but do not eliminate the urban wage premium. Observationally equivalent workers in urban areas make around five percent more than those in rural areas, with an additional four percent premium in medium sized or larger cities, and an additional 14 percent in the largest urban areas. This suggests that sorting into urban areas based on observables accounts for some of the urban wage premium, but not all of it.

Columns 4 – 7 of Table 2 report estimates from the individual fixed-effects regression (equation 3) with various additional controls for year, occupation and industry. The time-varying control variables retain their significance. However, the addition of the individual fixed effects further reduces the urban wage premium, suggesting again that some of the urban wage premium is an artifact of high ability individuals sorting into cities. In the smallest cities, the urban wage premium becomes insignificant. In medium-sized cities, the premium is driven down to only 2.5 percent in the model with industry, occupation and year controls, and the large-city premium is driven down to an additional 6.5 percent over medium-sized cities. Both these urban premia are highly significant. Interestingly, the addition of the year fixed effects reduces the medium-city premium (compare columns 4 and 5), while the addition of the occupation and industry effects lowers only the small-city premium (compare columns 4 and 6). The large-city premium is robust to the inclusion of

---

[17] They can thus be compared to the education coefficient, which measures the effect of a year's amount of schooling on log wages.

these additional effects.[18] The substantial reduction in the urban wage premium after the inclusion of the individual fixed effects means that sorting of unobservedly able earners into cities is an important component of the observed wage differentials in cities. Whether this sorting derives from a preference among the highly able to live in cities, or higher returns to unobservable ability in cities, these results cannot address. However, the medium- and large-city premia are still significant: the urban wage premium does not appear to be an artifact of sorting.

Table 3 introduces the experience, tenure and education interaction terms as in equation (4). In Table 3, urban residence is entered as a single dummy variable. For exposition's sake, regressions both with and without the interactions are reported.[19] Columns 1 and 2 treat the *Urb* dummy variable as *Urban*, so all metropolitan areas are considered urban. Columns 3 and 4 define *Urb* as *Medium & Big Urban*, and columns 5 and 6 code *Urb = Big Urban* (i.e., only cities with populations over one million). Across all measures of urban residence, urban residence, urban experience and urban tenure are significant, while the urban education interaction is insignificant. The tenure and experience interactions are such that they suggest both learning and matching are important factors in differential urban wage growth. While this broad pattern is consistent across definitions of urban residence,[20] the importance of learning relative to matching varies considerably (ranging from about .5 to .75, by our rough measure). These results are consistent with those reported in Yankow (2006) and Wheeler (2006), although the specification does not facilitate the comparison of effects across city sizes.

---

[18] This statement applies to the large-city premium over medium cities. The total premium for living in large cities relative to rural areas, which is the sum of the three "urban" coefficients, is thus reduced in the same way as the smaller-city premia are.

[19] All the models in table 3 include individual fixed effects, and year, industry and occupation fixed effects.

[20] The broad pattern for this coding of urban residence is robust to the exclusion of the year, occupation and industry controls, and to estimation using OLS instead of fixed effects.

Table 4 enters urban residence as a vector of three urban sizes as described above. The urban dummies are the same ones used in the previous tables, so that the total urbanization effect for the largest cities (and the total urban interaction for these cities) is thus the sum of all three urban coefficients (or urban interaction coefficients). Column 1 reports results from a fully-specified OLS model, while columns 2 – 5 report coefficients from fixed-effect models with various controls for industry, occupation and year.

Before turning to the learning vs. matching issue, the other results bear some description (focusing on the fixed-effects models). The main effects of experience, tenure and education have the expected signs. Factoring in the interaction terms, education, tenure and experience all have positive effects in rural areas and in all city sizes. Relative to rural areas, the additional urban return to education is either negative (for smaller cities) or insignificant (for the largest cities). Relative to small cities, however, the additional returns to education in the largest cities (*Education*Big Urban + Education*Medium & Big Urban*) are significantly greater than zero across all models in Table 4 at the 5 percent level or better. This result resembles that of Adamson et al. (2004). Overall, while there are some interesting patterns across city sizes, there is little evidence of enhanced returns to education in cities relative to rural areas. These data thus fail to support at least one interpretation of the Duranton and Puga (2001) "nursery cities" model.

