Ability Sorting and Consumer City

Sanghoon Lee*

University of Minnesota and Federal Reserve Bank of Minneapolis

February 2, 2005

Abstract

Average wages tend to increase with city size. Most explanations of this urban wage premium focus on productivity spillovers. This paper proposes a consumption-side explanation. The claim is that the wide consumption variety found in large cities is more important to high-skill (hence high-income) workers than low-skill workers, and thus the higher wages found in large cities are due to the selection of high skill workers choosing to live there. A testable implication of my theory, distinguished from productivity-based theories, is that urban wage premiums may be negative for high-skill workers. This implication is confirmed by data on the medical profession. At the top skill level, there is substantial urban wage discount: doctors in large cities are paid 9 percent less than their peers in small cities.

*Address: Department of Economics, University of Minnesota, 1035 Heller Hall, 271-19th Avenue South, Minneapolis, MN 55455; E-mail: sanghoon@econ.umn.edu. I am deeply indebted to my advisor Thomas Holmes for his guidance and patience. I also thank James Schmitz, Samuel Kortum, Zvi Eckstein, Andrea Moro, Robert Town, Jonathan Eaton, Hwikwon Ham, Kim-Sau Chung, Kim Ruhl, Mark Gibson, and participants in applied microeconomics workshop at university of Minnesota for valuable suggestions and discussions. All errors are mine.

1 Introduction

There is large urban economics literature on the connection between average pay and city size. Empirical work has documented that workers in large cities are paid substantially more than those in small cities, with differences on the order of 30 percent or more. (e.g., Glaeser and Mare (2001), Rauch (1993)). Theoretical work has emphasized the importance of productivity spillover in accounting for why workers in large cities may be more productive (e.g. Lucas (1988), Henderson (1974)). These productivity side theories certainly capture one part of the story, but they tend to neglect another important aspect of cities.

Large cities have wide consumption variety. They have museums, professional sports teams, and nice French restaurants, which small cities do not have. An extensive literature accounts for this wide consumption variety in cities through scale economies (see Fujita, Krugman, and Venables (1999) and Fujita and Thisse (2002) for recent textbook treatments). However, this literature on variety has not connected with the literature on the urban wage premium. In fact, there is a tension here. Since rural areas have little product variety, we might even expect a compensating premium to be paid to workers in rural areas.

This paper presents a consumption-side explanation for the urban wage premium that accounts for both the substantial variety found in large cities and the higher pay of the workers that live there. The theory has two crucial ingredients. First, workers vary in skill and hence earning power. Second, preferences are such that product variety found in large cities 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 average high wage found in large cities in the model is due to the selection of the relatively higher-skill individuals choosing to live there.

The theory also delivers a number of new implications that I can examine with the data. The most striking implication is that the wage differential between large cities and small cities, *for* *fixed skill*, should be decreasing in skill. In fact, for high enough skill levels, the differential turns negative and workers in large cities are paid *less* than their small city counterparts. Such workers, with great earning power and thus great demand for urban amenities, have to be heavily bribed to move to rural areas.

I examine this implication using data from the health care sector. I picked the health care sector because it is the best example of an industry for which it is absolutely necessary to employ extremely high skill workers (doctors) in small cities. This is different from the legal industry, for example, where there is no need for patent attorneys in small cities. I find that doctors who live in large cities are paid substantially *less* than their peers in small cities. Doctors in large cities are paid 9 percent *less* than doctors in small cities. Moreover, across doctors, those in the highest paid specialties get the largest pay discounts. Surgeons in large cities are paid 18 percent *less* than surgeons in small cities. Across the health care sector as a whole, the urban wage premium sharply decreases in skill levels. Nurses are paid more in large cities. Doctors, dentists, etc. are paid less in large cities.

The finding that doctors are paid less in large cities is particularly striking because it is sensible to expect that doctors might vary in quality and the better ones would locate in larger cities, in accordance with my theory. In fact, I provide evidence of such selection using data on doctors from different medical schools of varying quality. I find that doctors from better medical schools tend to be more concentrated in large cities.

A second implication from the theory is that firms in large cities substitute relatively cheaper high skill workers for relatively more expensive low-skill workers, as compared to firms in small cities. In the health care sector I find that high skill professions such as doctors are more concentrated in large cities, as compared with relatively low skill professions such as nurses, and that this selection drives the average pay of health care workers to be substantially higher in large cities.

In order to see the importance of this selection effect on the urban wage premium, I decompose urban wage premium for the whole health care sector into wage variation within occupation, quantity variation across occupations, and covariance term. The second term - the quantity variation across occupations - captures the selection effect, i.e. the effect that large cities have relatively more doctors than nurses, as compared with small cities. I find that 63% of the urban wage premium is accounted for by this selection effect.

The assumption that consumption variety is an income-elastic good plays the key role in my theory. I am unaware of any previous empirical work that directly estimates this parameter, but there exists previous empirical work that indirectly substantiates this claim. First, there is evidence that high-income people consume more high-quality goods and that cities have higher-quality products, which together would imply my claim. Bils and Klenow (2001) show that high-income people consume high-quality goods by constructing Engel curves for quality. Berry and Waldfogel (2003) show that the range of qualities on offer substantially increases in city size looking at the restaurant industry. The second line of evidence that substantiates the claim comes from estimates on income elasticity of the demand for residential land. What turns out to be crucial for my theory is that expenditure shares on the variety of goods relative to land increase with income, i.e. that the demand for variety is more income elastic than the demand for land. (Higher variety is the gain for cities, higher land price is the cost, so differences in expenditures shares are key.) There is strong evidence that the demand for land is income inelastic. Glaeser, Kahn, and Rappaport (2000) find that the income elasticity of demand for land is unlikely to be more than 0.4.

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, Kolko, and Saiz (2000), Costa and Kahn (2000), Tabuchi and Yoshida (2000)). However, the view that cities are better places for living implies that there should be, on average, an urban wage discount after adjustments for cost of living difference across urban and rural areas, because workers in rural areas have to be compensated for the lack of urban amenities.[1] This does not hold in data.[2] This paper reconciles the tension by introducing different levels of skills and ability sorting.

There are other papers in the empirical urban wage premium literature that have argued that ability sorting may play a role in accounting for the observed urban wage premium (e.g. Glaeser and Mare (2001)). This paper differs from the literature in two ways. First, I propose an explicit theoretical mechanism for why the selection takes place. Second, I account for the fact that the actual premium can be dramatically different for different skills. Previous work that I am citing here assumes a constant relationship between city size and skill. My key finding is that the size-skill relationship depends in an important way on skill.

The rest of paper proceeds as follows. Section 2 provides the model and its theoretical implications. Section 3 provides evidence for the theoretical implications using the health care sector as a whole. First, urban wage premiums are decreasing in skill across different professions, and are negative for high-skill professions including doctors and dentists. Second, high-skill professions like doctors are more concentrated in large cities than low-skill professions. Section 4 provides evidence only with doctors, but across different specialties. Urban wage premiums are negative across all specialties, and are even lower for doctors in high-pay specialties. Section 5 provides the evidence for physician quality even within specialties. Large cities have better doctors, as measured by their medical school quality. Section 6 provides the urban wage premium decomposition for the whole health care sector. Section 7 concludes.

[1] See Roback (1982).

[2] See Glaeser and Mare (2001).

2 The Theory

This section provides a partial equilibrium model of location decisions in the medical profession, and studies its implications on the equilibrium wage and skill distribution across different-size cities.

2.1 The Model

There is a continuum of cities, and they are indexed by their population sizes $n \in N$ where $N \subset \mathbb{R}_+$ is an interval. Every economic activity takes places in one of the cities. City of population size n has exogenously given consumption variety level $v(n)$ and rent (or land price) $r(n)$. I assume that large cities have higher consumption variety levels and higher rents.[3]

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

Workers are heterogeneous in their abilities. The heterogeneity in ability is modeled in such a way that workers with ability $\theta \in [\underline{\theta}, \bar{\theta}]$ have an outside option offering utility $\bar{u}(\theta)$, and $\bar{u}(\theta)$ is increasing in θ.

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

Workers first decide which city to live in and whether to take the outside option. Once in city n, ability θ workers get paid the wage $w(\theta, n)$, have to consume one unit of land paying the 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 set this

[3] These assumptions can be easily endogenized in a bigger model. See Krugman (1991) and Duranton, Gilles and Diego Puga (2004) for examples.

price as numeraire.[4] In summary, ability θ workers face the following optimization problem.

