Ability Sorting and Consumer City

Sanghoon Lee*

University of British Columbia

October 20, 2009

Abstract

This paper proposes a consumption-side explanation for the urban wage premium. The key claim is that the wide consumption variety in large cities is a luxury good that is more important to high-skill workers, and thus the higher average wages in large cities are due to the selection of high skill workers choosing to live there. A unique implication is that urban wage premiums are decreasing in skills and can even be negative for very high-skill workers. I confirm this implication using data on the health care workers.

*Address: Sauder School of Business, University of British Columbia, 2053 Main Mall, Vancouver, BC, Canada V6T 1Z2; E-mail: sanghoon.lee@sauder.ubc.ca. I am indebted to Thomas Holmes for his guidance and patience. I also thank two anonymous referees, James Schmitz, Samuel Kortum, Andrea Moro, Robert Town, Edward Glaeser, Bob Helsley, Hwikwon Ham, Mark Gibson, and seminar participants at University of Minnesota, University of British Columbia, University of Rochester, NBER, FRB Minneapolis, FRB Boston, and Econometric Society World Congress 2005 for valuable suggestions and discussions. All remaining errors are mine.

1 Introduction

Empirical work has documented that average wages are substantially higher in large cities than in small cities, with differences on the order of 30 percent or more (e.g., Glaeser and Mare (2001)). Theoretical work has emphasized the importance of productivity spillovers in accounting for the higher wages in large cities: workers become more productive once they move to large cities and thus earn higher wages (e.g., Henderson (1974)).

This paper proposes a consumption-side theory for the urban wage premium. Large cities offer a wide consumption variety. They have museums, professional sports teams, and French restaurants that small cities do not have. The key idea of this paper is that this urban consumption variety is deemed to be an income elastic good. That is, convenient access to French restaurants is something that is relatively more important to a rich person than a poor person. The high wages found in large cities are due in part to high-skill, thus high-income, workers choosing to live there.

The theory delivers new testable implications. Suppose that there is an industry that is required in both large cities and small cities and that needs to employ both high-skill workers and low-skill workers. In this industry, the wage differential between large cities and small cities should be decreasing in skill, and may even be negative at top skill levels so that workers in large cities can be paid less than their small-city counterparts. Such workers, with great earning power and thus great demand for urban amenities, need significant financial inducements to move to small cities. A second implication is that the ratios of high-skill workers to low-skill workers are higher in larger cities. Firms in large cities substitute relatively cheaper high-skill workers for relatively more expensive low-skill workers. This higher share of high-skill workers in large cities is the ability sorting effect that drives up average wages in large cities.

I examine these implications with the healthcare sector because it is the best example of an industry that is required in both large cities and small cities and that needs to employ both high-skill workers (doctors and dentists) and low-skill workers. This is different from the legal industry, for example, where small cities do not need patent attorneys and thus financial inducements to relocate them are unnecessary.

The results show that the urban wage premium sharply decreases as skill level rises. The negative urban wage premiums exist at the top skill levels. Nurses and physician assistants, for example, are paid more in large cities than in small cities while dentists and optometrists are paid less. The finding that high-skill workers can be paid less in large cities is particularly striking because I also find evidence that better ones, even within an occupation, tend to locate in large cities. Large-city doctors are more likely to be high-skill specialists from top medical schools.

The results also show that high-skill healthcare workers are more concentrated in large cities relative to low-skill workers. For example, the ratio of doctors to nurses is higher in large cities. This ability sorting, a high proportion of high-skill healthcare workers in large cities, causes the average pay of healthcare workers to be substantially higher in larger cities. I find that this ability sorting effect accounts for at least 72 percent of the urban wage premium in the healthcare sector.

Although I use the healthcare sector to test my theory, the theory applies more generally. If urban consumption variety is a luxury good, it attracts high-skill workers into large cities, whether they are in the healthcare sector or not. In fact, the ability sorting effect would be even stronger outside the healthcare sector because they do not have to bring high-skill workers to small cities.

What turns out to be crucial for my theory is that expenditure shares on consumption variety, relative to residential land, increase with income. In other words, the demand for differentiated consumption goods should be more income elastic than the demand for land. Higher variety is the gain for cities, high land price is the cost, so differences in expenditures shares by income are the key. I am not aware of any previous empirical work that directly estimates these parameters, but there exists previous empirical work that indirectly substantiates this assumption. First, it is well-known among urban economists that the demand for residential land is income inelastic. For example, Glaeser et al. (2008) estimate the income elasticity of land demand to be less than 0.4.[1] Second, two pieces of evidence jointly suggest that the demand for consumption variety is income elastic. Bils and Klenow (2001) show that high-income people consume more high-quality goods. Berry and Waldfogel (2003) show that high quality goods can be found in large cities using the restaurant and news paper industries.

The key theory of this paper is that the demand for consumption variety is a luxury good. The role of cities as consumption centers started drawing attention recently. This "consumer city" literature emphasizes urban amenities as the centripetal force attracting workers into cities (e.g. Glaeser et al. (2001), Tabuchi and Yoshida (2000)). However, this literature has not been connected to the urban wage premium literature. In fact, if cities are better places to live, workers should be paid more in small cities and we should see an urban wage *discount,* not premium. This paper reconciles this tension by introducing different skills and ability sorting.

There are other urban economics papers that look at the skill distribution across different-size cities (e.g., Combes et al. (2008), Bacolod et al. (2009a), Bacolod et al. (2009b)). Especially, Combes et al. (2008) argue that ability sorting plays an important role in accounting for the urban wage premium. This paper differs from Combes et al. (2008) in two ways. First, this paper proposes an explicit theoretical mechanism why the sorting occurs and tests this particular mechanism, while Combes et al. (2008) focus on identifying the overall sorting effect regardless of its driving mechanisms. Second, this paper looks at urban wage premiums by skills, while Combes et al. (2008) assumes that urban wage premium does not depend on skills. Bacolod et al. (2009a) look at the distribution of horizontal skills: cognitive skills, social skills, and physical skills and find that the cognitive skills and the social skills

[1] I use demand for land (lot size) instead of demand for housing because housing consumption includes components that may not be relevant in this analysis. For example, rich households live in better structures in better neighborhoods. However, the structure and neighborhood quality should not matter because the structure consists of tradable goods and the neighborhood quality is a local consumption amenity that is not necessarily more costly to obtain in large cities.

are rewarded more in large cities but the physical skills are not. Bacolod et al. (2009b) focus on the social skills and find that social skill levels tend to spread with city size.

The dominant explanation for the urban wage premium is the productivity spillover theories (e.g., Glaeser and Mare (2001), Rosenthal and Strange (2008), Wheaton and Lewis (2002), Wheeler (2006)). One common implication of the productivity spillover theories is that workers, regardless of their types, make higher nominal wages in large cities. The negative urban wage premium found in this paper is the unique prediction of my theory and make the key evidence.

The rest of this paper is structured as follows. Section 2 presents the model and its theoretical implications. Section 3 provides empirical evidence from the healthcare sector as a whole. Section 4 provides evidence only from doctors but differentiated by their specialties and quality. Section 5 discusses alternative explanations for the empirical evidence found in this paper and considers other high-skill occupations outside the healthcare sector.

2 Theory

The key theory in this paper is that the demand for consumption variety is a luxury good. This section derives its theoretical implications, on a non-tradable good industry that is required in both large and small cities and needs to employ both high-skill and low-skill workers. I call this industry the healthcare industry because it is the best example of such an industry and I use it in the empirical analysis. The first implication is on price. The wage differential between large cities and small cities is decreasing in skill and can even be negative at top skill levels. The second implication is on quantity. High skill workers are more concentrated in large cities relative to small cities.

The model uses workers' equal utility condition across cities to pin down equilibrium wages. This is different from Roback (1982) which uses firms' equal cost condition across cities as well as the equal utility condition. The equal cost condition is important for footloose tradable goods industries because their firms would all move away from high cost cities. However, this condition is less important here because we look at a non-tradable good sector that is required everywhere.

The model is a partial equilibrium model. I assume that there are cities with different sizes and that consumption variety and rent increase with city size. The implicit assumption here is that the healthcare sector is relatively small compared to the whole economy and thus does not significantly affect these aggregate variables. With these assumptions, I solve for the equilibrium wages and the location choices of the healthcare workers. Note that the consumption variety here concerns all differentiated goods, not just the healthcare service.

2.1 The Model

There is a continuum of cities, indexed by their population size n. A city of size n has an exogenously given consumption variety level $v(n)$ and rent $r(n)$. Larger cities have higher consumption variety levels and higher rents.[2]

$$v'(n) > 0 \text{ and } r'(n) > 0.$$

A city of population size n requires n units of the healthcare service to be locally provided.

Health care workers differ in their skill $\theta \in \left[\underline{\theta}, \bar{\theta}\right]$. There is an infinite supply of workers for each skill θ. Skill θ workers have an outside option offering utility $\bar{u}(\theta) > 0$. I assume that $\bar{u}(\theta)$ is increasing in θ.[3]

$$\bar{u}'(\theta) > 0.$$

Workers consume land h and a set of differentiated local goods $\{q(x) | x \in \{0, v\}\}$. The

[2] These assumptions can be easily endogenized in a bigger model. See Krugman (1991) and Duranton and Puga (2003) for examples.

[3] This assumption makes workers' equilibrium wage increasing in skill everywhere. The reason I do not directly assume that wages increase in skill is that the workers of the same skill can have different equilibrium wages in different cities.

following utility function describes their preferences.

$$\begin{array}{ll} \left(\int_0^v q(x)^\mu \, dx\right)^{\frac{1}{\mu}} & \text{if } h \geq 1 \\ -\infty & \text{if } h < 1 \end{array}$$

where $\mu < 1$ is the local good complementarity parameter. The condition $\mu < 1$ makes workers care about the consumption variety level v. Note that workers demand one unit of land regardless of their income and this perfectly inelastic demand for land makes the demand for differentiated local goods income elastic.[4]

Workers first decide whether to take the outside option or not and which city to live in. Once in city n, ability θ workers are paid a wage of $w(\theta, n)$, consume one unit of land paying rent $r(n)$, and spend the rest of their income $w(\theta, n) - r(n)$ on the range $v(n)$ of diversified local goods. The local goods are assumed to have the same prices across all types and locations and I use this price as the numeraire price.[5] In summary, ability θ workers solve the following optimization problem.

$$\max_n \left\{ \bar{u}(\theta), \tilde{U}(\theta, n) \right\} \text{ s.t.}$$

$$\tilde{U}(\theta, n) = \max_{q(x)} \left(\int_0^{v(n)} q(x)^\mu \, dx \right)^{\frac{1}{\mu}} \text{ s.t. } \int_0^{v(n)} q(x) \, dx = w(\theta, n) - r(n). \tag{1}$$

Firms have the following standard CES production function.

$$\left(\int_{\underline{\theta}}^{\bar{\theta}} g(\theta, n) \, l(\theta, n)^\gamma \, d\theta \right)^{\frac{1}{\gamma}}$$

where $l(\theta, n)$ is the number of type θ workers employed in the sector in city n, $g(\theta, n)$ is the productivity of a skill θ worker in city n, and $\gamma < 1$ is the input complementarity parameter.

[4] It is only for simplicity that I assume the *perfect* income inelastic demand for land. All the theoretical results hold with more general non-homothetic preferences such as Stone-Geary preference.

[5] This assumption can be endogenized by local good production technology of constant unit marginal cost. In addition, this assumption can be relaxed so that local good prices are different from city to city. However, in that case I would need to introduce exchange rates between local goods in different cities.