Table 4 offers quite detailed description of how the sources of urban wage growth vary across city sizes. "Growth" here refers to the association of within-worker variation in wages and within-person variation in time spent working. While the results vary somewhat between OLS and fixed-effects estimates, within the fixed-effects models they are quite consistent. In the smallest cities, the urban tenure

coefficients are over 60 percent of the urban experience coefficients.[21] By our rough

measure, learning makes up only about a third of the urban wage growth effect. The

marginal effect of moving from a small city to a medium city is actually perverse in

terms of matching so that learning accounts for the entire urban wage growth effect in

medium cities.[22] However, in the largest cities, matching again becomes important.

In fact, the marginal increase in wage growth associated with moving from a medium

to a large city arises completely from the matching effect. The large-city tenure

coefficients are about the same magnitude as the large-city experience coefficients,

meaning that (relative to medium cities) experience gained at one firm in a large city

has no effect on wages.[23] The total urban growth effect in large cities (relative to

rural areas) is about 60 percent learning, 40 percent matching.[24]

VI. Discussion

These results can be interpreted in terms of urbanization and localization

economies. Assuming that Glaeserian learning occurs mostly within industries,[25] the

learning dynamic can be seen as a localization economy, while the matching dynamic

seems to represent an urbanization economy with the larger variety of jobs resulting

from the urbanization of a variety of firms and industries. The results suggest that

smaller cities are the result of strong urbanization economies, while medium-sized

---

[21] The significance of these urban-tenure interactions are marginal: in the fullest specification they are only significant at the .1 level. However, I am unable to reject the null hypothesis that the *Tenure* interaction is equal in magnitude to the experience interaction (*p-value* = .2853), so that I am unable to reject the pure matching hypothesis for the smallest urban areas.

[22] Learning actually accounts for over 100% of the wage growth effect by our rough measure, but not significantly more than 100% (*p-value* =.7345 for this test).

[23] The *p-value* of the test that the sum of the large city experience and tenure coefficients is zero (representing no learning relative to medium cities) is .4393.

[24] Both the pure matching and pure learning hypotheses can be rejected at the 1 percent significance level in the largest cities relative to rural areas.

[25] That is, human capital is fairly industry-specific. Glaeser (1999) suggests this through the inclusion of a parameter for the number of industries, although he does not make much of it in the final paper. Empirical evidence for industry-specific human capital in a job search setting can be found in Neal (1995, 1999).

cities are the result of deep localization economies and the largest urban areas are results of a mix of each. Compared to rural areas, there are a very wide variety of jobs in small cities, but these small cities do not offer the deep industry-specific knowledge pools from which young workers can learn much. In medium-sized cities, on the other hand, the specialization in one or few industries does not significantly broaden the array of possible jobs, but these cities host concentrated stocks of workers with deep knowledge in the relevant sub-sectors. Finally, relative to medium-sized cities, the largest cities offer intense urbanization economies. The largest cities offer every kind of employment imaginable, improving matches, but the increased variety does not engender additional learning (relative to medium-sized cities). That additional learning does not occur could be a result of the knowledge base in larger cities being spread widely across industries rather than more deeply, of congestion or rapidly declining returns in learning mechanisms or of time constraints. After several hundred thousand residents, adding more residents does not increase the likelihood of learning because most of a person's interactions occur among a smaller group of close friends, family and colleagues.