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

$$\tilde{U}\left(\theta,n\right)=\max_{q(x)}\left(\int_{0}^{v(n)}q\left(x\right)^{\frac{1}{\mu}}dx\right)^{\mu}\text{ s.t. }\int_{0}^{v(n)}q\left(x\right)dx=w\left(\theta,n\right)-r\left(n\right). \tag{1}$$

where $\mu > 1$ is the local good complementarity parameter. Note that the demand for land is perfectly income inelastic, which makes the demand for consumption variety income elastic.[5] Note also that workers living in city n can consume the local goods only within the range $[0, v(n)]$. The complementarity parameter μ greater than 1 implies that workers care about consumption variety and thus their utilities increase in consumption variety level v.

A city of population size n requires n units of medical service to be locally provided. Medical service production technology is given by the standard CES function. The CES technology requires all ability types for production, and this captures the characteristic of the health care sector that they need to have doctors even in small cities.

The medical service provider in city n solves the following profit maximization problem.

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

where $p^m(n)$ is the unit price for medical service in city n, $l(\theta, n)$ is the measure of type θ workers employed in the medical service sector in city n, and $\gamma > 1$ is the input complementarity parameter.

[4]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. In this case I need to introduce exchange rates between local goods in different cities.

[5]The demand for land does not have to be *perfectly* income inelastic. All the results hold with any income inelastic demand for land.

2.2 Equilibrium

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

To begin, I calculate the indirect utility of ability type θ workers living in city n. Since both the prices and the 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 $x \in [0, v]$ solves the workers' optimization problem (1). Substituting q for $q(x)$ in the optimization problem (1), I obtain the following indirect utility function.

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

Since workers can freely choose where to live and whether to take the outside option, the indirect utilities across all cities have to be equal to the reservation utility $\bar{u}(\theta)$ offered by the outside option.

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

Thanks to the constant returns to scale medical service technology, I can consider one aggregate medical service provider in each city that employs constant shares of ability types and sets a constant price regardless of its output. The first order conditions for medical service producers imply the following share of ability types for production and the medical service price.[6]

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

$$p^m(n) = \left(\int_{\underline{\theta}}^{\bar{\theta}} w(\theta, n)^{\frac{1}{1-\gamma}} d\theta\right)^{1-\gamma}. \tag{5}$$

The medical service market clearing condition is simple. The aggregate medical service provider in city n produces n units of medical service.

$$n = \left(\int_{\underline{\theta}}^{\bar{\theta}} l(\theta, n)^{\frac{1}{\gamma}} d\theta\right)^{\gamma}. \tag{6}$$

An equilibrium of this model consists of the list $\left\{(p^m(n), w(\theta, n), l(\theta, n)) | \theta \in \left[\underline{\theta}, \bar{\theta}\right], n \in N\right\}$ satisfying conditions (2) to (6) for each city n. The wage schedules $w(\theta, n)$ are pinned down by the workers' utility maximization conditions (2) and (3). The geographic distribution of medical workers $l(\theta, n)$ is calculated from conditions (4) and (6). The unit medical prices $p^m(n)$ are calculated from condition (5).

2.3 Implications

This section derives two implications of the model. The first implication is on equilibrium prices that urban wage premiums are decreasing in skill, and may even be negative, and the second implication is on equilibrium quantities that high-ability workers are more concentrated in large cities compared to low skill workers.

First, I derive the first implication on equilibrium prices. To begin, I define urban wage premium, as the percentage by which the wage increases when city size doubles. I later estimate this urban

[6] The medical service price $p^m(n)$ is derived from $p^m(n) = \int_{\underline{\theta}}^{\bar{\theta}} w(\theta, n) \hat{l}(\theta, n) d\theta$ such that $\left(\int_{\underline{\theta}}^{\bar{\theta}} \hat{l}(\theta, n)^{\frac{1}{\gamma}} d\theta\right)^{\gamma} = 1$ and the first order condition $\hat{l}(\theta_1, n) / \hat{l}(\theta_2, n) = (w(\theta_1, n) / w(\theta_2, n))^{\frac{\gamma}{1-\gamma}}$ for any $\theta_1, \theta_2 \in \left[\underline{\theta}, \bar{\theta}\right]$.

wage premium using data.

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

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

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

$$w(\theta, n) = r(n) + \bar{u}(\theta) \cdot \frac{1}{v(n)^{\mu-1}}. \tag{7}$$

The wage schedule (7) shows that there are two types of wage compensations in order to give the same level of utility to the workers of same ability living across different sizes of cities. The first term $r(n)$ captures the compensation for land price, and the second term $\bar{u}(\theta)/v(n)^{\mu-1}$ captures the compensation for consumption variety level. Workers in large cities have to be compensated for the high land prices there, and workers in small cities have to be compensated for the lack of consumption variety. The compensations for land prices are equal across different ability types of workers because workers consume one unit of land regardless of their ability types. However, the compensation for consumption variety increases in workers' ability types because the demand for differentiated local goods is income elastic. Thus, urban wage premiums are positive for low skill workers because the compensations for land prices dominate the compensations for consumption variety, and urban wage premiums are negative for high skill workers because the consumption variety compensations dominate the land price compensations.

Proposition 1 *1) The urban wage premium* $\beta^w(\theta, n)$ *for fixed ability* θ *at city* n *is decreasing in* *ability θ.*

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

2) The urban wage premium $\beta^w(\theta,n)$ for ability type θ at city n is negative if and only if

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

Proof. The first result is obtained by differentiating $\beta^w(\theta,n)$ with respect to θ.

$$\frac{\partial\beta^w(\theta,n)}{\partial\theta} = \frac{\partial}{\partial\theta}\frac{\partial w(\theta,n)}{\partial n}\frac{n}{w(\theta,n)} = -\frac{nv(n)^\mu u'(\theta)\left(v(n)\,r'(n)+(\mu-1)\,r(n)\,v'(n)\right)}{\left(r(n)\,v(n)^\mu+u(\theta)\,v(n)\right)^2} < 0.$$

The second result follows directly from wage equation (7). ■

Now I derive the second implication on equilibrium quantities.. To begin, I define urban concentration rate, as how much percentage the number of workers increases when city size doubles. Later I estimate the urban concentration rates using data.

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

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

The second implication follows straight from the first implication and medical service providers' first order conditions. The first implication means that high skill workers are relatively cheaper in large cities. Medical service providers in large cities substitute these relatively cheaper high ability workers for relatively more expensive low ability workers.

Proposition 2 *The urban concentration rate* $\beta^q(\theta, n)$ *at city* n *is increasing in ability* θ.

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

Proof. I obtain the equilibrium distribution of health care sector workers $l(\theta, n)$ from condition (4) and condition (6).

$$l(\theta, n) = n \cdot \frac{w(\theta, n)^{\frac{\gamma}{1-\gamma}}}{W(n)^{\gamma}}$$

where $W(n) = \int_{\underline{\theta}}^{\bar{\theta}} w(\theta, n)^{\frac{1}{1-\gamma}} d\theta$.

$$\frac{\partial}{\partial\theta}\beta^q(\theta, n) = \frac{\partial}{\partial\theta}\frac{\partial}{\partial \log n}\log\left(n \cdot \frac{w(\theta, n)^{\frac{\gamma}{1-\gamma}}}{W(n)^{\gamma}}\right) = \frac{\gamma}{1-\gamma}\frac{\partial\beta^w(\theta, n)}{\partial\theta} > 0.$$

The last inequality comes from Proposition 1. ∎

3 The Health Care Sector as a Whole

This section provides evidence for the theoretical predictions in Propositions 1 and 2, using data on the whole medical service sector across different occupations. 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 health care sector except veterinarians (34 occupations, Census occupation code 300-365). This paper uses the Metropolitan Statistical Area (MSA) as basic geographic units. The metropolitan areas used here contain about 76 percent of the US population.[7]

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

3.1 The evidence on equilibrium prices

This section confirms the prediction on equilibrium prices in Proposition 1 across different health care professions. I use the average annual income of an occupation as the measure for skill level, and show that urban wage premiums are decreasing in skill across health care professions and are negative for high income professions including doctors and dentists. The negative urban wage premiums for high skill professions are the key evidence in this paper that do not follow from productivity-based theories.