The productivity function $g(\theta, n)$ can capture the productivity spillover effect and help clarify which results are unique to the consumption variety theory and which are consistent with both the consumption variety theory and the productivity spillover theory. The CES technology with $\gamma < 1$ makes firms employ all skill types for production. In summary, firms in city n solve the following profit maximization problem.

$$\max_{l(\theta,n)} p^m(n) \cdot \left(\int_{\underline{\theta}}^{\bar{\theta}} g(\theta, n) \, l(\theta, n)^\gamma \, d\theta \right)^{\frac{1}{\gamma}} - \int_{\underline{\theta}}^{\bar{\theta}} w(\theta, n) \, l(\theta, n) \, d\theta$$

where $p^m(n)$ is the unit price for the nontradable good in city n.

2.2 Equilibrium

There are three conditions an equilibrium has to satisfy. First, workers maximize their utilities. Second, firms in each city maximize their profits. Third, the market in each city has to clear.

First, I characterize workers' utility maximization conditions. To begin, I calculate the indirect utility of skill θ workers living in city n. Since both prices and utility weights are equal across all types of local goods, workers consume the same quantity of local goods across all types. In other words, there is $q \in \mathbb{R}_+$ such that $q(x) = q$ for all $x \in [0, v(n)]$ solves workers' optimization problem (1). Substituting q for $q(x)$ in the optimization problem (1), I obtain the following indirect utility function.

$$\tilde{U}(\theta, n) = v(n)^{\frac{1-\mu}{\mu}} \left(w(\theta, n) - r(n) \right). \tag{2}$$

Since there is an infinite supply of workers who can freely choose where to live and whether to take the outside option or not, the indirect utilities across all cities have to be equal to the reservation utility $\bar{u}(\theta)$ offered by the outside option.

$$\tilde{U}(\theta, n) = \bar{u}(\theta) \text{ for all } n \in N. \tag{3}$$

Second, I characterize firms' profit maximization conditions. Since the CES production technology has a constant returns to scale (CRS) property, I can consider one aggregate representative firm for each city. In addition, the representative firm in each city employs a constant proportion of skill types regardless of output levels due to the property of CRS technology. The first order condition implies the following proportion of skill types for production.

$$\frac{l(\theta_1, n)}{l(\theta_2, n)} = \left(\frac{w(\theta_2, n)}{w(\theta_1, n)} \frac{g(\theta_1, n)}{g(\theta_2, n)}\right)^{\frac{1}{1-\gamma}} \quad \text{for any } \theta_1, \theta_2 \in \left[\underline{\theta}, \bar{\theta}\right] \tag{4}$$

Third, the market clearing condition is simple. The aggregate firm in city n has to produce n units of output.

$$n = \left(\int_{\underline{\theta}}^{\bar{\theta}} g(\theta, n)\, l(\theta, n)^{\gamma}\, d\theta\right)^{\frac{1}{\gamma}} \tag{5}$$

An equilibrium of this model consists of the list $\left\{(w(\theta, n), l(\theta, n)) \,|\, \theta \in \left[\underline{\theta}, \bar{\theta}\right], n \in N\right\}$ satisfying conditions (2) to (5) for each city n. Wage schedules $w(\theta, n)$ are pinned down by the workers' utility maximization conditions (2) and (3). The geographic distribution of workers $l(\theta, n)$ is determined by conditions (4) and (5).

2.3 Implications

This section derives two implications from the model. The first implication is on equilibrium prices and the second implication is on equilibrium quantities. First, I derive the price implication: urban wage premiums are decreasing in skill and can even be negative for high enough skills. To begin, I define urban wage premium, as how much percentage a wage increases when city size doubles. Later I estimate this urban wage premium with data.

Definition 1 *The urban wage premium $\beta^w(\theta, n)$ for skill θ at city size n is defined as the* *population size elasticity of wage.*

$$\beta^{w}(\theta,n)=\frac{\partial\log w(\theta,n)}{\partial\log n}.$$

The wage schedule $w(\theta,n)$ is pinned down by equations (2) and (3).

$$w(\theta,n)=r(n)+\bar{u}(\theta)\cdot\frac{1}{v(n)^{\frac{1-\mu}{\mu}}}. \tag{6}$$

Wage schedule (6) shows that there are two types of wage compensations that ensure workers of the same type achieve the same level of utility across different cities. One is the land price compensation $r(n)$ and the other is the consumption variety compensation $\bar{u}(\theta)/v(n)^{(1-\mu)/\mu}$. Note that the two compensations work in the opposite direction. As city size n rises, the land price compensation $r(n)$ increases while the consumption variety compensation $\bar{u}(\theta)/v(n)^{(1-\mu)/\mu}$ decreases. The land price compensation $r(n)$ is equal across different skill types of workers because all workers consume one unit of land regardless of their skills. On the other hand, the consumption variety compensation $\bar{u}(\theta)/v(n)^{(1-\mu)/\mu}$ is increasing in skill θ because the demand for local goods is income elastic. Therefore, urban wage premiums are positive for low-skill workers, for whom the land price compensation dominates the consumption variety compensation, and negative for high-skill workers, for whom the consumption variety compensation dominates the land price compensation.

Proposition 1 *1) Urban wage premium $\beta^{w}(\theta,n)$ for skill θ at city n is decreasing in skill θ.*

$$\frac{\partial}{\partial\theta}\beta^{w}(\theta,n)<0 \text{ for all } n\in N$$

2) Urban wage premium $\beta^{w}(\theta,n)$ for skill θ at city n is negative if and only if

$$\theta>\bar{u}^{-1}\left(\frac{\mu}{1-\mu}\frac{v(n)^{\frac{1}{\mu}}}{v\prime(n)}r\prime(n)\right).$$

Proof. 1)

$$\begin{aligned}
\frac{\partial \beta^{w}(\theta, n)}{\partial \theta} &= \frac{\partial}{\partial \theta} \frac{\partial w(\theta, n)}{\partial n} \frac{n}{w(\theta, n)} \\
&= -\frac{n v(n)^{\frac{1}{\mu}} \bar{u}\prime(\theta) \cdot (\mu v(n) r\prime(n) + (1-\mu) r(n) v\prime(n))}{\mu \left(r(n) v(n)^{\frac{1}{\mu}} + v(n) \bar{u}(\theta)\right)^{2}} < 0.
\end{aligned}$$

2)

$$\beta^{w}(\theta, n) = \frac{\partial \log w(\theta, n)}{\partial \log n} = \frac{n\left(\mu v(n)^{\frac{1}{\mu}} r\prime(n) + (-1+\mu) \bar{u}(\theta) v\prime(n)\right)}{\mu\left(r(n) v(n)^{\frac{1}{\mu}} + v(n) \bar{u}(\theta)\right)}$$

Thus, $\beta^{w}(\theta, n) < 0$ if and only if $\theta > \bar{u}^{-1}\left(\frac{\mu}{1-\mu} \frac{v(n)^{\frac{1}{\mu}}}{v\prime(n)} r\prime(n)\right)$. ■

The decreasing urban wage premiums in skills in Proposition 1 mean that high-skill workers have a greater preference to live in large cities than low-skill workers. If their preferences for large cities are strong enough, they may be paid even less in large cities than in small cities. Note that Proposition 1 holds regardless of workers' productivity distribution function $g(\theta, n)$. This means that the wage implications are the unique prediction of my theory, distinguished from any productivity-based theories.

Now I derive the quantity implication: high-skill workers are more concentrated in large cities as compared to low-skill workers. Since high-skill workers are relatively cheaper in large cities, firms in large cities employ relatively more high-skill workers. To begin, I define urban concentration rate as the percentage by which the number of workers increases when city size doubles. Later, I estimate the urban concentration rates with data.

Definition 2 *The urban concentration rate $\beta^{q}(\theta, n)$ for type θ at city size n is defined as the population size elasticity of the number of skill θ workers.*

$$\beta^{q}(\theta, n) = \frac{\partial \log l(\theta, n)}{\partial \log n}$$

I obtain $l(\theta, n)$ from condition (4) and condition (5).

$$l(\theta, n) = n \cdot \frac{\left(g(\theta, n) / w(\theta, n)\right)^{\frac{1}{1-\gamma}}}{\Gamma(n)^{\frac{1}{\gamma}}} \tag{7}$$

where $\Gamma(n) = \left(\int_{\underline{\theta}}^{\bar{\theta}} g(\theta, n)^{\frac{1}{1-\gamma}} \left(\frac{1}{w(\theta,n)}\right)^{\frac{\gamma}{1-\gamma}} d\theta\right)$. Using equation (7) I obtain the following.

$$\begin{aligned} \frac{\partial}{\partial\theta}\beta^{q}(\theta, n) &= \frac{\partial}{\partial\theta}\frac{\partial \log l(\theta, n)}{\partial \log n} = \frac{\partial}{\partial\theta}\frac{\partial l(\theta, n)}{\partial n}\frac{n}{l(\theta, n)} \\ &= \frac{1}{1-\gamma}\left\{\frac{\partial}{\partial\theta}\beta^{g}(\theta, n) - \frac{\partial}{\partial\theta}\beta^{w}(\theta, n)\right\} \end{aligned} \tag{8}$$

where $\beta^{g}(\theta, n) \equiv \partial \log g(\theta, n) / \partial \log n$. $\beta^{g}(\theta, n)$ represents the percentage increase of productivity of skill θ workers as city size doubles, and this captures the productivity spillover effect for skill θ workers. Since $\frac{\partial}{\partial\theta}\beta^{w}(\theta, n)$ is negative due to Proposition 1, I obtain the following proposition.

Proposition 2 *Suppose that $\beta^{g}(\theta, n) \geq 0$. The urban concentration rate $\beta^{q}(\theta, n)$ at city n is increasing in skill θ.*

$$\frac{\partial}{\partial\theta}\beta^{q}(\theta, n) > 0 \text{ for all } n \in N$$

Proposition 2 captures the ability sorting effect that high-skill workers are relatively more concentrated in large cities compared to low-skill workers. Equation (8) shows that this ability sorting can arise for two reasons. First, the consumption variety effect; since high-skill workers are relatively cheaper in large cities (Proposition 1), firms hire relatively more high-skill workers in large cities. Second, the productivity spillover effect; if high-skill workers benefit more from productivity spillover than low-skill workers, the relative productivity of high-skill workers to low-skill workers is higher in large cities and this makes firms hire relatively more high-skill workers in large cities. Therefore, the quantity implication is consistent with both my theory and the productivity-based theories. On the other hand, the price implication in Proposition 1 is unique to my theory and thus makes it the key implication of this paper.

3 Health Care Sector as a Whole

This section presents empirical evidence using the healthcare sector data. I use the healthcare sector to examine the theoretical implications because it is the best example of a non-tradable good industry that is required both in large and small cities and that needs to employ both high and low skill workers.

The primary data set used in this section is the census 2000 5 percent Public Use Micro Samples (PUMS). I look at all occupations in the healthcare sector except veterinarians (34 occupations, census occupation codes 300-365). I restrict the sample so that the respondents are employed, working in private sector, not attending school, working more than 20 hours per week, and less than 65 years of age in contiguous U.S. I use the Metropolitan Statistical Area (MSA) as geographic units. The metropolitan areas used here contain about 76 percent of the U.S. population.[6]

3.1 On Prices

I first examine the price implication in Proposition 1. I use the average annual income of an occupation as the measure for skill, and show that urban wage premiums are decreasing in skill across healthcare occupations and can even be negative at the top skill levels.