These results paint an interesting picture for cities. Localization advantages for learning persist across city scales but plateau at the largest city size. Urbanization advantages for matching manifest for small cities and are particularly strong in large cities (relative to smaller cities or rural areas). That the age-earnings profile grows steeper from small to large cities is expected and consistent with Baum-Snow and Pavan's (forthcoming) findings. The fading of the matching effect for medium-sized cities is also consistent with Yankow (2006), who shows negligible matching effects in moving from small to medium citiss and no increased propensity to change jobs in medium-sized markets compared to smaller towns. Relatively specialized medium-

sized cities offer better learning opportunities but thinner job markets, consistent with a version of Duranton and Puga's (2001) model where the largest (and smallest) cities are "nurseries" for innovation and specialized firms relocate to mid-sized cities to exploit that learning.

The difference across city sizes in the importance of urbanization vs. localization economies is also consistent with the differences in urban wage effects across education levels. Medium-sized cities concentrated in a few industries are likely production-oriented cities, with high demand for production workers (high-school graduates) but no substantial increased demand for high-skill workers. Highly urbanized cities, on the other hand, have greater demand for all sorts of laborers. Thus, we see that high-school graduates with no urban experience earn an urban wage premium in small, medium and large cities, while inexperienced college degree holders earn a premium only in the largest cities.

Taking the results from tables 2 - 4 together, the story emerges that wage differences (and thus productivity differences) across city sizes arise from a variety of factors. Sorting based on observable and unobservable characteristics explains a substantial amount of urban wage differences. Similarly, straight agglomeration economies are important across all urban sizes. For college graduates, large cities offer substantial wage premiums. Finally, there is evidence supporting enhanced learning along the lines of Glaeser (1999) in all cities sizes, while substantial productivity increases due to enhanced match quality along the lines of Helsley and Strange (1990) also seem quite important at least in the smallest and largest cities.

These results are consistent with many previous findings (e.g., Glaeser and Maré 2001, Gabe 2004, Adamson et al. 2004, Glaeser and Resseger 2010) and offer some new results in two ways. First, some significant effects are found where

previous studies (Yankow 2006, Wheeler 2006) found insignificant results. This may be due to a longer panel or other data construction differences, alternative model specifications, or different identifying assumptions. Violations of the fixed-effects assumptions and different counterfactuals may also account for some of the difference. Second, Table 4 reveals complex relationships between city size and urban wage growth effects that eluded previous studies that imposed more restrictive specifications. The flexible estimator used here shows an intercept shift (for wage level effects), convexity (for matching effects) and concavity (for learning effects) over city size.

VII. Conclusion

Intuitively, we know that there is something different about cities. How and why they are different is the subject of perennial debate. Recent theory has begun to formalize the intuitions laid out by Alfred Marshall in the 19th century and Jane Jacobs in the mid-twentieth century. This paper has mainly tried to shed light on the source of one of these aspects of cities: the fast wage growth urban workers experience.

Several plausible explanations for this phenomenon are supported by the data. The diminishment of the urban level effect when observable and unobservable characteristics are controlled for suggests that there is sorting into large labor markets. The persistently significant coefficient on the un-interacted urban status variables suggest that productive amenities or agglomeration economies deriving from scale economies in the production of goods, services and/or public infrastructure may also be contributing to high urban wages. However, it is likely that at least some of these level effects are deriving from the matching and learning dynamics discussed in the text. The wage data used in the paper comes from surveys conducted some time after

23

respondents would have moved to or from the city. If rural-urban migrants are able to find better matches in cities, even with their first job, then some of the urban level effect would come from a matching effect. Similarly, if the respondents have had time to live and learn in the city before the survey was conducted, some of the urban level effect could be coming from learning as modeled by Glaeser (1999).

The results also suggest that the dynamics of wage growth differ depending on the size of the city. The smallest cities offer richer and more varied markets than rural areas, but not much opportunity to gain expertise from knowledgeable co-workers or neighbors. In contrast, medium-sized cities offer intense concentrations of industry-relevant knowledge which leads to substantial wage growth as workers pick up useful skills. In the largest cities, workers retain the learning advantages of medium sized cities in terms of learning, but get the added benefit of extremely thick labor markets, which allow them to find better and better matches to their skills. These patterns can be understood as urbanization/localization effects. The results shed light on how the importance of these effects varies across the urban hierarchy, and on the varying functions of cities of different sizes.