Table 1 reports the average annual income by metropolitan area size for selected professions[8] These professions constitute the top and the bottom three professions in the skill hierarchy. Each group of metropolitan areas has about the same size of population.[9] Rural areas have about 21 percent of the whole U.S. population. Table 1 clearly shows that high-skill workers are paid less in large cities while low-skill workers are paid more. The ratios of incomes between the top third metropolitan areas and the bottom third are strictly less than 1 for all the high-skill professions, while the ratios are greater than 1 for all low skill professions. The same relationship also holds for the top third metropolitan areas and rural areas.

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

$$\log w_i = \alpha^w(\theta) + \beta^w(\theta) \cdot \log n_i \tag{8}$$

where w_i is individual i's annual total income and n_i is the population size of the metropolitan

[8] Refer to Table A.1 for other health care occupations.

[9] The top third metropolitan areas include New York CMSA, Los Angeles CMSA, Chicago CMSA, Washington-Baltimore CMSA, San-Francisco CMSA, Philadelphia CMSA, and Boston CMSA.

area where individual i lives. The regression coefficient $\beta^w(\theta)$ captures the urban wage premium for occupation θ. The second column of Table 2 reports the urban wage premium $\beta^w(\theta)$ and its standard error for the selected professions.[10] The urban wage premiums are all negative for the top three professions and all positive for the bottom three.

Now I report the urban wage premiums for all health care professions. Figure 8 shows the relationship between the urban wage premium and average annual income across all health care professions. There exists a clear negative relationship between the urban wage premium and skill. In addition, urban wage premiums are all negative for high-skill professions with average annual income of $80,000 or more. This confirms the prediction in Proposition 1.

However, this downward pattern and negative premiums for high-skill workers might be due to some other factors that might change systematically with metropolitan area size and average occupation income. I control for other possible factors by running Mincer regression for each occupation. I add to the previous regression the standard control variables such as working hours, age and squared age, sex, and education.[11] The third column of Table 2 reports the urban wage premiums after controlling for these factors for the selected occupations. Figure 8 shows the relationship between the urban wage premium and average annual income for all occupations after controlling for these factors. The downward pattern remains strong. The urban wage premiums for most high skill professions stay negative.[12]

[10] Refer to Table A.1 for other health care occupations.

[11] I add as regressors, logged total hours worked, age, squared age, dummy variable for sex being female, dummy for race being not white, and six dummies for education levels. The six education groups include 8th grade or lower, 12th grade or lower, high school graduates, some college or associates degree, bachelor degree, and post college degree.

[12] The urban wage premium for doctors becomes virtually zero (0.00071 with standard error 0.00354). The urban wage premium for doctors seems to increase in the Mincer regression because small cities have relatively fewer female doctors and more old doctors. However, the urban wage premium is expected to go negative once I control for doctors' specialties on which the census data do not have information. Specialists on average make more income than generalists and large cities have relatively more specialists. In addition, an urban wage premium of zero is still a large discount once the price level is considered.

3.2 The evidence on equilibrium quantities

This section confirms the prediction on equilibrium quantities in Proposition 2 across different health care professions. I show that high-income professions are more concentrated in large cities compared to low-income professions. Table 3 shows the number of workers per 100,000 population by the size of metropolitan areas for the top and the bottom three professions in the health care sector.[13] It is clear that the top three professions are more concentrated in large cities than are the bottom three. The count ratio between the top third of metropolitan areas and the bottom third are uniformly higher for the high-skill medical professions. The same relationship holds between the top third of metropolitan areas and rural areas.

I summarize the results using the urban concentration rates defined in Definition 2. I assume that the elasticities are constant across the cities, and calculate the urban concentration rate $\beta^q(\theta)$ for occupation θ by running the following metropolitan-area-level regression for each profession θ.

$$\log l(\theta, n_j) = \alpha^q(\theta) + \beta^q(\theta) \cdot \log n_j \tag{9}$$

where $l(\theta, n_j)$ is the number of workers in occupation θ in metropolitan area j, and n_j is the population size of metropolitan area j. The regression coefficient $\beta^q(\theta)$ captures the urban concentration rate for profession θ. The urban concentration rate $\beta^q(\theta)$ is equal to 1 if workers in profession θ are proportionately distributed to population size. The urban concentration rate $\beta^q(\theta)$ is greater than 1 if workers in profession θ are disproportionately concentrated in large cities, and vice versa. The fourth column of Table 2 reports the urban concentration rate $\beta^q(\theta)$ and its standard error for the top and bottom three occupations.[14] The urban concentration rates are higher for high skill workers as Proposition 2 predicted.

[13] Refer to Table A.2 for other health care occupations.

[14] Refer to Table A.2 for other health care occupations.

Now I report the urban concentration rates for all health care professions. Figure 8 shows the relationship between the urban concentration rate and average income for each occupation. There exists a clear positive relationship between the urban concentration rate and income, and this confirms the theoretical prediction in Proposition 2. In addition, the urban concentration rates are greater than 1 for most high-skill occupations with annual income of $80,000 or more, and less than 1 for most low-skill occupations, which implies that high-skill professions are substituted for low-skill professions in large cities, and vice versa in small cities.

4 A Focus on Physicians

The previous section examined the theoretical implications across different professions. The same idea may be applicable to even within professions. This section examines the theoretical implications only with doctors but across different specialties, and along the way reconfirms the key evidence of this paper that urban wage premiums are negative for doctors.

The primary data sets used in this section are the Community Tracking Study (CTS) physician survey 2000 - 2001 and a year 2000 version of the American Medical Association (AMA) physicians' master file. I use the CTS data set to confirm the implication on equilibrium prices. 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.[15] It has more detailed occupation specific information than the census data, such as specialty, board certification, etc.

I use the AMA data set to confirm the implication on equilibrium quantities. The AMA data set has very detailed information on most physicians in the US, such as practice locations, specialty,

[15] They restrict their sample to the physicians providing direct patient care more than 20 hours per week, excluding federal employees, foreign medical school graduates who are only temporarily licensed to practice in U.S., and specialists in fields where the primary focus is not direct patient care.

and medical school.[16] I use the AMA data set instead of the CTS data set to confirm the implication on quantities because the CTS data does not cover all of US. I can not use the AMA data to confirm the implication on prices because it does not have information on income.

I aggregate specialties into 4 groups for both data sets following the standard classification scheme - general practice, medical specialties, surgical specialties, and other specialties. The second column of Table 4 shows the average annual income for each specialty. Across specialties, doctors in surgical specialties make the highest income, and doctors in general practice make the least.

4.1 The evidence on equilibrium prices

This section confirms the prediction in Proposition 1, across doctors with different specialties. I show that urban wage premiums are negative across all specialties, and are even lower for high pay specialties. As in Section 3.1, I calculate urban wage premiums by running the individual level regression (8) for each specialty θ. The fourth column of Table 4 reports the urban wage premiums across specialties. The urban wage premiums are negative across all specialties. In addition, urban wage premiums decrease in skill across different specialties. The urban wage premium is lowest for surgeons at -7.1 percent. This not only confirms the prediction in Proposition 1, but also reconfirms the key evidence in this paper that urban wage premiums are negative for doctors.

This negative relationship between urban wage premium and income might be due to other factors that might systematically change with metropolitan area size and skill. As in the previous section, I run the Mincer regression for each specialty controlling for a standard set of variables and some doctor specific variables.[17] The fifth column of Table 4 reports the urban wage premiums from the Mincer regression. The negative relationship remains strong.

[16]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.

[17]I added as regressions, dummy variable for international medical graduates, dummy for board certification, in addition to the standard variables I controlled for in section 3.1.

4.2 The evidence on equilibrium quantities

This section confirms the prediction in Proposition 2, across doctors with different specialties. I show that high skill specialties are more concentrated in large cities. As in Section 3.2, I calculate the urban concentration rate for each specialty θ by running the metropolitan area level regression (9). The sixth column of Table 4 shows the urban concentration rate for each specialty. The results are a little mixed, but consistent with my theory in general. Medical specialties, which ranks second in average income, are the ones that are most concentrated in large cities. However, the results are consistent with my theory in that generalists are least concentrated in large cities. Firms in large cities substitute higher-skill specialists for lower-skill generalists.