I present the results using the urban wage premiums defined in Definition 1. I assume that the elasticities are constant across the different-size cities and I calculate the urban wage premium $\beta^w(\theta)$ for each occupation θ by running the following individual level regression

[6] The Public Use Microdata Areas (PUMAs), the smallest geographic units in the census 5 percent PUMS, are not fine enough to fully identify MSAs. I approximate each metropolitan area with the group of PUMAs contained within the metropolitan area. I lose about 3 percent of population by dropping the PUMAs that stretch across metropolitan area borders. The original metropolitan areas have about 79 percent of the US population.

for each occupation θ.

$$\log w_i = \alpha^w(\theta) + \beta^w(\theta) \cdot \log n_i + \varepsilon_i \tag{9}$$

where w_i is individual i's annual total income, n_i is the population size of the metropolitan area where individual i lives, ε_i is the individual error term. I cluster the error term ε_i by metropolitan areas. The regression coefficient $\beta^w(\theta)$ captures the urban wage premium for occupation θ.

Figure 1 shows the urban wage premiums across all healthcare occupations against their average annual incomes. (See Table 1 for precise estimate values.) There exists a clear negative relationship between urban wage premiums and skill levels. I summarize the negative relationship by running generalized least squares (GLS) regression, since the urban wage estimates have different standard errors across different occupations. The GLS coefficient is -4.7 percent with a standard error 0.53 percent. All urban wage premiums of high-skill occupations making more than $80,000 per year are negative, with those of doctors, dentists, and optometrists being significantly negative at 95 percent confidence level. The solid symbols in Figure 1 indicate that their urban wage premiums are significantly different from zero at a 95 percent confidence level.

Note that these price regressions are run with nominal incomes. In order to show that workers prefer large cities to small cities, it suffices to show that urban wage premiums are negative in real income, as in Tabuchi and Yoshida (2000). I use nominal income here because the negative urban wage premiums in nominal terms can never be obtained only with the productivity spillover theories, and thus make it one of the key implications of my theory; one common implication among the productivity spillover theories is that firms pay higher nominal wages to workers in large cities, regardless of their skills. The cost of living tends to increase at 5.2 percent when city size doubles, according to ACCRA cost of living index (2000).[7] Roughly speaking, the occupations that are below the ACCRA line in Figure

[7] I obtained this number by running metropolitan area level regression where logged composite price indexes in the ACCRA data are regressed on logged population size.

1, would have negative urban wage premiums in real income.

However, the downward pattern in Figure 1 might be due to some other factors that are correlated with metropolitan area size and individual income. I control for other possible factors by running a Mincer regression for each occupation. I add to the previous regression the standard control variables such as gender, race, working hours, age, squared age, self-employed, marital status, having a child.[8] Figure 2 shows urban wage premiums after controlling for these factors. (See Table 1 for precise estimate values for urban wage premiums. See Table 2 for the coefficient estimates for the control variables.) The negative relationship between urban wage premiums and skills decreases in absolute value but still remains strong: the GLS coefficient is -2.9 percent with a standard error 0.5 percent.[9] Urban wage premiums of dentists and optometrists stay significantly negative at a 95 percent confidence level. The urban wage premium of doctors becomes statistically zero, but this is still a big discount if the cost of living differences are considered.

Even with the Mincer regression controls, there can still be other characteristics that are unobservable with the Census data. For example, doctors in large cities may be inferior to their counterparts in small cities in a way that is not observable with the Census data and this may drive the negative correlation between urban wage premiums and skills. I look at some of these unobservable characteristics in section 4 with more specialized data sets relating to doctors and find the opposite pattern: doctors in large cities are more likely to be specialists from better medical schools. This strengthens the negative correlation I find with the Census data.

One may try to control for these unobservable characteristics using panel data by looking at workers moving across different-size cities over time, as in Combes et al. (2008). However,

[8] I do not include education-level control variables because there is not much education-level variation within an occupation. Moreover, the Mincer regression I ran with full education-level dummies shows the same pattern as the one reported in this paper.

[9] It is interesting that adding the control variables flattens the slope in the second-stage regression. This means that spatial sorting by these characteristics are negative; for example, young workers tend to make lower wages and high skill occupations have relatively more young workers in large cities compared to low skill occupations. A potential explanation for this pattern is that training venues for high skill workers (e.g., medical schools) are relatively more concentrated in large cities.

this imposes formidable data requirements which are hard to satisfy in this paper. I would need a panel data set with a large enough sample of workers in each occupation moving across different cities over time. In addition, even with this approach, there still remains the issue that moving decisions are endogenous.[10]

3.2 On Quantities

Now I test the quantity implication in Proposition 2. I present the results using the urban concentration rates defined in Definition 2. I assume that the elasticities are constant across the different size of cities, and calculate the urban concentration rate $\beta^q(\theta)$ for occupation θ by running the following metropolitan-area-level regression for each occupation θ.

$$\log l(\theta, n_j) = \alpha^q(\theta) + \beta^q(\theta) \cdot \log n_j + \nu_{\theta,j} \tag{10}$$

where $l(\theta, n_j)$ is the number of workers in occupation θ in metropolitan area j, n_j is the population size of metropolitan area j, and $\nu_{\theta,j}$ is the error term. The regression is weighted by the metropolitan area population size n_j. The regression coefficient $\beta^q(\theta)$ captures urban concentration rate for occupation θ, i.e. the percentage change in the number of occupation θ workers increase when city size doubles. The urban concentration rate $\beta^q(\theta)$ is equal to 1 (or 100%) if workers in occupation θ are distributed proportionately to population size. The urban concentration rate $\beta^q(\theta)$ is greater than 1 if workers in occupation θ are disproportionately concentrated in large cities, and vice versa.

Figure 3 shows urban concentration rates for all healthcare occupations against their average annual incomes. (See Table 1 for precise estimate values.) There exists a clear positive relationship between urban concentration rates and incomes. The GLS regression coefficient between average incomes and urban concentration rates is 6.5 percent with a standard error of 1.5 percent. In addition, the urban concentration rates are greater than

[10] These moving decisions across cities can be correlated with better job opportunities, promotions, unemployment, retirement, etc., which are correlated with wages.

1 for most high-skill occupations with annual incomes of $80,000 or more, and less than 1 for most low-skill occupations. This implies that firms substitute high-skill occupations for low-skill occupations in large cities, and vice versa in small cities.

These results are consistent with the theoretical implications in Proposition 2. As mentioned in the theory section, this upward pattern is consistent with both my theory and the productivity spillover theories. On the other hand, the price results are unique to the consumption variety theory, and thus make it the key evidence of this paper.

3.3 How Much Does Ability Sorting Account for the Urban Wage Premium in the Health Care Sector?

The higher percentage of high-skill occupations in large cities reported in the previous section is the ability sorting effect on the urban wage premium and this drives up the average pay of the whole healthcare sector in large cities. This section shows that this ability sorting effect accounts for 72 percent of the urban wage premium in the healthcare sector. I group cities into large cities and small cities so that each group has similar total population size.[11]

First of all, there exists an urban wage premium for the healthcare sector as a whole. The average annual income of healthcare workers living in the large cities is 15 percent higher than those living in the small cities ($56,797 vs. $49,525). The average income difference $\bar{W}^L - \bar{W}^S$ between the large cities and the small cities can be expressed as

$$\bar{W}^L - \bar{W}^S = \sum_\theta \left(\phi_\theta^L - \phi_\theta^S\right) \bar{W}_\theta^S + \sum_\theta \phi_\theta^S (\bar{W}_\theta^L - \bar{W}_\theta^S) + \sum_\theta \left(\phi_\theta^L - \phi_\theta^S\right) \left(\bar{W}_\theta^L - \bar{W}_\theta^S\right) \quad (11)$$

where $\bar{W}_\theta^i$ is the average income of occupation θ in area i ($i = L$ for the large cities and $i = S$ for the small cities), and ϕ_θ^i is the quantity share of occupation θ in area i ($\phi_\theta^i \equiv l_\theta^i / \sum_\theta l_\theta^i$ where l_θ^i is the number of workers in occupation θ in area i).

[11] The large cities are the top 17 MSAs with the total population size of about 112 million. The smallest MSA among the large cities is San Diego. The small cities are the remaining MSAs with total population size of about 117 million.

There are three terms in the decomposition (11). The first term $\sum_\theta \left(\phi_\theta^L - \phi_\theta^S\right) \bar{W}_\theta^S$ is the contribution of quantity variation among occupations. This term captures the ability sorting effect in the urban wage premium. This term would disappear if the skill distribution were equal between the large cities and the small cities. The second term $\sum_\theta \phi_\theta^S(\bar{W}_\theta^L - \bar{W}_\theta^S)$ is the contribution of wage variation within occupations. This term would disappear if the average wages for each occupation were equal between the large cities and the small cities. The third term $\sum_\theta \left(\phi_\theta^L - \phi_\theta^S\right) \left(\bar{W}_\theta^L - \bar{W}_\theta^S\right)$ is the covariance term which shows how the wage changes between the large cities and the small cities are correlated with the quantity share changes. My theory predicts this covariance term to be negative because high-skill workers are concentrated more in large cities but paid less there. The following table summarizes the decomposition results in absolute terms and percentage terms.

$\bar{W}^L - \bar{W}^S$	$\sum_\theta \left(\phi_\theta^L - \phi_\theta^S\right) \bar{W}_\theta^S$	$\sum_\theta \phi_\theta^S(\bar{W}_\theta^L - \bar{W}_\theta^S)$	$\sum_\theta \left(\phi_\theta^L - \phi_\theta^S\right) \left(\bar{W}_\theta^L - \bar{W}_\theta^S\right)$
\$7,272	\$5,244	\$2,619	-\$590
100%	72%	36%	-8%

The decomposition result shows that the ability sorting effect — quantity variations across occupations — accounts for 72 percent of the urban wage premium. The negative covariance term confirms the theoretical prediction that the wage changes within occupations are negatively correlated to the quantity share changes between occupations.

As a robustness check, I run this decomposition exercise using a different set of definitions for the large cities and the small cities. I divide cities into three groups: large cities, medium cities, and small cities so that all the three groups have similar total population sizes.[12] I

[12] The large cities are the top 7 MSAs with total population size about 76.4 million. The smallest MSA among the large cities is Boston-Worcester-Lawrence CMSA. The small cities are the bottom 240 MSAs with total population size 73.4 million. The largest MSA among the small cities is Raleigh-Durham-Chapel Hill MSA.

compare the large cities and the small cities. The following table shows the result.

$\bar{W}^L - \bar{W}^S$	$\sum_\theta \left(\phi_\theta^L - \phi_\theta^S\right) \bar{W}_\theta^S$	$\sum_\theta \phi_\theta^S (\bar{W}_\theta^L - \bar{W}_\theta^S)$	$\sum_\theta \left(\phi_\theta^L - \phi_\theta^S\right) \left(\bar{W}_\theta^L - \bar{W}_\theta^S\right)$
\$9,440	\$7,174	\$3,389	-\$1123
100%	76%	36%	-12%

With the new definitions, the large cities have \$9,440 higher annual income than the small cities (\$58,313 vs. \$48,873). The ability sorting term (the quantity variation among occupations) accounts for 76 percent of the gap, the within occupation term accounts for 36 percent, and the covariance term accounts for -12 percent.

4 Focus on Doctors

The previous section examined the theoretical implications across different occupations in the healthcare sector. However, there exists sizeable heterogeneity even within an occupation, that are not observable with the Census data. For example, within the doctors there are different specialties, and even within the same specialties, doctors may vary in their quality. This section examines the theoretical implications just with doctors, but in two different subdimensions: the doctors with different specialties and the doctors with different levels of quality measured by the quality of medical schools they attended.