While the paper has concentrated on comparing two models of urban wage growth, models can be interpreted in various ways. For instance, the learning model could be interpreted as meaning that in cities it is easier to learn about better job opportunities. Conversely, one could interpret matching as a match to the firm that employs people from whom a person can learn the most. These models are not necessarily mutually exclusive, as the empirical results show. There is plenty different about cities, be it air, markets or some third factor. It should not be surprising that the measurable differences, such as wage or productivity differences, are made up of many small factors. What may be more surprising is that almost all

these factors seem to point in the same direction, towards advantage for urban residence and production.

Bibliography:

Adamson, D., D. E. Clark and M. D. Partridge (2004). "Do Urban Agglomeration Effects and Household Amenities Have a Skill Bias?" <u>Journal of Regional Science</u> **44**(2): 201-23.

Baum-Snow, N. and R. Pavan (forthcoming). "Understanding the City Size Wage Gap." <u>Review of Economic Studies</u>. Accessed 20 August 2011 online http://www.restud.com/paper/understanding-the-city-size-wage-gap/

Charlot, S. and G. Duranton (2004). "Communication externalities in cities." <u>Journal of Urban Economics</u> **56**(3): 581-613.

Duranton, G. and D. Puga (2001). "Nursery cities: Urban diversity, process innovation and the life cycle of products." <u>American Economic Review</u> **91**(5): 1454-77.

Duranton, G. and D. Puga (2004). "Micro-foundations of urban agglomeration economies." <u>Handbook of regional and Urban Economics</u>, J.V. Henderson and J.F. Thisse, eds. Elsevier, Amsterdam. Pp. 2063-117.

Finny, Miles M. and Janet E. Kohlhase (2008). "The effect of urbanization on labor turnover." <u>Journal of Regional Science</u> **48**(2): 311-28.

Fujita,M., P. Krugman and A. Venebles (1999) <u>The Spatial Economy: Cities Regions and International Trade</u>. Cambridge: MIT.

Gabe, T. M. (2004). "Establishment Growth in Small Cities and Towns." <u>International Regional Science Review</u> **27**(2): 164-186.

Glaeser, E. L. (1999). "Learning in cities." <u>Journal of Urban Economics</u> **46**: 254-277.

Glaeser, E. L. and D. C. Maré (2001). "Cities and skills." <u>Journal of Labor Economics</u> **19**(2): 316-42.

Glaeser, E. L. and M. G. Resseger (2010)."The Complementarity between Cities and Skills." <u>Journal of Regional Science</u> **50**(1): 221-44.

Gould, E. D. (2007). "Cities, Workers, and Wages: A Structural Analysis of the Urban Wage Premium." <u>Review of Economic Studies</u> **74**: 477-506.

Head, K and T. Mayer (2004). "The empirics of agglomeration and trade." <u>Handbook of regional and Urban Economics</u>, J.V. Henderson and J.F. Thisse, eds. Elsevier, Amsterdam. Pp. 2609-69.

Helsley, R.W. and W. C. Strange (1990). "Matching and agglomeration economies in a system of cities." <u>Regional Science and Urban Economics</u> **20**: 189-212.

Heueremann, D., B. Halfdanarson and J. Suedekum (2010). "Human Capital Externalities and the Urban Wage Premium: Two Literatures and their Interrelations."

Urban Studies **47**(4): 749-67.

Holtz-Eakin, D. and M. E. Lovely (1996). "Scale economies, returns to variety and the productivity of public infrastructure." Regional Science and Urban Economics **26**(2): 105-23.

Krugman, P. (1991). "Increasing returns and economic geography." Journal of Political Economy **99**(3): 483-99.

Matano, A. and P. Naticchioni (2011). "Wage distribution and the spatial sorting of workers." Journal of Economic Geography. published online. doi: 10.1093/jeg/lbr013

Neal, D. (1995). "Industry-specific human capital: Evidence from displaced workers." Journal of Labor Economics **13**: 653-77.