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

5 Physician Quality

This section shows that better doctors are more concentrated in large cities, even within specialties. This not only confirms the quantity implication of my theory, but also reinforces the key result in this paper. Doctors in large cities are paid less even though they are better doctors. A recent health economics literature examines physician quality, but most of the literature has focused on the difference across states rather than across urban and rural areas (e.g. Baicker and Chandra (2004), Jencks et al. (2000), Fisher and Skinner (2001)). As a measure of doctor quality, I use the average Medical College Admission Test (MCAT) scores of the medical schools of 2001 entrants.[18]

[18] The average MCAT scores are obtained from US news & world report - best graduate schools. The results obtained in this section are robust with the other years' MCAT scores I ran - 2002 and 2003.

I show that doctors from better medical schools are relatively more concentrated in large cities.

Using the AMA data set, I calculate urban concentration rate for each medical school θ by running the metropolitan area regression (9). Figure 8 shows the relationship between urban concentration rates and average MCAT scores of medical schools. There exists a clear positive relationship, which means that doctors from better medical schools are more concentrated in large cities.

This upward trend may arise for other reasons. First, top medical schools tend to produce relatively more specialists than generalists, and specialists are needed more in large 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 8 shows the geographic distribution of doctors from University of Illinois medical schools. They are heavily concentrated in Illinois, where their medical school are located. I resolve this issue by focusing only on migrant doctors who practice more than 500 miles away from their medical schools. For example, I exclude the observations inside the 500 mile circle in Figure 8 when I calculate the urban concentration rate for University of Illinois medical school graduates. Figure 8 shows the urban concentration rates across medical schools for each specialty after controlling for doctors' geographic concentration near their medical schools. The positive pattern remains strong for each specialty.

6 Urban Wage Premium Decomposition

There exists an urban wage premium for the health care sector as a whole. The average annual income of health care workers living in top third metropolitan areas is 17% higher than that of those living in the bottom third metropolitan areas ($45,986 vs $39,293) according to the Census 2000 data. This section provides a decomposition of this urban wage premium into three terms -

wage variations within occupations, quantity variations between occupations, and covariance. The between occupation term indicates the size of ability sorting effect, and it turns out that ability sorting accounts for 63 percent of the urban wage premium in the health care sector.

The average income difference $\bar{W}^T - \bar{W}^B$ between the top third metropolitan areas and the bottom third can be expressed as

$$\bar{W}^T - \bar{W}^B = \sum_\theta s_\theta^B (\bar{W}_\theta^T - \bar{W}_\theta^B) + \sum_\theta \left(s_\theta^T - s_\theta^B\right) \bar{W}_\theta^B + \sum_\theta \left(s_\theta^T - s_\theta^B\right) \left(\bar{W}_\theta^T - \bar{W}_\theta^B\right). \tag{10}$$

where $\bar{W}_\theta^i$ is the average income of occupation θ in area i ($i = T$ for the top third metropolitan areas, and $i = B$ for the bottom third), and s_θ^i is the quantity share of occupation θ in area i, i.e. $s_\theta^i \equiv l_\theta^i / \sum_\theta l_\theta^i$ (l_θ^i is the number of workers in occupation θ in area i.)

There are three terms in the decomposition (10). The first term is the contribution of wage variation within occupations. The second term is the contribution of quantity variation between occupations. Large cities have relatively more high skill workers, and this drives up the wages in large cities. This second term captures the ability sorting effect across occupations. The third term is the covariance term, which shows how the changes in wages between large cities and small cities are correlated with the changes in quantity shares. My theory predicts that this covariance term should be negative because high skill workers are concentrated more in large cities but paid less there, and vice versa for low skill workers. The following table summarizes the decomposition results in absolute terms and percentage terms.

$\bar{W}^T - \bar{W}^B$	$\sum_\theta s_\theta^B (\bar{W}_\theta^T - \bar{W}_\theta^B)$	$\sum_\theta \left(s_\theta^T - s_\theta^B\right) \bar{W}_\theta^B$	$\sum_\theta \left(s_\theta^T - s_\theta^B\right) \left(\bar{W}_\theta^T - \bar{W}_\theta^B\right)$
\$6,700	\$3,120	\$4,250	-\$670
100%	47%	63%	-10%

The decomposition result shows that ability sorting effect - quantity variations across occupations - accounts for 63% of the urban wage premium. The covariance term is negative, and this confirms the theoretical prediction that wage changes between large and small cities are negatively correlated with share (quantity) changes.

7 Conclusion

This paper provides a new consumption-side explanation for the urban wage premium. The explanation is based on selection of high-skill workers into cities. The demand for consumption variety is more income elastic than the demand for residential land. That is, as you become richer and richer, you spend higher fraction of your income on the differentiated local goods. Therefore, high skill workers care more about consumption variety level, and select to live in large cities. A testable implication of the theory, distinguished from other theories, is that there may exist an urban wage discounts for high-skill workers. This implication is confirmed by the health care sector data. High-skill health care professions, like doctors, in large cities are paid less than their peers in small cities. Returning back to the urban wage premium I started with, I decompose the urban wage premium for the whole health care sector into wage variations within occupations, quantity variations between occupations, and covariance term. I find that ability sorting effect, captured by the quantity variations between occupations, accounts for 63 percent of the urban wage premium for the whole health care sector.

8 References

Baicker, Katherine, and Amitabh Chandra. 2004. "Medicare Spending, The Physician Workforce, And Beneficiaries' Quality Of Care," *Health Affairs,* April 7.

Baumgardner, James R. 1988. "The Division of Labor, Local Markets, and Worker Organization," *Journal of Political Economy* 96, no. 3:509-527.

Berry, Steven, and Waldfogel, Joel. 2003. "Product Quality and Market Size," Ph.D. diss., National Bureau of Economic Research, Inc., NBER Working Papers.

Bils, Mark, and Peter J. Klenow. 2001. "Quantifying Quality Growth," *American Economic Review* 91, no. 4:1006-1030.

Center for Studying Health System Change. COMMUNITY TRACKING STUDY PHYSICIAN SURVEY, 2000-2001: [UNITED STATES] [Computer file]. ICPSR version. Washington, DC: Center for Studying Health System Change [producer], 2003. Ann Arbor, MI: Inter-university Consortium for Political and Social Research, 2003.

Costa, Dora L., and Matthew E. Kahn. 2000. "Power Couples: Changes in the Locational Choice of the College Educated," 1940-1990. *Quarterly Journal of Economics* 115, no. 4:1287-1315.

Fisher, Elliot and Jonathan Skinner, "Comparing the Health Care of States," *Providence Journal-Bulletin*, 17 March 2001.

Fujita, Masahisa, Paul Krugman, and Anthony J. Venables (1999), *The Spatial Economy: Cities, Regions, and International Trade*, Cambridge, Mass.: MIT Press.

Fujita, M., and Thisse, J.F. (2002) *Economics of Agglomeration: Cities, Industrial Location and Regional Growth.* Cambridge: Cambridge University Press.

Giles Duranton and Diego Puga, 2003. "Micro-Foundations of Urban Agglomeration Economies," NBER Working Papers 9931, National Bureau of Economic Research, Inc.

Glaeser, Edward L., Kahn, Matthew E., and Rappaport, Jordan. 2000. "Why Do the Poor Live in Cities?" Ph.D. diss., National Bureau of Economic Research, Inc., NBER Working Papers.

Glaeser, Edward L., Jed Kolko, and Albert Saiz. 2001. "Consumer City," *Journal of Economic Geography* 1, no. 1:27-50.

Glaeser, Edward L., and David C. Mare. 2001. "Cities and Skills." *Journal of Labor Economics* 19, no. 2:316-342.

Henderson, J V, 1974. "Optimum City Size: The External Diseconomy Question," *Journal of Political Economy*, University of Chicago Press, vol. 82(2), pages 373-88, Part I, M.

Jencks, Stephen, Barbara Fleming, and Nilasena David, "Quality of Medical Care Delivered to Medicare Beneficiaries," *Journal of the American Medical Association* 284, no. 13 (2000): 1670–1676.

Krugman, Paul. 1991. "Increasing Returns and Economic Geography," *Journal of Political Economy* 99, no. 3:483-499.

Lucas, Robert E.,Jr. 1988. "On the Mechanics of Economic Development," *Journal of Monetary Economics* 22, no. 1:3-42.

Rauch, James E. 1993. "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities," *Journal of Urban Economics* 34, no. 3:380-400.