The primary data sets used in this section are the Community Tracking Study (CTS) physician survey 2000 - 2001 and the year 2000 version of the American Medical Association (AMA) physicians' master file. I use the CTS data set to test the price implications. The CTS is a micro data set that contains 12,406 physicians from 60 sites (51 metropolitan areas and 9 nonmetropolitan areas) randomly selected to be representative of the nation as a whole. The CTS data set has more detailed occupation specific information for doctors compared to the census data, such as specialty, board certification, hospital ownership, etc.

I use the AMA data set to test the quantity implications. The AMA data set has very detailed information on most physicians in the U.S., such as practice locations, specialty, and medical school attended.[13] I use the AMA data set to test the quantity implication because the CTS data does not cover all of the U.S. However, the AMA data set does not have income information so I cannot use it to test the price implications.

I aggregate medical specialties into 4 groups, for both data sets, following the standard classification scheme — general practice/family physician, medical specialties, surgical specialties, and other specialties. The second column of Table 3 shows the average annual income for each specialty. Across specialties, surgical specialties make the highest income whereas general practice/family physician make the lowest income.

4.1 Across Different Medical Specialties

To begin, I calculate the urban wage premium and the urban concentration rate for all doctors.[14] The results with the CTS data can differ from the results with the Census data because the CTS data set uses different sample selection rule. The CTS data set restricts its sample to the physicians providing direct patient care for more than 20 hours per week, excluding federal employees, foreign medical school graduates who are only temporarily licensed to practice in the U.S., and specialists in fields where the primary focus is not direct patient care. The second row of Table 3 shows the urban wage premiums and the urban concentration rates for all doctors. The urban wage premiums for doctors are significantly negative, with or without Mincer controls. (See Table 4 for the full regression table.) The urban concentration rate is also significantly greater than 1.

Now I look at the doctors by specialties. The fourth and fifth columns of Table 3 report urban wage premiums by specialties.[15] Urban wage premiums tend to decrease in skill. Without controls, the urban wage premium for surgical specialties is -7.1 percent and the

[13] The AMA data set covers most physicians, including AMA members and nonmembers, and graduates of foreign medical schools who satisfy the requirements to be recognized as physicians in the U.S.

[14] The CTS data set has a complicated sample design. I use the sample design parameters for combined sample given in its user guide Table 4.1.

[15] The full regression table is provided in Table A.4.

urban wage premium for general practice is -3.9 percent. With controls, the urban wage premium for surgical specialties is -4.4 percent, while the urban wage premium for general practice is -1.9 percent. These results are consistent with the theory, but the differences across specialties are not statistically significant.

The sixth column of Table 3 shows urban concentration rates by skills. The results are a little mixed. The medical specialties group, which ranks second in average income, is most concentrated in large cities. However, the results are generally consistent with my theory in that the general practice/family physicians are least concentrated in large cities. Hospitals in large cities substitute specialists for generalists, who are relatively low-skill.

This paper is not the first to report this pattern that large cities have relatively more specialists than small cities. Baumgardner (1988) explains this phenomenon as the division of labor arising through scale economies. This argument is certainly valid for the quantity results, but does not explain the price results that urban wage premiums are negative and decreasing in skill across specialties.

4.2 Doctor Quality

This section tests the quantity implications with different qualities of doctors. Using the quality of medical schools attended as the measure of doctor quality, I show that doctors from better medical schools tend to locate more in large cities. I use the average Medical College Admission Test (MCAT) score of each medical school as the measure for medical school quality.[16]

Using the AMA data set, I calculate urban concentration rate for the graduates of each medical school θ by running the metropolitan area regression in equation (10). Figure 4 shows the result. There exists a clear positive relationship, meaning that doctors from better medical schools are more concentrated in large cities. The GLS regression coefficient

[16] The average MCAT scores are obtained from *US News & World Report - Best Graduate Schools.* In this section I use MCAT scores from year 2001, but the results are robust with all the other years I tried - 2002 and 2003.

between the urban concentration rates and the MCAT scores is 0.19 with a standard error of 0.03.

This upward trend may arise for other reasons. First, top medical schools tend to produce relatively more specialists than generalists, and large cities need relatively more specialists compared to small cities. I resolve this issue by showing that the upward trend persists within each specialty. Second, top medical schools tend to be located in large cities and their graduates locate near their medical schools. Figure 5 shows the geographic distribution of doctors from University of Illinois medical schools; they are heavily concentrated in Illinois. I resolve this issue by focusing only on migrant doctors. More specifically, I calculate the urban concentration rate for each medical school using only doctors practicing at least 500 miles away from their schools.[17] Figure 6 shows the urban concentration rates across medical schools in each specialty, after controlling for doctors' geographic concentrations near their medical schools. The positive pattern remains strong for each specialty.

There is health economics literature examining physician quality across geographic areas but most of the studies look at the quality differences across different states (e.g., Baicker and Chandra (2004), Jencks et al. (2000), Fisher and Skinner (2001)). This paper differs from these papers in that I look at the quality differences across different-sized cities and I use a different measure for doctor quality.

5 Discussion

The key evidence in this paper is that urban wage premiums are decreasing in skill and can even be negative at top skill levels. However, there may be other explanations for this result. Moreover, it turns out that other high-skill occupations, such as lawyers, have significantly positive urban wage premiums. This section examines two alternative explanations for the

[17] For each medical school I drop all observations in metropolitan areas within a 500 mile radius of the school. The distances between medical schools and metropolitan areas are calculated using the ZIP codes of the medical schools and the latitudes and longitudes of the metropolitan areas provided in the census data geographic file.

price results and discusses why lawyers may have a larger urban wage premium compared to doctors.

5.1 Human capital accumulation

Glaeser and Mare (2001) show that wages tend to grow more quickly in large cities. One alternative explanation for the decreasing urban wage premiums in skill is that young high-skill workers anticipating this faster wage growth may stay in large cities despite the low current wages. I examine this hypothesis by looking at those aged at least 50 years old for whom this human capital accumulation effect should not play a big role.

I calculate the urban wage premiums by running the Mincer regression in section 3.1, but this time only with those aged between 50 and 65. Figure 7 shows the urban wage premiums against the average occupation incomes. The negative relationship between urban wage premiums and skills remains strong. The GLS regression coefficient is -2.6 percent with a standard error of 0.6 percent. This is not statistically different from the GLS coefficient I obtained with the full sample (-2.9 percent with a standard error of 0.5 percent).

5.2 Colocation problem

Another explanation for the decreasing urban wage premiums in skills is the colocation problem for high-skill workers. The spouses of high-skill workers are also likely to be high-skill workers. These high-skill dual career couples may have difficulty finding jobs for both of them in small cities, because their jobs tend to be specialized and small cities do not have many high-skill jobs (See Costa and Kahn (2000)). This colocation problem can decrease the supply of high-skill workers in small cities, thereby raising their wages there.

There is controversy over the importance of the colocation problem in workers' location choices. Compton and Pollak (2007) find that the colocation problem does not play an important role in accounting for the large share of highly educated dual career couples in large cities. Even if it does, the colocation problem has the same implication for the urban wage premium as my theory. That is, both theories imply that the demand for living in big cities increases with skill levels, and both lead to ability sorting across different-sized cities.

5.3 Lawyers

Lawyers are arguably the most similar occupation to doctors outside the healthcare sector; they earn high incomes and their jobs are highly specialized. However, it turns out that their urban wage premium is much larger than that of doctors. According to the census data, the urban wage premium of lawyers is 8.4 percent while that of doctors is not statistically different from 0 when I run the Mincer regressions from section 3.1. I claim that the urban wage premium of lawyers is much higher than that of doctors for two reasons.

First, there is not much need for lawyers in small cities as compared with doctors. The census data shows that the small cities have only half as many lawyers per capita as the large cities while they have 83 percent as many doctors.[18] This relatively small demand for lawyers in small cities lowers the financial inducement necessary to bring them to live in small cities, and makes the urban wage premium of lawyers higher than that of doctors.

Second, the lawyers in small cities are noticeably less skilled than their peers in large cities, as compared with doctors. For example, the smallest MSA in the census is Enid, Oklahoma with a population of 57,813 people. Even this small city needs to have top-skill specialty doctors because patients may not have time to drive to big cities for emergencies. Anesthesiologists and surgeons, for example, are the top two medical specialties with respect to income.[19] The AMA data set shows that Enid has three anesthesiologists and five surgeons, including one cardiovascular surgeon. In contrast, there is not much demand for top-skill specialty lawyers in small cities because their clients, if any, can go to the nearby big cities. For example, Enid did not have any lawyers in finance, investment, or intellectual

[18] The large cities have 4.6 lawyers per 1000 population while the small cities have 2.3 lawyers. In contrast, the large cities have 29 doctors per 1000 and the small cities have 24 doctors.

[19] According to the Occupational Outlook Handbook 2004 by the U.S. Department of Labor, the average incomes for anaesthesiologists and surgeons are $306,964 and $255,438 respectively.

property in 2005 according to Martindale-Hubbel legal directory.[20]

6 Conclusion

The key idea in this paper is that the consumption variety in large cities is an income elastic good. This leads to the selection of high-skill workers into large cities. A testable implication of the theory, distinguished from the productivity spillover theories, is that urban wage premiums are decreasing in skills and that there may exist urban wage discounts for high-skill workers. I test this implication with the healthcare sector data. I find that urban wage premiums are in fact decreasing in skill, which implies that the preference for large cities rises in skill, and that high-skill healthcare workers in large cities can be paid less than their peers in small cities. I also find that ability sorting accounts for 72 percent of the urban wage premium for the whole healthcare sector.

References

Bacolod, M., Blum, B. S., and Strange, W. C. 2009a. Skills in the city. Journal of Urban Economics, 65(2):136–153.

Bacolod, M., Blum, B. S., and Strange, W. C. 2009b. Urban interactions: Soft skills versus specialization. Journal of Economic Geography, 9(2):227–262.

Baicker, K. and Chandra, A. 2004. Medicare spending, the physician workforce, and beneficiaries' quality of care. Health Affairs, 10.

Baumgardner, J. R. 1988. The division of labor, local markets, and worker organization. Journal of Political Economy, 96(3):509–27.

[20] Available at http://www.martindale.com/.

Berry, S. and Waldfogel, J. 2003. Product Quality and Market Size. National Bureau of Economic Research, Inc, NBER Working Papers: 9675.

Bils, M. and Klenow, P. J. 2001. Quantifying quality growth. American Economic Review, 91(4):1006–30.

Combes, P.-P., Duranton, G., and Gobillon, L. 2008. Spatial wage disparities: Sorting matters! Journal of Urban Economics, 63(2):723–742.

Compton, J. and Pollak, R. A. 2007. Why are power couples increasingly concentrated in large metropolitan areas? Journal of Labor Economics, 25(3):475–512.

Costa, D. L. and Kahn, M. E. 2000. Power couples: Changes in the locational choice of the college educated, 1940-1990. Quarterly Journal of Economics, 115(4):1287–1315.

Duranton, G. and Puga, D. 2003. Micro-Foundations of Urban Agglomeration Economies. National Bureau of Economic Research, Inc, NBER Working Papers: 9931.

Fisher, E. and Skinner, J. 2001. Comparing the health care of states. Providence Journal-Bulletin.

Glaeser, E. L., Kahn, M. E., and Rappaport, J. 2008. Why do the poor live in cities? the role of public transportation. Journal of Urban Economics, 63(1):1–24.

Glaeser, E. L., Kolko, J., and Saiz, A. 2001. Consumer city. Journal of Economic Geography, 1(1):27–50.

Glaeser, E. L. and Mare, D. C. 2001. Cities and skills. Journal of Labor Economics, 19(2):316–42.

Henderson, J. V. 1974. Optimum city size: The external diseconomy question. Journal of Political Economy, 82(2):373–88.