Neal, D. (1999). "The complexity of job mobility among young men." Journal of Labor Economics **17**: 237-61.

Ottaviano, G., T. Tabuchi and J. Thisse (2002). "Agglomeration and Trade Revisited." International Economic Review **43**(2): 409-435.

Ottaviano, G.I.P. and J.F. Thisse (2004). "Agglomeration and economic geography." Handbook of regional and Urban Economics, J.V. Henderson and J.F. Thisse, eds. Elsevier, Amsterdam. Pp. 2563-608.

Rauch, J. E. (1993). "Productivity gains from geographic concentration of human capital: Evidence from the cities." Journal of Urban Economics **34:** 380-400.

Rosenthal, S. and W. Strange (2004). "Evidence on the nature and sources of agglomeration economies." Handbook of regional and Urban Economics, J.V. Henderson and J.F. Thisse, eds. Elsevier, Amsterdam. Pp. 2119-71.

Teulings, C.N. and P.A. Gautier (2009). "Search and the city." Regional Science and Urban Economics **39**(3): 251-65

Topel, Robert H. and Michael P. Ward (1992). "Job mobility and the careers of young men." The Quarterly Journal of Economics **107**(2): 439-79.

Wheeler, Christopher H. (2006). "Cities and the growth of wages among young workers: Evidence from the NLSY." Journal of Urban Economics **60**: 162-84.

Wheeler, Christopher H. (2008). "Local market scale and the pattern of job changes among young men." Regional Science and Urban Economics **38**(2): 101-18.

Yankow, Jeffrey J. (2006). "Why do cities pay more? An empirical examination of some competing theories of the urban wage premium." Journal of Urban Economics **60**: 139-61.

Figure I: 2000-Census-based urban-rural codes.
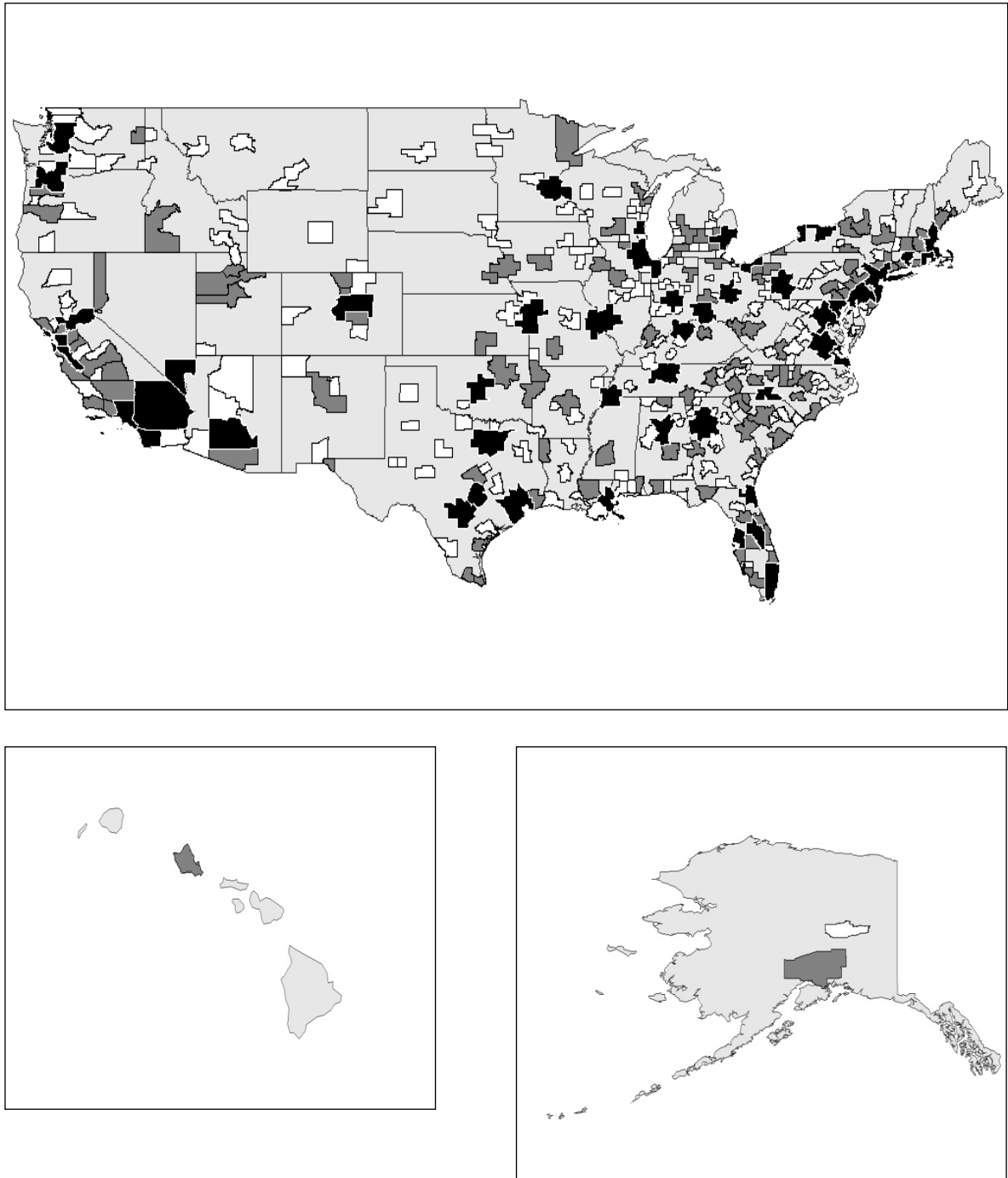
Note: Metropolitan areas with populations over one million shaded in black. Metropolitan areas with populations between 250,000 and one million shaded with dark grey. Metropolitan areas with populations between 50,000 and 250,000 shaded with white.