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life," *Journal of Political Economy* 90, no. 6:1257-1278.

Tabuchi, Takatoshi, and Atsushi Yoshida. 2000. "Separating Urban Agglomeration Economies in Consumption and Production," *Journal of Urban Economics* 48, no. 1:70-84.

Figure 1. The Relationship Between The Urban Wage Premium And Skill Across Medical Service Occupations

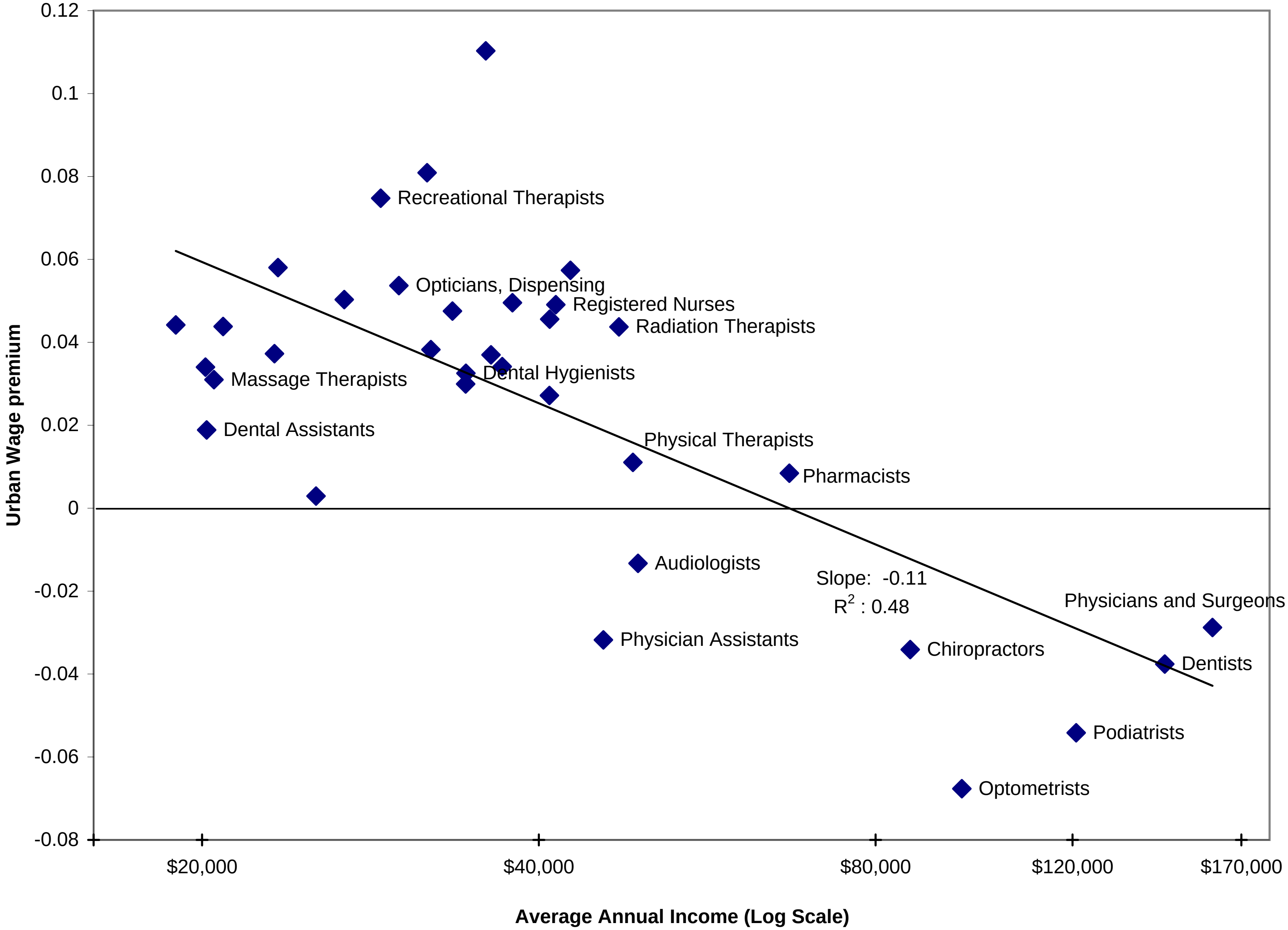

Figure 2. The Relationship Between The Urban Wage Premium And Skill Across Medical Service Occupations From The Mincer Regressions

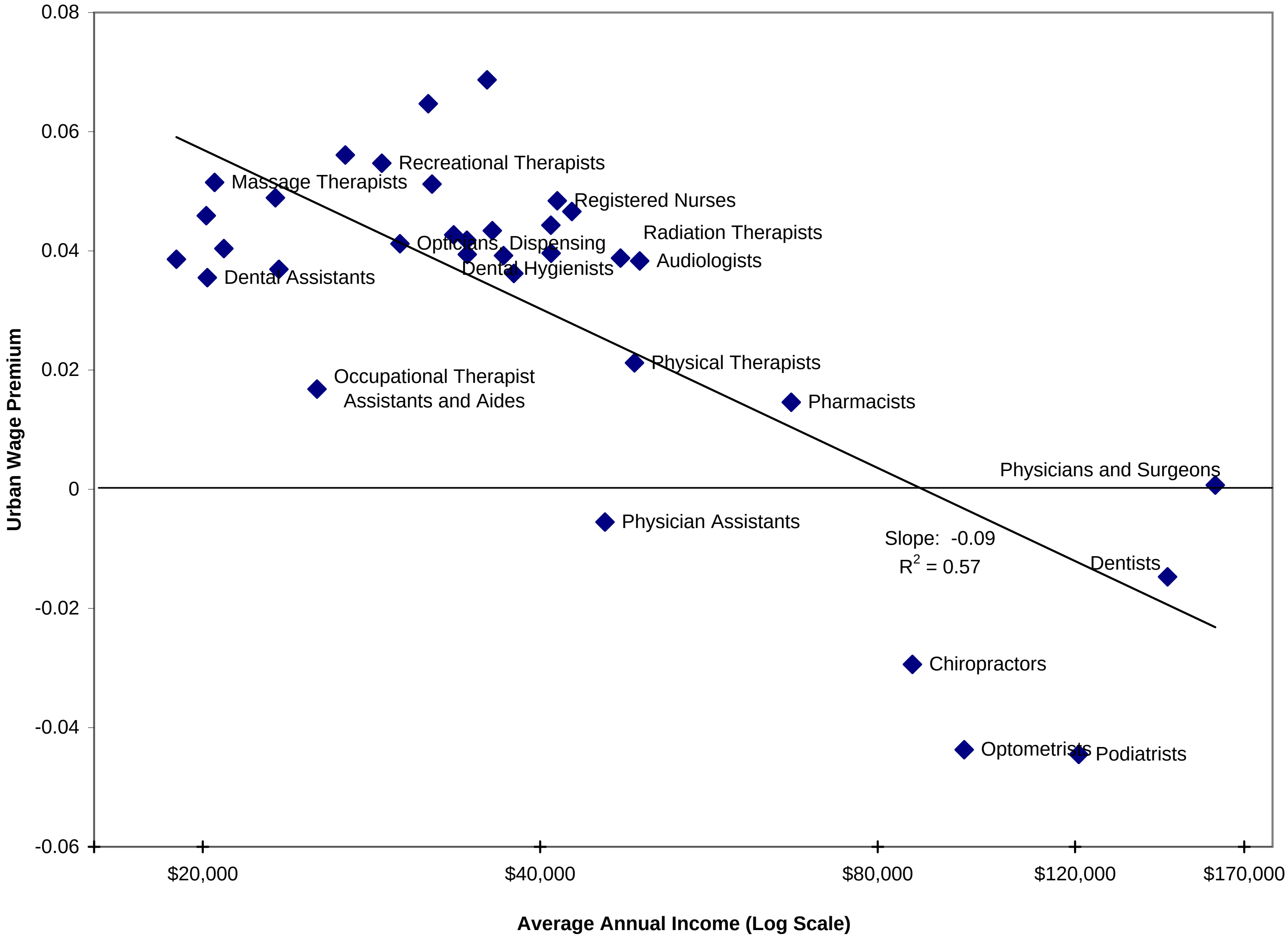

Figure 3. The Relationship Between Urban Concentration Rate And Skill Across Medical Service Occupations