Jencks, S. F., Cuerdon, T., Burwen, D. R., Fleming, B., Houck, P. M., Kussmaul, A. E., Nilasena, D. S., Ordin, D. L., and Arday, D. R. 2000. Quality of medical care delivered to medicare beneficiaries a profile at state and national levels. The Journal of the American Medical Association, 284(13):1670–1676.

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

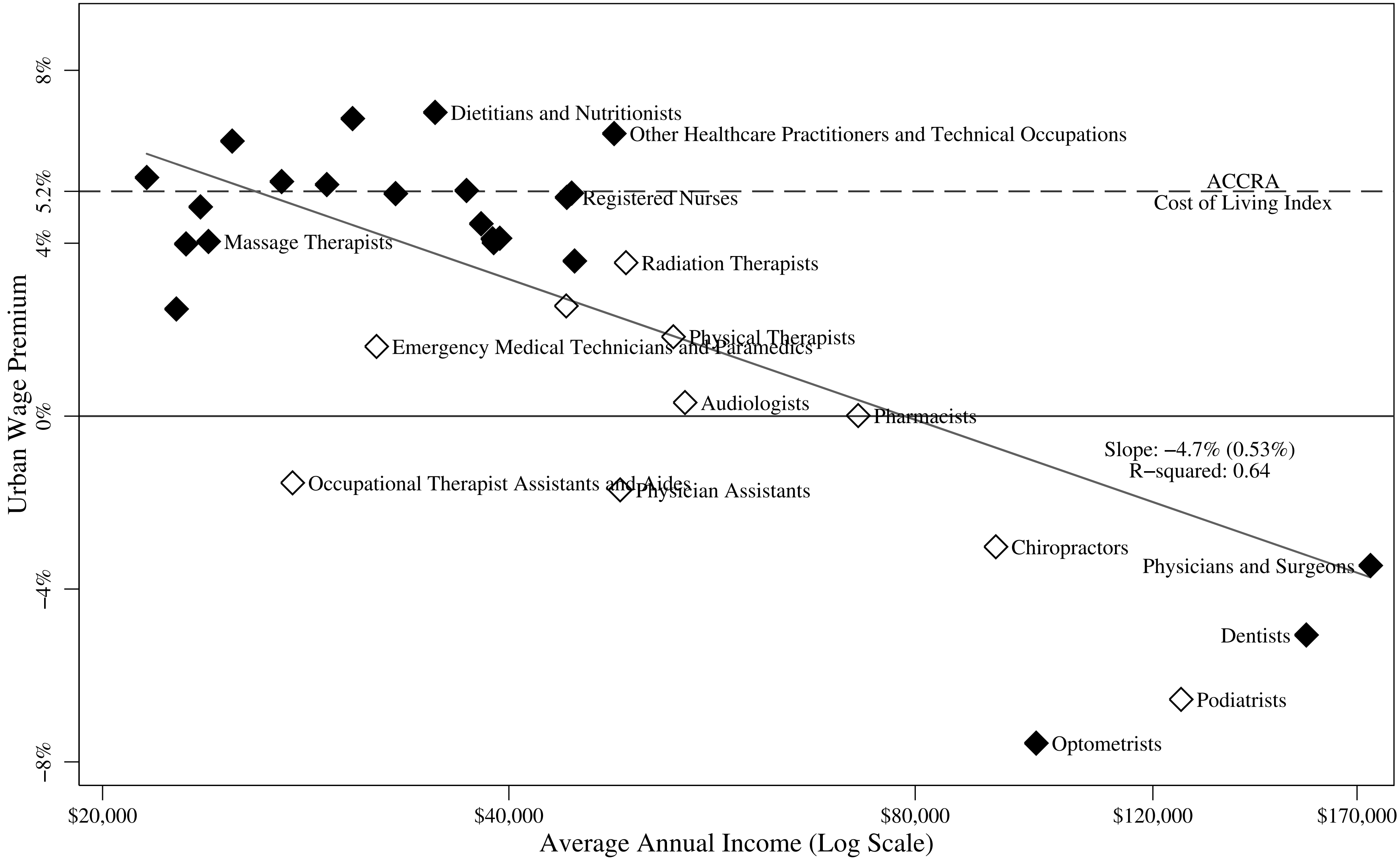
Rosenthal, S. S. and Strange, W. C. 2008. The attenuation of human capital spillovers. Journal of Urban Economics, 64(2):373 – 389.

Tabuchi, T. and Yoshida, A. 2000. Separating urban agglomeration economies in consumption and production. Journal of Urban Economics, 48(1):70–84.

Wheaton, W. C. and Lewis, M. J. 2002. Urban wages and labor market agglomeration. Journal of Urban Economics, 51(3):542 – 562.

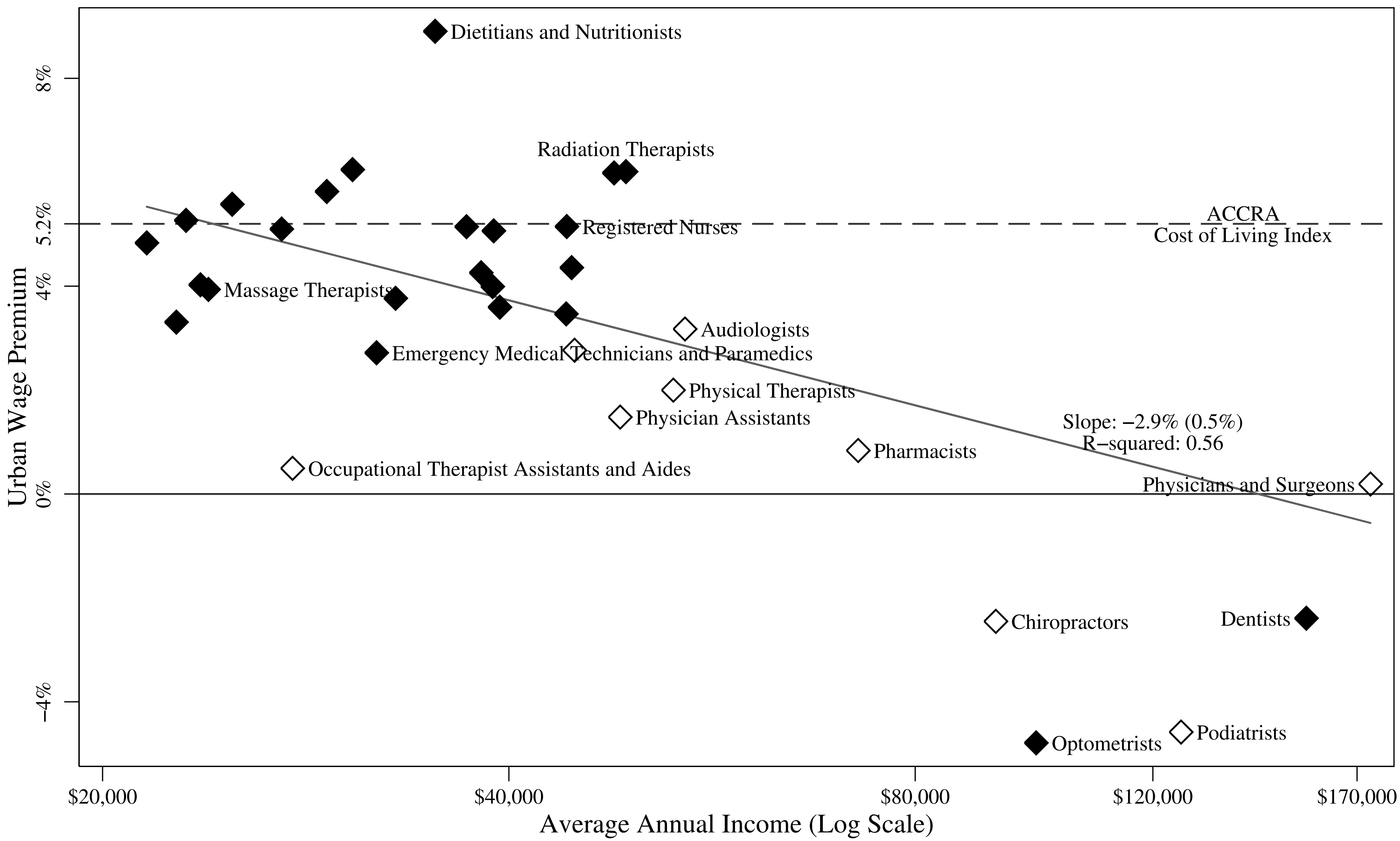
Wheeler, C. H. 2006. Cities and the growth of wages among young workers: Evidence from the nlsy. Journal of Urban Economics, 60(2):162 – 184.

Figure 1. Urban Wage Premium and Skill



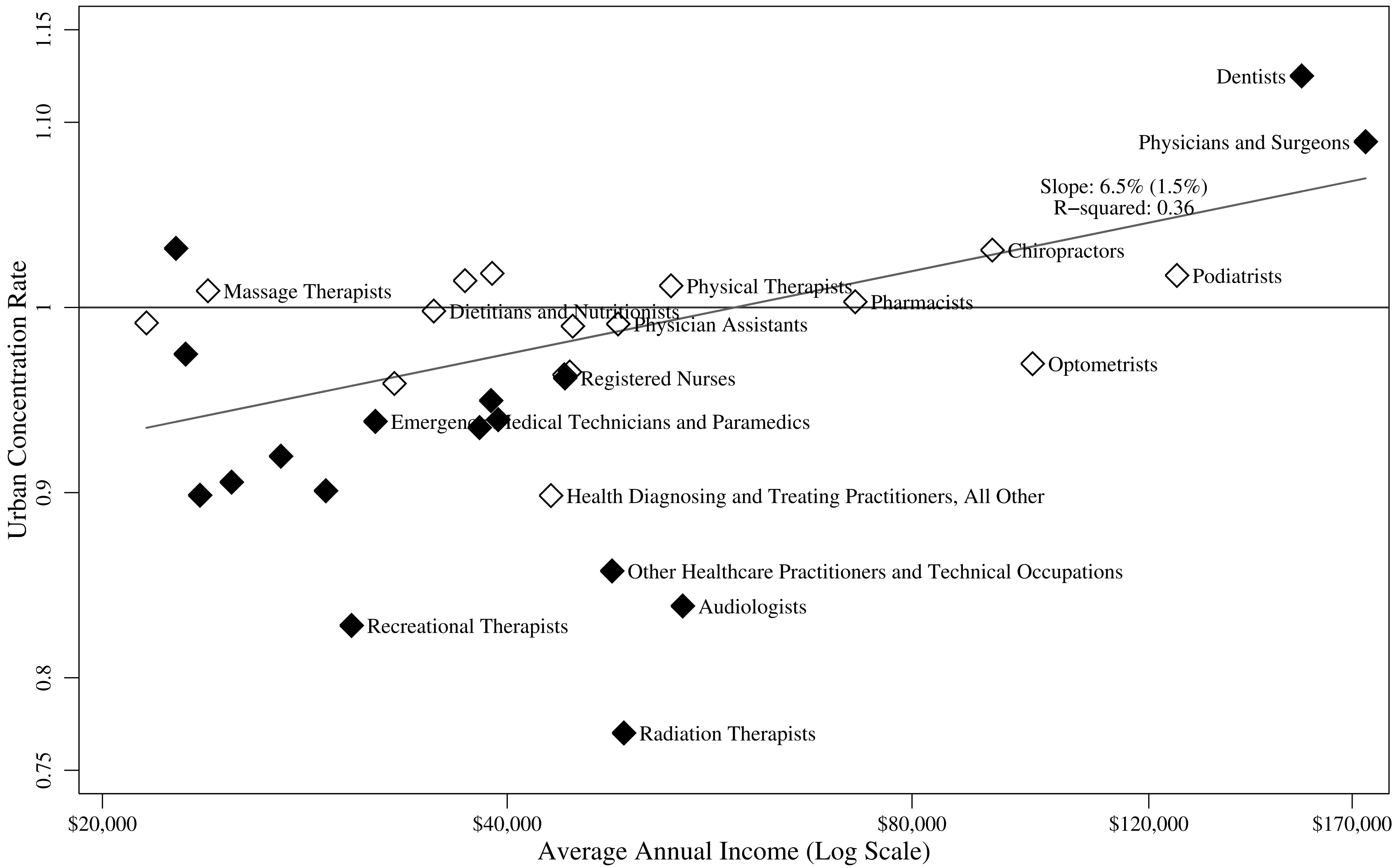
1. Solid symbols indicate that the urban wage premiums are significantly different from zero at a 95 percent confidence level. See Table 1 for precise estimate values.
2."Health Diagnosing and Treating Practioners" ($41,762, 14%) is dropped for visual clarity.