Table 1: Descriptive Statistics.

Variable Name	full sample	Metropolitan areas Big	Medium	Small	non-metro
logwage	2.09	2.22	2.04	1.99	1.88
	0.63	0.63	0.59	0.60	0.60
Urban	0.800	1	1	1	0
	0.400	-	-	-	-
Medium & Big Urban	0.696	1	1	0	0
	0.460	-	-	-	-
Big Urban	0.475	1	0	0	0
	0.499	-	-	-	-
Total Experience	470.30	489.81	461.49	468.59	434.63
	308.20	309.21	299.92	317.40	306.36
Weeks worked	56.55	57.98	55.44	56.98	54.18
	25.75	25.75	25.10	26.90	25.60
Tenure	183.24	185.46	181.04	188.81	177.51
	213.98	214.92	209.95	226.09	209.54
Education	12.93	13.17	13.02	12.66	12.40
	2.30	2.36	2.33	2.15	2.12
Male	0.560	0.56	0.55	0.54	0.58
	0.496	0.50	0.50	0.50	0.49
Black	0.251	0.28	0.24	0.20	0.21
	0.434	0.45	0.43	0.40	0.40
Age	28.19	28.50	28.05	28.30	27.56
	6.35	6.36	6.16	6.58	6.36
AFQT	42.40	43.71	42.65	41.87	39.31
	28.49	28.68	28.56	27.68	28.14
Hours per week	42.96	42.78	42.90	42.99	43.45
	7.53	7.23	7.26	7.63	8.42
Marital	0.47	0.43	0.47	0.51	0.52
	0.50	0.50	0.50	0.50	0.50
No. Child	0.97	0.91	0.98	1.11	1.03
	1.20	1.18	1.18	1.25	1.20
Observations (approx)	87,000	41,000	19,000	9,000	17,500

Note: Standard Deviations in small font below each average.