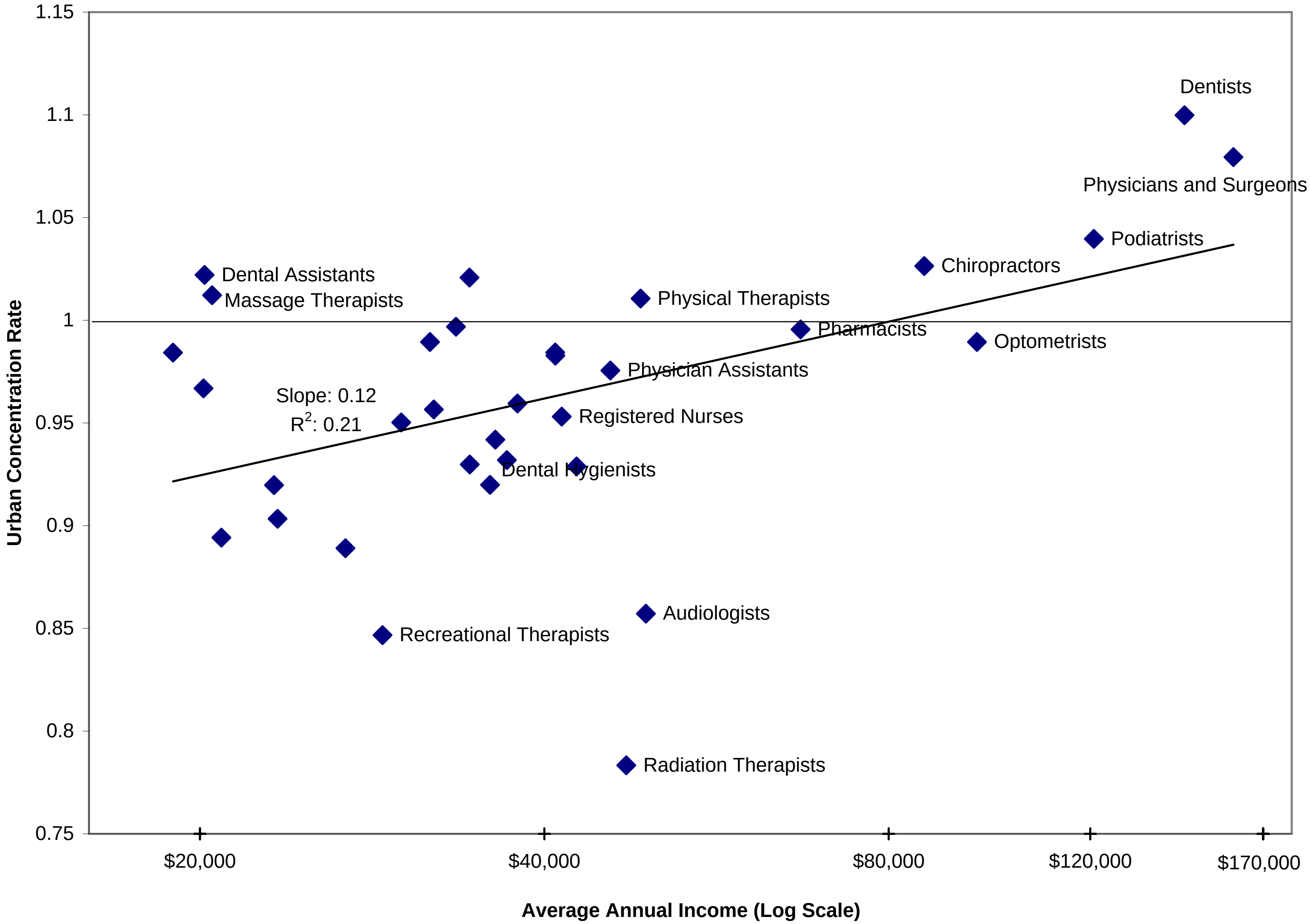

Figure 4. Urban Concentration Rate And Average MCAT Score Across Medical Schools

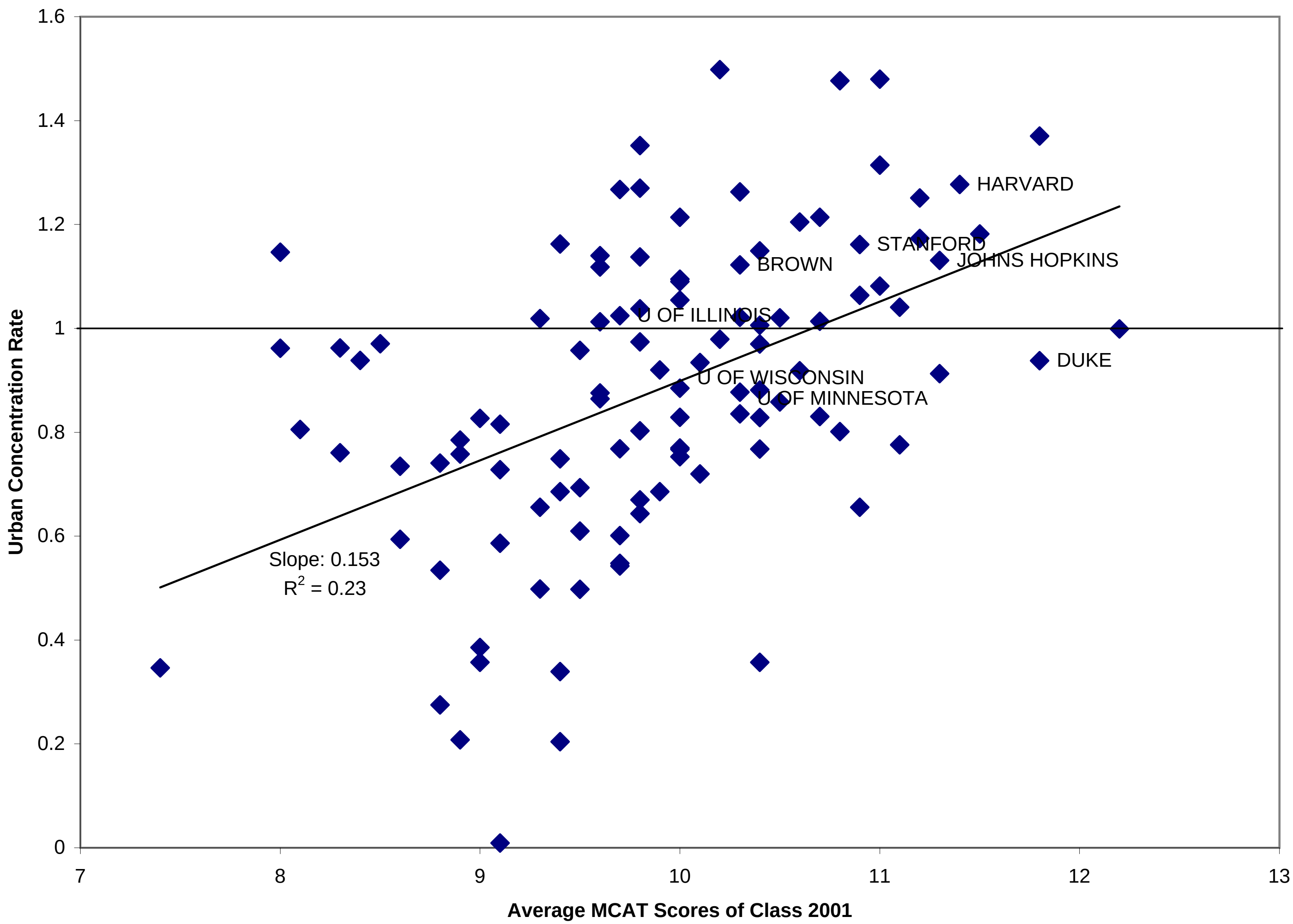

Figure 5. Geographic Distribution Of Doctors
University Of Illinois Medical School Graduates