Figure 2. Urban Wage Premium and Skill
Mincer Regressions



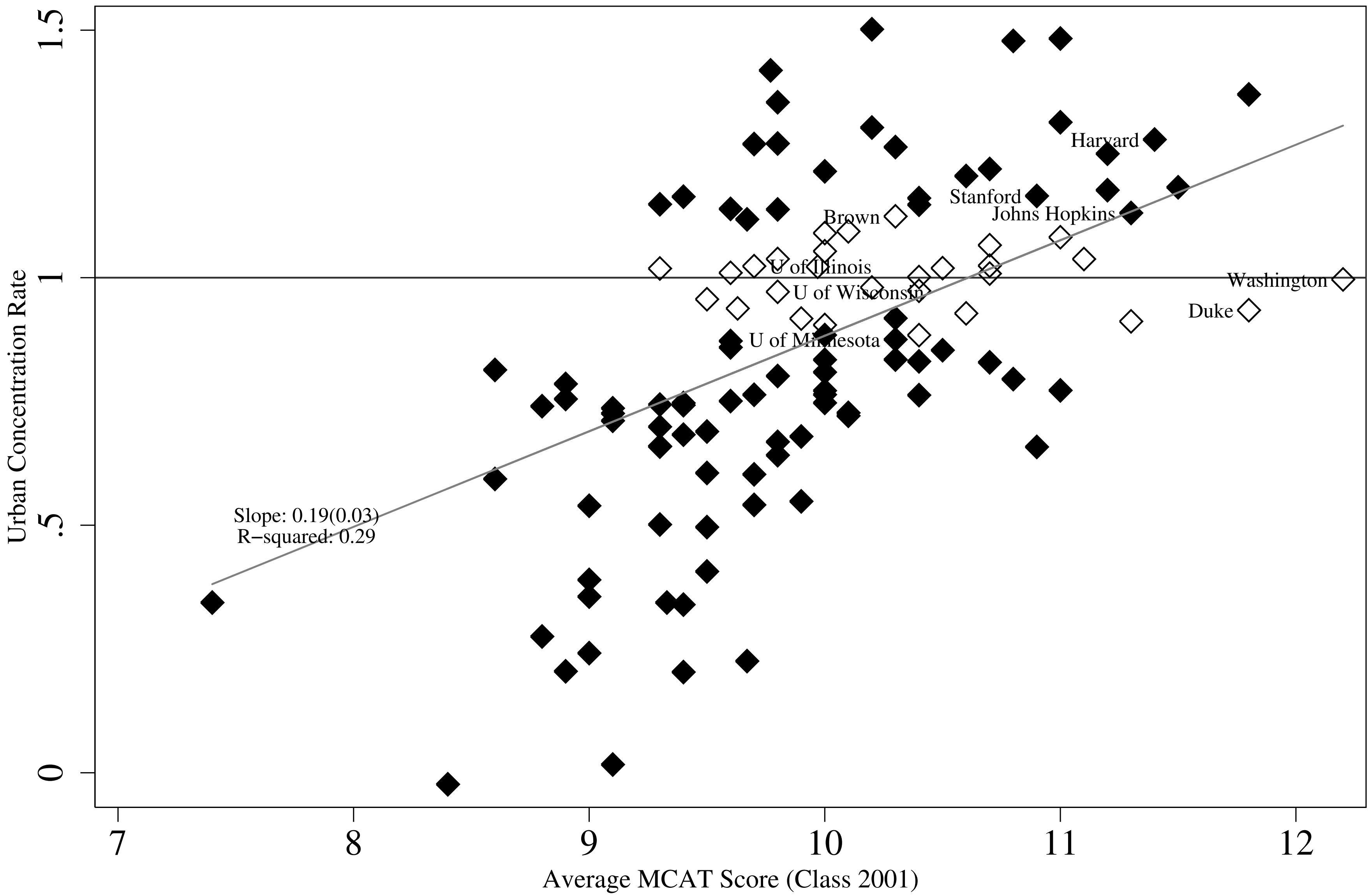
1. Solid symbols indicate that the urban wage premiums are significantly different from zero at a 95 percent confidence level. See Table 2 for precise estimate values.
2. "Health Diagnosing and Treating Practioners" ($41,762, 14%) is dropped for visual clarity.

Figure 3. Urban Concentration Rate and Skill



1. Solid symbols indicate that the urban wage premiums are significantly different from one at a 95 percent confidence level. See Table 1 for precise estimate values.
2. "Radiation Therapists" ($47,340, 0.67) and "Occupational Therapist Assistants and Aides" ($28,363, 0.48) is dropped for visual clarity.

Figure 4. Urban Concentration Rate and Average MCAT Score



Solid symbols indicate that the urban wage premiums are significantly different from one at a 95 percent confidence level.

Figure 5. Geographic Distribution Of Doctors
University of Illinois Medical School Graduates

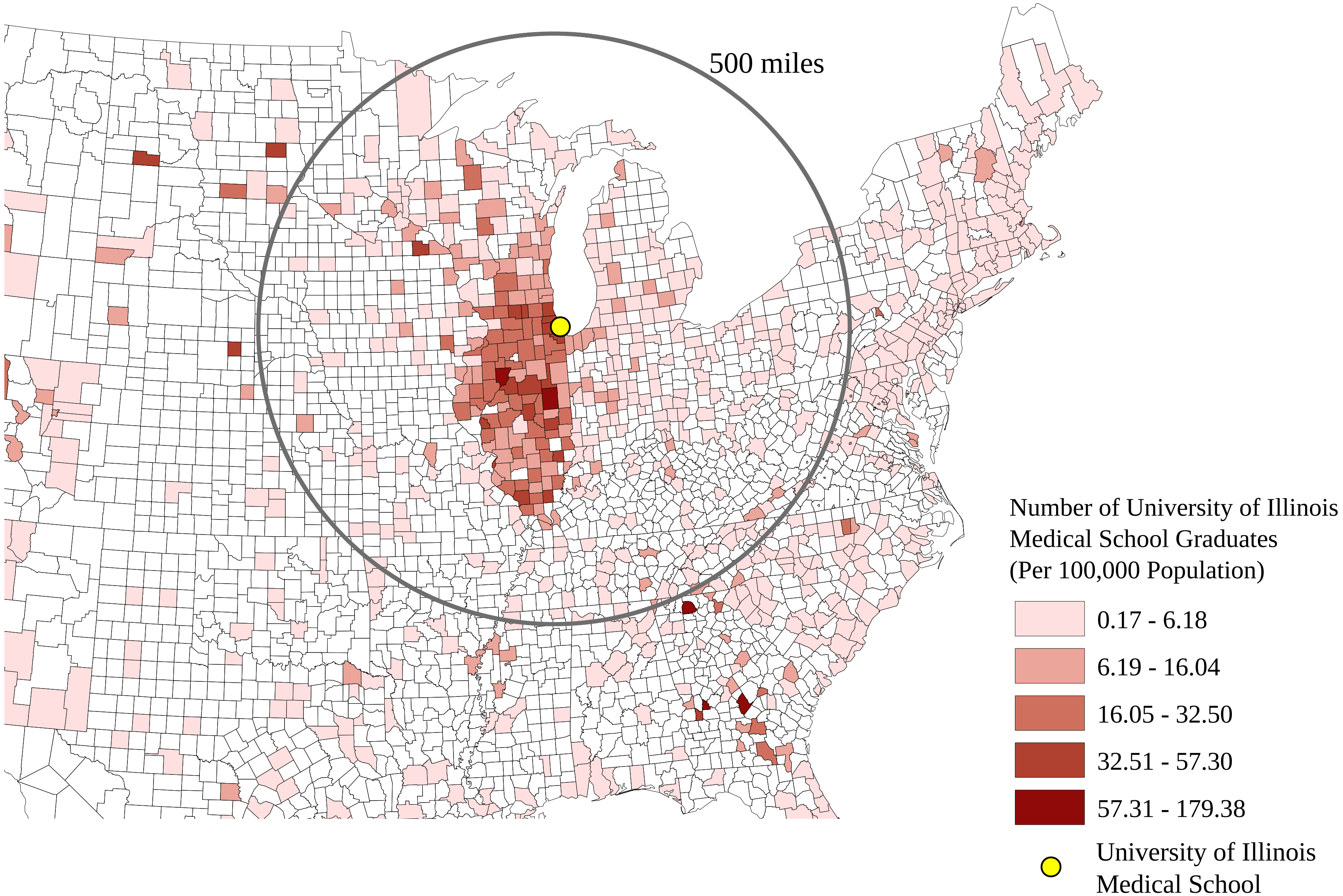
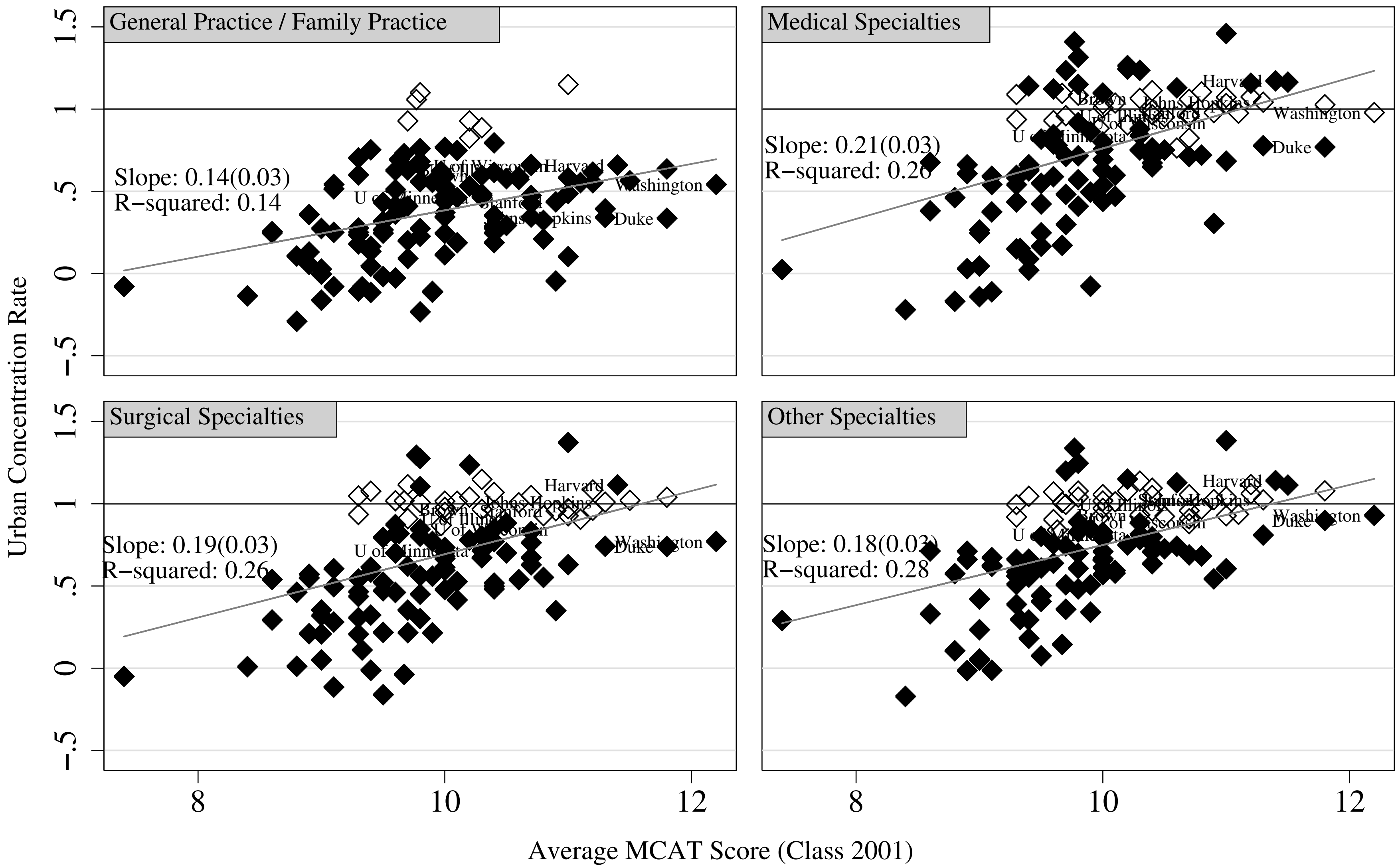
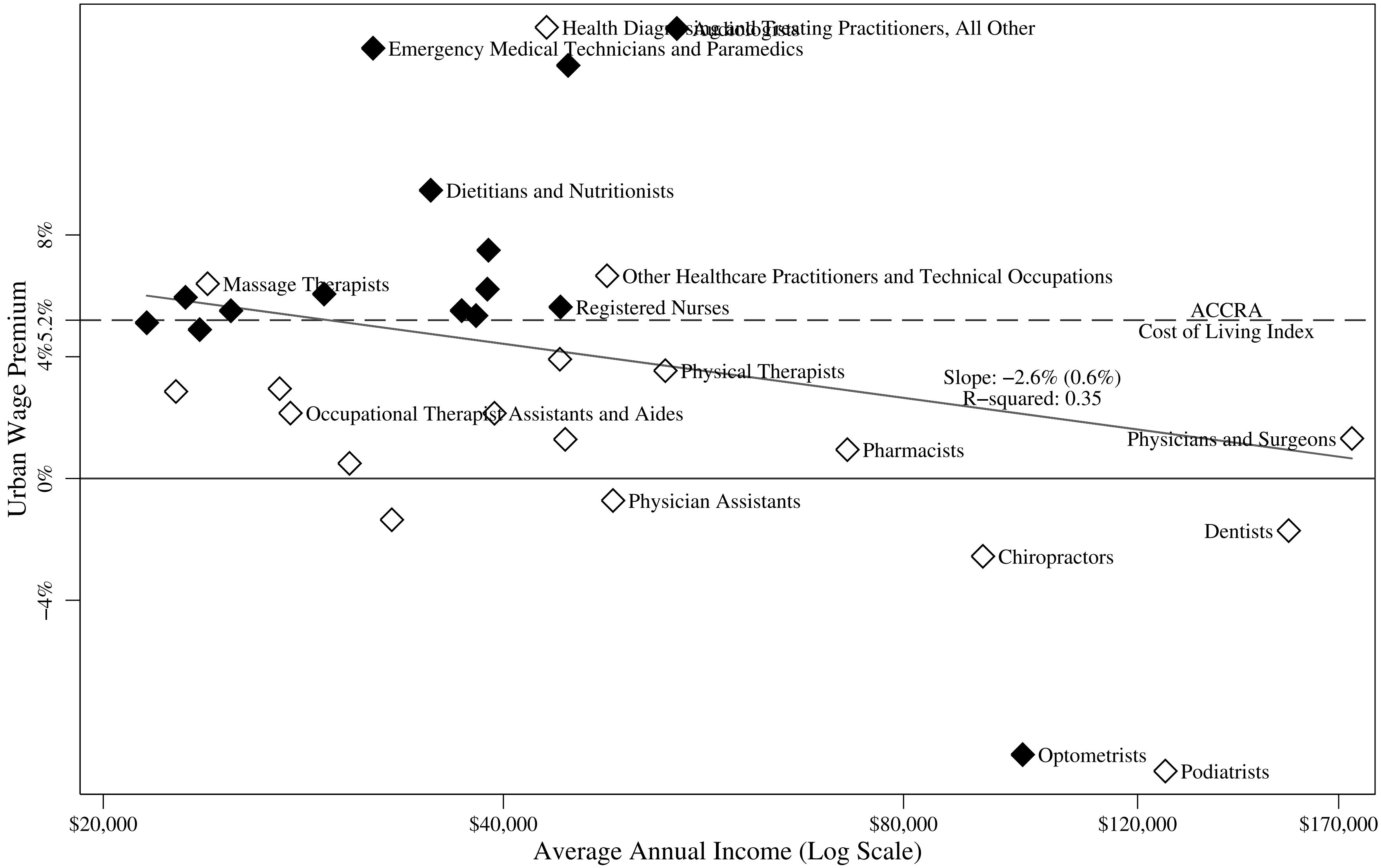