Table 2: Urban Level effects without interactions

	1 coef	1 t-stat	2 coef	2 t-stat	3 coef	3 t-stat	4 coef	4 t-stat	5 coef	5 t-stat	6 coef	6 t-stat	7 coef	7 t-stat
Urban	0.1176**	17.16	0.0511**	9.52	0.0501**	10.29	0.0093	1.23	0.0097	1.28	0.0038	0.52	0.0043	0.59
Medium & Big Urban	0.0457**	6.86	0.0510**	10.01	0.0388**	8.30	0.0390**	5.47	0.0223**	3.14	0.0395**	5.77	0.0248**	3.64
Big Urban	0.1821**	40.18	0.1412**	39.74	0.1309**	39.9	0.0683**	11.05	0.0656**	10.73	0.0654**	10.93	0.0634**	10.69
Experience	-		0.0377**	55.16	0.0285**	43.62	0.0379**	26.94	0.0378**	26.72	0.0370**	27.86	0.0373**	27.96
Tenure	-		0.0148**	32.31	0.0153**	35.43	0.0093**	20.38	0.0094**	21.04	0.0082**	18.17	0.0092**	20.60
Education	-		0.0594**	68.04	0.0452**	49.34	0.0864**	34.56	0.0759**	30.68	0.0771**	31.85	0.0689**	28.67
Age	-		0.0152**	24.40	-0.0075**	-10.02	0.0178**	14.00	0.0195**	12.05	0.0180**	15.02	0.0187**	12.08
Marital	-		0.0772**	24.62	0.0638**	22.04	0.0631**	19.66	0.0462**	14.47	0.0612**	19.76	0.0467**	15.11
No. Child	-		0.0079**	5.05	0.0051**	3.53	0.0085**	3.78	-0.0038	-1.71	0.0070**	3.23	-0.0034	-1.54
Male	-		0.1832**	63.69	0.1367**	43.99	-		-		-		-	
Black	-		-0.0451**	-12.42	-0.0380**	-11.31	-		-		-		-	
AFQT	-		0.0022**	32.09	0.0025**	37.64	-		-		-		-	
R^2	0.0359		0.5663		0.6391		0.5239		0.5294		0.5835		0.5870	
Estimation	OLS		OLS		OLS		FE		FE		FE		FE	
Occupation	No		No		Yes		No		No		Yes		Yes	
Industry	No		No		Yes		No		No		Yes		Yes	
Year	No		No		Yes		No		Yes		No		Yes	
Observations	116,890		84,668		84,481		86,533		86,533		86,345		86,345	
Individuals	--		--		--		9,210		9,210		9,208		9,208	

* Significant at the 5% level; ** Significant at the 1% level.

Table 3: Fixed effects estimates using different measures of urban, with and without interactions.

	All				Medium & Big				Big			
	1		2		3		4		5		6	
	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat
Urb	0.0476**	8.10	0.1109**	2.97	0.0591**	11.38	0.0579	1.81	0.0769**	14.46	0.0614	1.95
Experience	0.0377**	28.26	0.0291**	19.62	0.0376**	28.16	0.0306**	21.55	0.0373**	27.98	0.0328**	23.79
Experience×Urb	-		0.0102**	13.37	-		0.0093**	14.39	-		0.0085**	14.32
Tenure	0.0092**	20.50	0.0119**	13.06	0.0092**	20.61	0.0109**	14.99	0.0092**	20.55	0.0114**	19.52
Tenure×Urb	-		-0.0033**	-3.45	-		-0.0023**	-2.79	-		-0.0042**	-5.43
Education	0.0701**	29.06	0.0730**	20.57	0.0697**	28.92	0.0682**	21.95	0.0690**	28.73	0.0671**	24.42
Education×Urb	-		-0.0048	-1.62	-		-0.0000	-0.01	-		0.0014	0.59
Age	0.0186**	12.03	0.0191**	12.37	0.0187**	12.05	0.0194**	12.57	0.0188**	12.12	0.0194**	12.53
Marital	0.0462**	14.92	0.0459**	14.82	0.0463**	14.97	0.0455**	14.72	0.0467**	15.09	0.0460**	14.89
No. Child	-0.0043**	-1.94	-0.0042	-1.93	-0.0039	-1.79	-0.0038	-1.74	-0.0034	-1.57	-0.0037	-1.69
R^2	0.5775		0.5836		0.5806		0.5872		0.5857		0.5925	
Observations	86,345		86,288		86,345		86,260		86,345		86,229	
Individuals	9,208		9,201		9,208		9,198		9,208		9,196	

Note: Fixed effects estimates with unreported controls for occupation, industry and year.