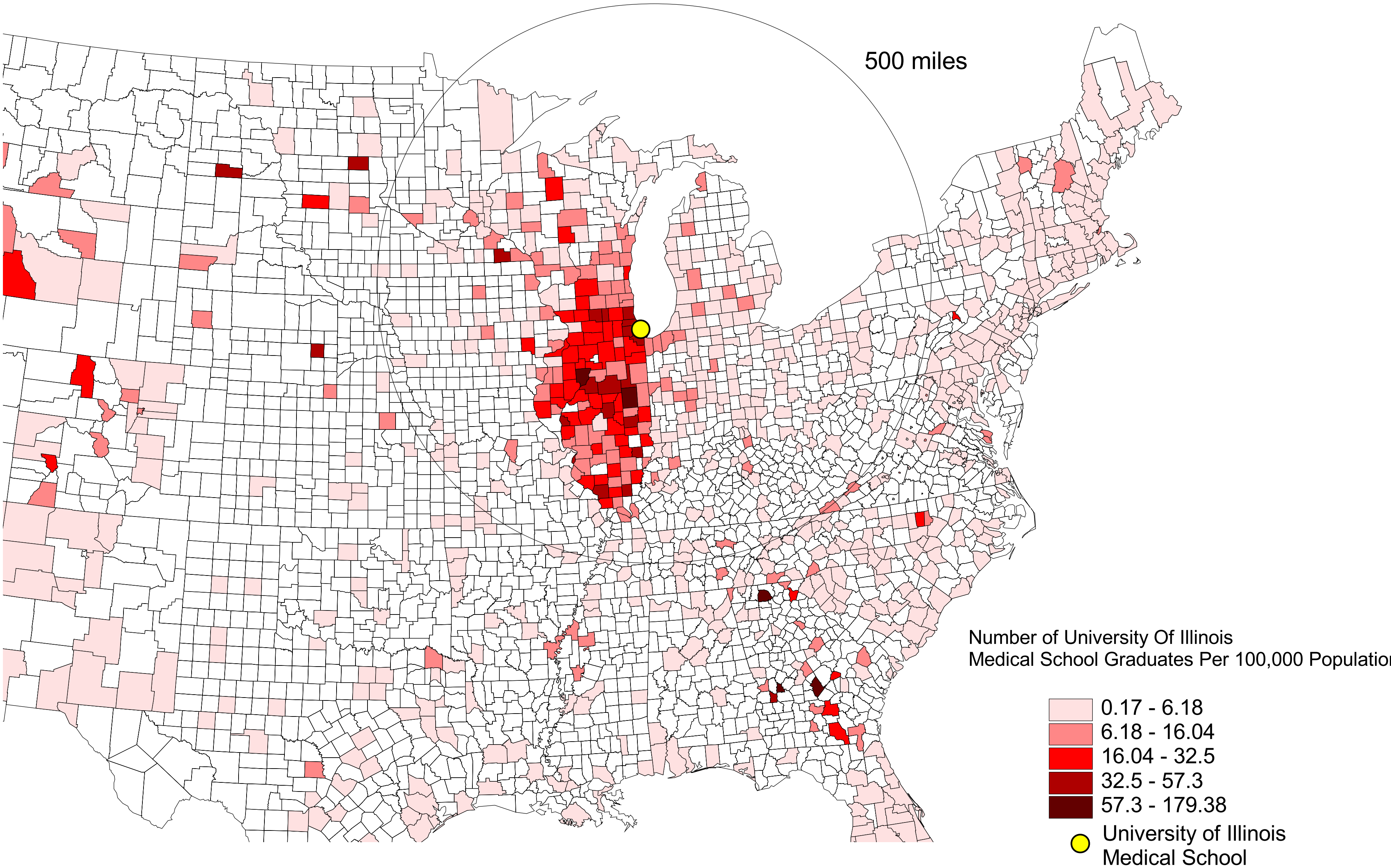

Figure 6. Urban Concentration Rate And Average MCAT Score By Specialty, For Doctors Practicing More Than 500 Miles Away From Their Medical Schools

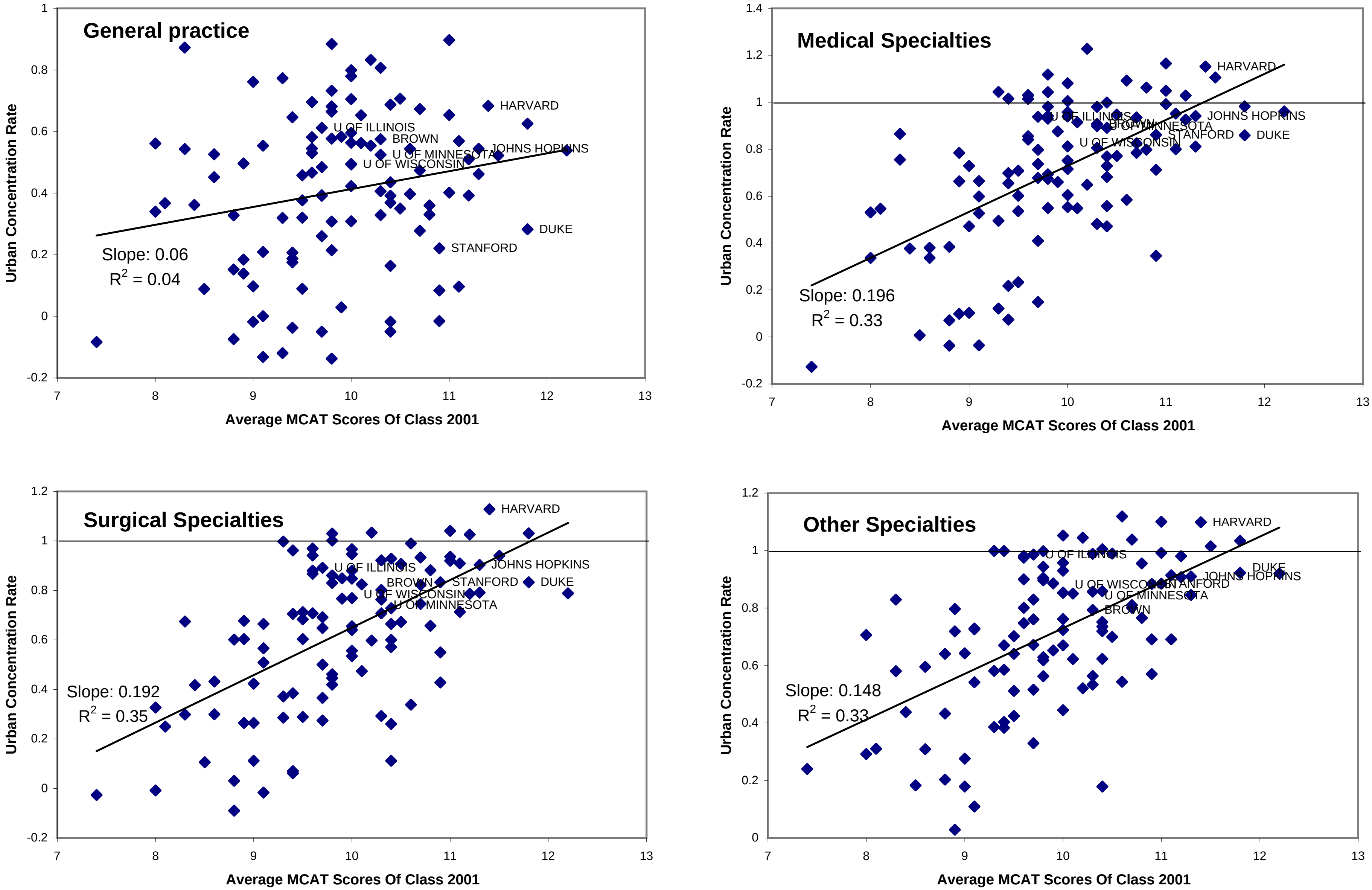

Table 1. Average Annual Income
By City Size, For Selected Medical Professions
(In Thousand of Dollars)

Occupation	Rural Areas	Bottom third metropolitan areas	Middle third metropolitan areas	Top third metropolitan areas	The ratio between top third and bottom third	The ratio between top third and rural areas
Physicians and Surgeons	157	158	151	145	0.92	0.92
Dentists	135	145	139	132	0.91	0.98
Podiatrists	113	125	110	110	0.88	0.97
Dental Assistants	16	17	20	18	1.06	1.13
Medical Assistants and Other Healthcare Support Occupations	15	17	19	20	1.16	1.32
Nursing, Psychiatric, and Home Health Aides	14	16	17	20	1.22	1.44

Source: Census 2000 5 percent Public Use Micro Sample

Table 2. Urban Wage Premiums and Urban Concentration Rates For Selected Medical Professions