Figure 6. Urban Concentration Rate and Average MCAT Score
By Specialty, Doctors Practicing more Than 500 miles Away from their Medical Schools



Solid symbols indicate that the urban wage premiums are significantly different from one at a 95 percent confidence level.

Figure 7. The Relationship between Urban Wage Premium and Skill
Mincer Regressions for Workers Aged between 50 and 65



Solid symbols indicate that the urban wage premiums are significantly different from zero at a 95 percent confidence level.

Table 1. Urban Wage Premiums and Urban Concentration Rates

Occupation	Average Income	Urban Wage Premium		Urban Concentration Rate
		No Controls	With Controls	
Physicians and Surgeons	$173,951	-3.5% (0.96%)	0.2% (0.53%)	1.09 (0.01)
Dentists	$155,898	-5.1% (1.02%)	-2.4% (0.84%)	1.13 (0.02)
Podiatrists	$125,916	-6.6% (4.07%)	-4.6% (3.81%)	1.02 (0.05)
Optometrists	$98,340	-7.6% (1.69%)	-4.8% (1.58%)	0.97 (0.02)
Chiropractors	$91,792	-3.0% (2.86%)	-2.5% (2.22%)	1.03 (0.03)
Pharmacists	$72,582	0.0% (1.09%)	0.8% (1.00%)	1.00 (0.02)
Audiologists	$54,021	0.3% (3.15%)	3.2% (2.68%)	0.84 (0.04)
Physical Therapists	$52,953	1.8% (1.62%)	2.0% (1.41%)	1.01 (0.02)
Radiation Therapists	$48,842	3.5% (1.91%)	6.2% (1.55%)	0.77 (0.04)
Physician Assistants	$48,361	-1.7% (1.41%)	1.5% (1.23%)	0.99 (0.02)
Other Healthcare Practitioners and Technical Occupations	$47,865	6.5% (2.09%)	6.2% (1.53%)	0.86 (0.03)
Occupational Therapists	$44,745	3.6% (1.73%)	2.8% (1.57%)	0.99 (0.03)
Miscellaneous Health Technologists and Technicians	$44,521	5.2% (1.24%)	4.4% (1.03%)	0.97 (0.02)
Registered Nurses	$44,154	5.1% (0.87%)	5.1% (0.78%)	0.96 (0.01)
Speech-Language Pathologists	$44,110	2.5% (1.65%)	3.5% (1.63%)	0.96 (0.03)
Health Diagnosing and Treating Practitioners, All Other	$43,108	13.6% (5.77%)	13.6% (5.24%)	0.90 (0.09)
Respiratory Therapists	$39,386	4.1% (1.39%)	3.6% (1.33%)	0.94 (0.02)
Therapists, All Other	$38,977	4.0% (1.11%)	5.1% (1.08%)	1.02 (0.02)
Diagnostic Related Technologists and Technicians	$38,904	4.1% (0.65%)	4.0% (0.70%)	0.95 (0.01)
Dental Hygienists	$38,149	4.5% (1.11%)	4.3% (0.99%)	0.94 (0.02)
Clinical Laboratory Technologists and Technicians	$37,210	5.2% (0.70%)	5.1% (0.72%)	1.01 (0.02)
Dietitians and Nutritionists	$35,270	7.0% (1.20%)	8.9% (1.10%)	1.00 (0.02)
Opticians, Dispensing	$32,970	5.1% (2.36%)	3.8% (1.41%)	0.96 (0.02)
Emergency Medical Technicians and Paramedics	$31,915	1.6% (1.13%)	2.7% (1.00%)	0.94 (0.03)
Recreational Therapists	$30,638	6.9% (1.96%)	6.2% (2.09%)	0.83 (0.06)
Licensed Practical and Licensed Vocational Nurses	$29,319	5.4% (0.80%)	5.8% (0.80%)	0.90 (0.02)
Occupational Therapist Assistants and Aides	$27,657	-1.5% (2.74%)	0.5% (2.92%)	0.51 (0.07)
Physical Therapist Assistants and Aides	$27,146	5.4% (1.27%)	5.1% (1.36%)	0.92 (0.03)
Medical Records and Health Information Technicians	$24,949	6.4% (2.16%)	5.6% (1.93%)	0.91 (0.02)
Massage Therapists	$23,960	4.0% (1.42%)	3.9% (1.39%)	1.01 (0.03)
Health Diagnosing and Treating Practitioner Support Technicians	$23,636	4.8% (0.84%)	4.0% (0.71%)	0.90 (0.01)
Medical Assistants and Other Healthcare Support Occupations	$23,060	4.0% (0.81%)	5.3% (0.87%)	0.97 (0.01)
Dental Assistants	$22,683	2.5% (1.10%)	3.3% (0.81%)	1.03 (0.01)
Nursing, Psychiatric, and Home Health Aides	$21,565	5.5% (0.58%)	4.8% (0.68%)	0.99 (0.02)