* Significant at the 5% level; ** Significant at the 1% level.

Table 4: Estimates with flexible specification of urban status.

	OLS		Fixed Effects							
	1		2		3		4		5	
	Coef.	t-stat.	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
Urban	0.0994**	2.95	0.1630**	3.49	0.1583**	3.38	0.1468**	3.22	0.1431**	3.14
Medium & Big Urban	-0.0222	-0.74	-0.0056	-0.13	-0.0435	-1.04	0.0068	0.17	-0.0277	-0.68
Big Urban	0.0112	0.57	0.0105	0.30	0.0184	0.52	0.0075	0.22	0.0139	0.40
Experience	0.0198**	24.71	0.0308**	19.67	0.0299**	19.06	0.0296**	19.95	0.0293**	19.72
Experience×Urban	0.0054**	6.69	0.0029*	2.51	0.0039**	3.45	0.0026*	2.36	0.0034**	3.14
Experience×Medium & Big Urban	0.0005	0.62	0.0026*	2.41	0.0023*	2.16	0.0031**	2.99	0.0029**	2.80
Experience×Big Urban	0.0084**	14.34	0.0049**	6.29	0.0052**	6.80	0.0053**	6.99	0.0055**	7.36
Tenure	0.0161**	18.09	0.0114**	11.99	0.0117**	12.48	0.0106**	11.42	0.0117**	12.78
Tenure×Urban	0.0028*	2.15	-0.0020	-1.50	-0.0026*	-1.96	-0.0017	-1.33	-0.0022	-1.72
Tenure×Medium & Big Urban	0.0002	0.20	0.0031*	2.49	0.0032**	2.63	0.0026*	2.12	0.0026*	2.17
Tenure×Big Urban	-0.0065**	-7.08	-0.0051**	-5.23	-0.0047**	-4.95	-0.0053**	-5.61	-0.0049**	-5.24
Education	0.0443**	22.79	0.0904**	24.64	0.0784**	21.42	0.0808**	22.43	0.0711**	19.78
Education×Urban	-0.0070**	-2.57	-0.0113**	-3.07	-0.0108**	-2.93	-0.0105**	-2.92	-0.0100**	-2.80
Education×Medium & Big Urban	0.0042	1.79	0.0025	0.78	0.0041	1.27	0.0017	0.54	0.0032	1.01
Education×Big Urban	0.0062**	4.11	0.0048	1.83	0.0039	1.47	0.0048	1.87	0.0041	1.56
Age	-0.0069**	-9.24	0.0182**	14.30	0.0202**	12.56	0.0185**	15.35	0.0195**	12.64
Marital	0.0647**	22.38	0.0628**	19.54	0.0455**	14.22	0.0607**	19.59	0.0458**	14.81
No. Child	0.0046**	3.18	0.0084**	3.75	-0.0040	-1.78	0.0069**	3.20	-0.0035	-1.60
AFQT	0.0025**	37.95	--		--		--		--	
Male	0.1379**	44.58	--		--		--		--	
Black	-0.0373**	-11.17	--		--		--		--	
R^2	0.6425		0.5302		0.5368		0.5901		0.5947	
Occupation & Industry	Yes		No		No		Yes		Yes	
Year	Yes		No		Yes		No		Yes	
Observations	84,323		86,367		86,367		86,179		86,179	
Individuals	--		9,193		9,193		9,191		9,191	

* Significant at the 5% level; ** Significant at the 1% level.