Occupation	Urban Wage Premium	Urban Wage Premium (Mincer Regression)	Urban Concentration Rate
Physicians and Surgeons	-2.9% (0.4%)	0.07% (0.35%)	1.08 (0.01)
Dentists	-3.8% (0.9%)	-1.5% (0.8%)	1.10 (0.02)
Podiatrists	-5.4% (3.3%)	-4.5% (3%)	1.04 (0.05)
Dental Assistants	1.9% (0.6%)	3.6% (0.4%)	1.02 (0.01)
Medical Assistants and Other Healthcare Support Occupations	3.4% (0.4%)	4.6% (0.3%)	0.97 (0.01)
Nursing, Psychiatric, and Home Health Aides	4.4% (0.2%)	3.9% (0.2%)	0.98 (0.02)

Source: Census 2000 5 percent Public Use Micro Sample

Table 3. Number of Workers by City Size
For Selected Medical Professions
(Per 100,000 population)

Occupation	Rural Areas	Bottom third metropolitan areas	Middle third metropolitan areas	Top third metropolitan areas	The ratio between top third and bottom third	The ratio between top third and rural areas
Physicians and Surgeons	147	292	307	366	1.25	2.49
Dentists	43	59	61	81	1.36	1.89
Podiatrists	2	4	5	6	1.65	3.24
Dental Assistants	87	94	98	106	1.13	1.22
Medical Assistants and Other Healthcare Support Occupations	238	283	280	247	0.87	1.04
Nursing, Psychiatric, and Home Health Aides	1087	854	689	839	0.98	0.77

Source: Census 2000 5 percent Public Use Micro Sample

Table 4. 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 Specialties	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	131	2,438	-3.9%	(1.3%)	-1.9%	(1.2%)	0.90	(0.02)

Source: Community Tracking Study Physician Survey 2000-2001

Table A.1. Average Annual Income By City Size (In Thousand Dollars) and Urban Wage Premium For All Health Care Occupations

Occupation	Average Income	Rural Areas	Bottom third metropolitan areas	Middle third metropolitan areas	Top third metropolitan areas	Urban Wage Premium
Physicians and Surgeons	151	157	158	151	145	-2.9% (0.4%)
Dentists	138	135	145	139	132	-3.8% (0.9%)
Podiatrists	114	113	125	110	110	-5.4% (3.4%)
Optometrists	91	100	102	88	79	-6.8% (1.8%)
Chiropractors	82	81	92	82	74	-3.4% (1.8%)
Pharmacists	63	67	62	61	64	0.8% (0.6%)
Audiologists	47	36	45	51	48	-1.3% (2.5%)
Physical Therapists	45	43	45	43	47	1.1% (0.8%)
Radiation Therapists	44	46	46	40	45	4.4% (2.0%)
Physician Assistants	42	42	42	41	42	-3.2% (1.4%)
Other Healthcare Practitioners and Technical Occupations	40	36	37	41	46	5.7% (1.1%)
Speech-Language Pathologists	39	36	37	39	42	4.9% (0.2%)
Registered Nurses	39	33	37	38	44	4.6% (1.1%)
Occupational Therapists	38	36	38	37	40	2.7% (0.8%)
Respiratory Therapists	35	30	34	35	40	5.0% (1.1%)
Miscellaneous Health Technologists and Technicians	35	27	33	36	41	3.4% (0.7%)
Health Diagnosing and Treating Practitioners, All Other	35	25	27	31	43	3.7% (0.5%)
Diagnostic Related Technologists and Technicians	34	30	33	35	38	11.0% (4.5%)

Table A.1 Continued

Occupation	Average Income	Rural Areas	Bottom third metropolitan areas	Middle third metropolitan areas	Top third metropolitan areas	Urban Wage Premium
Therapists, All Other	32	29	30	33	35	3.3% (0.8%)
Dental Hygienists	32	30	30	33	35	3.0% (1.2%)
Clinical Laboratory Technologists and Technicians	31	28	29	30	34	4.8% (0.5%)
Emergency Medical Technicians and Paramedics	29	26	28	31	33	3.8% (0.9%)
Dietitians and Nutritionists	29	24	26	28	34	8.1% (1.1%)
Opticians, Dispensing	28	22	28	26	33	5.4% (1.2%)
Recreational Therapists	26	24	25	27	29	7.5% (2.0%)
Licensed Practical and Licensed Vocational Nurses	25	21	24	25	29	5.0% (0.3%)
Occupational Therapist Assistants and Aides	24	21	26	21	26	0.3% (2.2%)
Medical Records and Health Information Technicians	22	18	20	23	25	5.8% (0.9%)
Physical Therapist Assistants and Aides	21	20	19	23	21	3.7% (1.3%)
Massage Therapists	20	16	19	19	23	4.4% (0.5%)
Health Diagnosing and Treating Practitioner Support Technicians	19	17	19	19	23	3.1% (1.5%)
Dental Assistants	18	16	17	19	18	1.9% (0.6%)
Medical Assistants and Other Healthcare Support Occupations	18	15	17	19	20	3.4% (0.4%)
Nursing, Psychiatric, and Home Health Aides	17	14	16	17	20	4.4% (0.2%)

Source: Census 2000 5 percent Public Use Micro Sample

Table A.2. Number of Workers (Per 100,000) By City Size and Urban Concentration Rate For All Health Care Occupations

Occupation	Rural Areas	Bottom third metropolitan areas	Middle third metropolitan areas	Top third metropolitan areas	Urban Concentration Rate
Physicians and Surgeons	147	292	307	366	1.08 (0.01)
Dentists	43	59	61	81	1.10 (0.02)
Podiatrists	2	4	5	6	1.04 (0.05)
Optometrists	11	11	12	13	0.99 (0.02)
Chiropractors	16	19	21	21	1.03 (0.03)
Pharmacists	66	90	88	82	1.00 (0.02)
Audiologists	2	6	5	5	0.86 (0.04)
Physical Therapists	39	57	56	58	1.01 (0.02)
Radiation Therapists	3	5	4	4	0.78 (0.04)
Physician Assistants	19	27	27	23	0.98 (0.02)
Other Healthcare Practitioners and Technical Occupations	26	31	29	25	0.93 (0.02)
Speech-Language Pathologists	31	43	36	39	0.95 (0.01)
Registered Nurses	834	1027	939	934	0.98 (0.03)
Occupational Therapists	17	28	29	28	0.98 (0.02)
Respiratory Therapists	30	40	36	28	0.96 (0.02)
Miscellaneous Health Technologists and Technicians	28	43	38	34	0.93 (0.02)
Health Diagnosing and Treating Practitioners, All Other	2	3	4	7	0.94 (0.01)
Diagnostic Related Technologists and Technicians	81	104	94	81	0.92 (0.08)

Table A.2 Continued

Occupation	Rural Areas	Bottom third metropolitan areas	Middle third metropolitan areas	Top third metropolitan areas	Urban Concentration Rate
Therapists, All Other	25	29	31	33	0.93 (0.02)
Dental Hygienists	35	53	51	40	1.02 (0.02)
Clinical Laboratory Technologists and Technicians	94	142	132	134	1.00 (0.01)
Emergency Medical Technicians and Paramedics	66	48	40	40	0.96 (0.02)
Dietitians and Nutritionists	31	40	32	35	0.99 (0.02)
Opticians, Dispensing	18	23	24	20	0.95 (0.02)
Recreational Therapists	6	8	6	8	0.85 (0.05)
Licensed Practical and Licensed Vocational Nurses	363	286	223	191	0.89 (0.02)
Occupational Therapist Assistants and Aides	5	6	4	3	0.55 (0.06)
Medical Records and Health Information Technicians	42	52	47	35	0.90 (0.02)
Physical Therapist Assistants and Aides	24	24	21	20	0.92 (0.02)
Massage Therapists	22	28	37	30	0.89 (0.01)
Health Diagnosing and Treating Practitioner Support Technicians	119	152	134	107	1.01 (0.03)
Dental Assistants	87	94	98	106	1.02 (0.01)
Medical Assistants and Other Healthcare Support Occupations	238	283	280	247	0.97 (0.01)
Nursing, Psychiatric, and Home Health Aides	1087	854	689	839	0.98 (0.02)

Source: Census 2000 5 percent Public Use Micro Sample