Source: Census 2000 Public Use Microdata Sample

Table 2. Mincer Regression Estimates
Census Data

Occupation	Urban Wage Premium	Self-Employed	Female	White	Age	Age Squared	Logged Working Hours	Married	Having Child	N	R-squared
Physicians and Surgeons	0.00	0.10	-0.32	0.13	0.24	-0.002	0.29	-0.10	0.10	21,144	0.28
	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.00)	(0.03)	(0.02)	(0.02)		
Dentists	-0.02	0.17	-0.45	0.16	0.11	-0.001	0.30	-0.12	0.09	4,870	0.15
	(0.01)	(0.04)	(0.04)	(0.05)	(0.01)	(0.00)	(0.07)	(0.04)	(0.03)		
Podiatrists	-0.05	0.10	-0.38	0.28	0.22	-0.002	0.29	-0.03	0.12	386	0.22
	(0.04)	(0.09)	(0.16)	(0.23)	(0.06)	(0.00)	(0.14)	(0.12)	(0.10)		
Optometrists	-0.05	0.06	-0.21	0.13	0.05	0.000	0.63	-0.05	0.13	876	0.22
	(0.02)	(0.05)	(0.06)	(0.07)	(0.02)	(0.00)	(0.12)	(0.07)	(0.05)		
Chiropractors	-0.02	0.13	-0.32	0.07	0.16	-0.002	0.49	-0.19	0.13	1,571	0.15
	(0.02)	(0.05)	(0.06)	(0.12)	(0.03)	(0.00)	(0.15)	(0.06)	(0.05)		
Pharmacists	0.01	0.01	-0.18	0.07	0.05	0.000	0.75	-0.06	-0.03	5,440	0.19
	(0.01)	(0.05)	(0.02)	(0.04)	(0.01)	(0.00)	(0.04)	(0.02)	(0.02)		
Audiologists	0.03	0.09	-0.24	0.18	0.09	-0.001	0.79	-0.09	0.04	310	0.25
	(0.03)	(0.19)	(0.14)	(0.09)	(0.04)	(0.00)	(0.18)	(0.07)	(0.07)		
Physical Therapists	0.02	0.14	-0.11	0.01	0.09	-0.001	0.99	-0.06	0.04	3,704	0.24
	(0.01)	(0.05)	(0.02)	(0.04)	(0.01)	(0.00)	(0.05)	(0.02)	(0.03)		
Radiation Therapists	0.06	0.89	-0.13	0.07	0.06	-0.001	1.37	-0.10	0.04	295	0.38
	(0.02)	(0.41)	(0.07)	(0.08)	(0.02)	(0.00)	(0.17)	(0.06)	(0.05)		
Physician Assistants	0.01	0.08	-0.27	0.14	0.11	-0.001	0.83	-0.04	-0.02	1,465	0.21
	(0.01)	(0.27)	(0.05)	(0.05)	(0.02)	(0.00)	(0.12)	(0.05)	(0.05)		
Other Healthcare Practitioners and Technical Occupations	0.06	0.07	-0.22	0.18	0.09	-0.001	1.02	-0.12	0.00	949	0.30
	(0.02)	(0.09)	(0.05)	(0.07)	(0.02)	(0.00)	(0.13)	(0.05)	(0.07)		
Occupational Therapists	0.03	0.24	-0.15	-0.08	0.08	-0.001	0.95	-0.06	-0.02	1,556	0.24
	(0.02)	(0.06)	(0.04)	(0.06)	(0.02)	(0.00)	(0.06)	(0.03)	(0.03)		
Miscellaneous Health Technologists and Technicians	0.04	0.38	-0.34	0.15	0.05	0.000	0.93	-0.05	0.09	2,031	0.29
	(0.01)	(0.09)	(0.03)	(0.04)	(0.01)	(0.00)	(0.13)	(0.04)	(0.03)		
Registered Nurses	0.05	-0.11	-0.10	0.02	0.06	-0.001	0.80	0.03	0.00	56,720	0.20
	(0.01)	(0.03)	(0.01)	(0.02)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)		
Speech-Language Pathologists	0.03	0.03	-0.11	0.05	0.06	-0.001	0.81	0.00	-0.02	1,328	0.19
	(0.02)	(0.07)	(0.07)	(0.05)	(0.02)	(0.00)	(0.09)	(0.04)	(0.04)		
Health Diagnosing and Treating Practitioners, All Other	0.14	-0.51	-0.17	0.00	0.16	-0.002	0.71	0.51	0.44	270	0.19
	(0.05)	(0.17)	(0.14)	(0.19)	(0.05)	(0.00)	(0.24)	(0.13)	(0.12)		
Respiratory Therapists	0.04	-0.07	-0.12	0.04	0.03	0.000	0.76	0.00	-0.01	2,281	0.20
	(0.01)	(0.17)	(0.02)	(0.03)	(0.01)	(0.00)	(0.06)	(0.02)	(0.02)		
Therapists, All Other	0.05	0.10	-0.07	0.07	0.09	-0.001	0.83	-0.09	-0.04	1,608	0.20
	(0.01)	(0.06)	(0.05)	(0.05)	(0.01)	(0.00)	(0.10)	(0.04)	(0.03)		
Diagnostic Related Technologists and Technicians	0.04	0.20	-0.15	0.06	0.07	-0.001	0.89	0.00	0.00	6,286	0.25
	(0.01)	(0.07)	(0.02)	(0.02)	(0.01)	(0.00)	(0.04)	(0.01)	(0.01)		

Source: Census 2000 5 Percent Public Use Microdata Sample

Table 2. Continued

Occupation	Urban Wage Premium	Self-Employed	Female	White	Age	Age Squared	Logged Working Hours	Married	Having Child	N	R-squared
Dental Hygienists	0.04	-0.10	-0.08	0.10	0.07	-0.001	0.79	0.04	-0.06	3,156	0.17
	(0.01)	(0.09)	(0.10)	(0.05)	(0.01)	(0.00)	(0.06)	(0.02)	(0.02)		
Clinical Laboratory Technologists and Technicians	0.05	0.16	-0.11	0.09	0.07	-0.001	0.77	-0.03	-0.02	7,150	0.22
	(0.01)	(0.13)	(0.02)	(0.02)	(0.01)	(0.00)	(0.04)	(0.02)	(0.02)		
Dietitians and Nutritionists	0.09	-0.09	-0.07	0.37	0.07	-0.001	1.02	-0.01	0.06	1,660	0.23
	(0.01)	(0.11)	(0.07)	(0.04)	(0.01)	(0.00)	(0.07)	(0.04)	(0.04)		
Opticians, Dispensing	0.04	0.17	-0.25	0.09	0.06	-0.001	0.86	-0.01	0.00	1,437	0.24
	(0.01)	(0.08)	(0.03)	(0.06)	(0.01)	(0.00)	(0.11)	(0.03)	(0.04)		
Emergency Medical Technicians and Paramedics	0.03	0.21	-0.20	0.12	0.07	-0.001	0.76	-0.04	0.00	1,579	0.26
	(0.01)	(0.26)	(0.03)	(0.06)	(0.01)	(0.00)	(0.08)	(0.03)	(0.02)		
Recreational Therapists	0.06	-0.03	-0.16	0.00	0.09	-0.001	0.96	0.08	0.03	325	0.21
	(0.02)	(0.21)	(0.09)	(0.09)	(0.03)	(0.00)	(0.25)	(0.05)	(0.06)		
Licensed Practical and Licensed Vocational Nurses	0.06	-0.32	-0.14	0.09	0.05	0.000	0.84	0.02	0.02	12,622	0.17
	(0.01)	(0.07)	(0.02)	(0.01)	(0.00)	(0.00)	(0.03)	(0.01)	(0.01)		
Occupational Therapist Assistants and Aides	0.00	-0.63	0.14	-0.06	0.04	0.000	0.93	0.03	0.04	210	0.27
	(0.03)	(0.17)	(0.11)	(0.09)	(0.03)	(0.00)	(0.15)	(0.07)	(0.07)		
Physical Therapist Assistants and Aides	0.05	-0.08	-0.10	0.13	0.09	-0.001	0.90	0.03	0.06	1,077	0.20
	(0.01)	(0.25)	(0.04)	(0.06)	(0.02)	(0.00)	(0.09)	(0.04)	(0.04)		
Medical Records and Health Information Technicians	0.06	0.36	-0.12	0.06	0.07	-0.001	1.00	0.03	-0.07	2,238	0.20
	(0.02)	(0.20)	(0.06)	(0.04)	(0.01)	(0.00)	(0.11)	(0.03)	(0.03)		
Massage Therapists	0.04	-0.09	-0.12	0.10	0.06	-0.001	0.60	0.09	-0.04	1,742	0.09
	(0.01)	(0.04)	(0.05)	(0.08)	(0.01)	(0.00)	(0.06)	(0.05)	(0.05)		
Health Diagnosing and Treating Practitioner Support Technicians	0.04	-0.14	-0.17	0.02	0.07	-0.001	1.01	0.05	-0.02	6,402	0.23
	(0.01)	(0.33)	(0.02)	(0.02)	(0.00)	(0.00)	(0.05)	(0.02)	(0.02)		
Medical Assistants and Other Healthcare Support Occupations	0.05	-0.02	-0.11	0.11	0.07	-0.001	1.06	-0.01	-0.03	14,050	0.23
	(0.01)	(0.04)	(0.02)	(0.01)	(0.00)	(0.00)	(0.03)	(0.01)	(0.01)		
Dental Assistants	0.03	-0.23	-0.21	0.13	0.08	-0.001	0.88	-0.02	-0.06	5,568	0.17
	(0.01)	(0.22)	(0.09)	(0.02)	(0.01)	(0.00)	(0.04)	(0.02)	(0.01)		
Nursing, Psychiatric, and Home Health Aides	0.05	-0.12	-0.21	0.07	0.05	0.000	0.89	-0.02	0.01	33,516	0.15
	(0.01)	(0.04)	(0.02)	(0.01)	(0.00)	(0.00)	(0.03)	(0.01)	(0.01)		

Table 3. Urban Wage Premiums and Urban Concentration Rates
Across Doctors with Different Specialties

Occupation	Average Annual Net Income ($1,000)	Number of Observations	Urban Wage Premium	Urban Wage Premium (Mincer Regression)	Urban Concentration Rate
All	179	9,710	-5.3% (0.8%)	-3.1% (0.6%)	1.07 (0.02)
Surgical Specialties	232	1,394	-7.1% (1.9%)	-4.4% (1.5%)	1.04 (0.02)
Medical Specialties	171	5,470	-4.7% (1%)	-3.3% (1%)	1.14 (0.02)
Other Specialties	141	408	-2.4% (3.2%)	-1.5% (2.7%)	1.09 (0.02)
General Practice / Family Physicians	131	2,438	-3.9% (1.3%)	-1.9% (1.2%)	0.90 (0.02)

Source: Community Tracking Study Physician Survey 2000-2001

Table 4 Mincer Regression Estimates
CTS Data

	All	General Practice / Family Physicians	Other Specialties	Medical Specialties	Surgical Specialties
Average Annual Income (In Thousand Dollars)	179	131	141	171	232
Urban Wage Premium (Logged Population Size)	-3.28% (0.64%)	-1.94% (1.22%)	-1.37% (2.63%)	-3.27% (0.98%)	-4.27% (1.51%)
Logged Working Hours	0.33 (0.03)	0.38 (0.06)	1.01 (0.18)	0.19 (0.04)	0.38 (0.06)
Age	0.06 (0.01)	0.03 (0.01)	0.05 (0.03)	0.07 (0.01)	0.06 (0.02)
Squared Age	-0.0005 (0.0001)	-0.0003 (0.0001)	-0.0004 (0.0003)	-0.0007 (0.0001)	-0.0006 (0.0002)
Dummy variables:					
Female	-0.33 (0.03)	-0.34 (0.06)	-0.14 (0.09)	-0.36 (0.05)	-0.27 (0.09)
Non-White	-0.03 (0.04)	0.01 (0.05)	0.09 (0.10)	0.02 (0.05)	-0.23 (0.10)
Foreign Medical School Graduates	-0.02 (0.04)	0.01 (0.04)	-0.04 (0.10)	-0.06 (0.05)	-0.02 (0.07)
Board Certified	0.18 (0.03)	0.19 (0.04)	0.15 (0.10)	0.19 (0.06)	0.21 (0.06)
Owner of Hospital	0.09 (0.02)	0.03 (0.04)	-0.06 (0.09)	0.14 (0.04)	0.06 (0.04)
Working for Universities	0.01 (0.07)	-0.14 (0.07)	0.44 (0.18)	0.08 (0.08)	-0.06 (0.21)
General Practice / Family Physicians	-0.08 (0.05)				
Medical Specialties	0.13 (0.05)				
Surgical Specialties	0.36 (0.05)				
R^2	0.17	0.11	0.23	0.12	0.13
N	9682	2436	408	5445	1393

Source: Community Tracking Study (CTS) Physician Survey 2000-2001