E C O N O M I C S   B U L L E T I N

City-Size and Health Outcomes: Lessons from the USA

Achintya Ray
Tennessee State University

Soumendra N. Ghosh
Tennessee State University

Abstract

In this paper, we compare health outcomes in cities of different sizes. Using 2001 National Health Interview Survey data for adult urban-US population, it is shown that individual health is better in bigger cities compared to small or medium sized ones. This result holds after controlling for potentially confounding variables including age, gender, education, marital status, smoking, income, asset-ownership, and race. Possible sources of selection bias are controlled using many model specifications and population sub-groupings. Although, stiff challenges for healthcare delivery exist for large cities, an aggressive urban health policy should also put strong emphasis on improving health in small and medium sized cities to reduce urban health disparities in the USA. Policy implications for other developed and developing countries are also hypothesized.

We are grateful to James E. Foster, Mark Cohen, William Collins, and an anonymous referee for many useful comments and suggestions on earlier drafts of the paper. Responsibilities for all remaining errors lie with the authors only.
Citation: Ray, Achintya and Soumendra N. Ghosh, (2007) "City-Size and Health Outcomes: Lessons from the USA." *Economics Bulletin,* Vol. 9, No. 5 pp. 1-7
Submitted: February 15, 2007. **Accepted:** February 16, 2007.
URL: http://economicsbulletin.vanderbilt.edu/2007/volume9/EB-07I10003A.pdf

I. Introduction

Individual health attainments in American cities of different sizes are studied in this paper. Understanding health gradient among different sections of the urban population may be crucial to partly explain the overall urban health inequality.[1] Different health outcomes are found between different regions. For example, it has been argued that urban regions have better health compared to rural areas (Blumenthal (2002)). Access to healthcare facilities and resources are argued to be core reasons for such a gradient (Rabinowitz and Paynter (2002)).[2]

We find that after controlling for physical risk factors, demographic, behavioral, income/asset ownership and access to healthcare, health attainments are systematically better in larger urban areas. This is true across a number of population sub-groups defined along income, education and racial lines. This consistency attests to the robustness of the results obtained. [3]

The results of this study are different in many ways. *Firstly*, the health outcomes in cities of different sizes are compared. This is distinct from more traditional discussions on rural-urban health differentials. *Secondly*, this paper models the impact of city sizes on individual health after controlling for demographic and economic confounders. This is distinct from considering average health outcome in different locations. *Thirdly*, this paper looks at a broad array of health outcomes (like diabetes, arthritis, heart attack etc.) that greatly contribute to disability, death, and higher health spending. Thus the results of this paper have broad policy implications compared to those that look at either extreme outcomes like death or a small number of conditions.

Differences in health outcomes are found in many contexts for different population groups. People who are richer or more educated or both tend to have better health compared to their poorer or lower educated counterparts (Pincus et al (1987 & 1994), Pritchett and Summers (1997), Strauss and Thomas (1998)). Given these studies, it seems that health attainment is an increasing function of income and/or education. Studies in racial health disparities often argue that whites have better health than blacks (For example, Cohen et al (2004)).

If health attainment is an increasing function of income then societies having high income inequality will also have high inequality in the health dimension (Contoyannis and Forster (1999)). Similarly, a society where educational attainments are widely distributed may also have high health inequality.

---

[1] A health gradient is said to exist when health varies systematically with respect to another variable. For example, health is usually an increasing function of income. Thus a health gradient exists income. Also, health is usually positively related to education. In this case a health gradient exists in education.

2 Health gradient is also important for understanding health disparities. For example, if health is an increasing function of income then larger income inequalities may be related to higher health disparities. The same holds true for education also (Ray, 2007.)

[3] Different population sub-groups are separately studied to rule out any bias coming from selection problems. A selection problem is said to exist if the distribution of the sample in not random. In other words, selection problem will be a major concern if people of the same description are non randomly “bunched” in the same place.

In a related context, it has been argued that people tend to have better health if they live in more equal societies (Wilkinson (1996), Lynch et al (1998), Kennedy et al (1996 and 1998), Kaplan et al (1996)). For example, mortality is lower in societies where income is more equally distributed. In other words, these important studies indicate that income inequality may have hazardous health impacts. In this paper we present another possibility. Health disparities might persist if city size has an independent effect on individual health outcomes.

II. Data and Methods

Data for this paper came from two files of the year 2001 version of the National Health Interview Surveys' (henceforth, NHIS01) files on All persons and the file on the sample adults only. In the Original NHIS01 file distribution, these files were named PERSONX.DAT and SAMADULT.DAT respectively. These two simple ASCII files were formatted in STATA SE/8.0 software program for statistical analysis. Only Sample Adults were analyzed in this paper.

Wherever a variable of interest was not available in the sample adult file, it was collected from the file on all persons. The person and the sample adult files could be merged using the household and person identifiers. In this paper, we will restrict the sample to people falling in the age group of 18-65 only (boundaries included). This is to ensure that chronic health problems occurring especially among older people do not bias the results.

This paper looks at urban people living in Metropolitan Statistical Areas (MSA) only. We will use a dichotomous variable BIGCITY to represent MSA that have at least 1 million people living in them. The dichotomous variable BIGCITY will be the main variable of interest in this paper.

Six different measures of individual health are studied in this paper. Two are related to Self Rated Health Status (SRHS) and details of their construction will be described below. Other four are Arthritis (AR), Heart Attack (HA), Diabetes (DI) and Obesity (OB). Note that obesity could be a health problem in itself and additionally, it could lead to other forms of diseases. For example, hypertension, a major reason for HA is partly caused by obesity.[4]

To record SRHS, people were asked to self-evaluate their overall health status and report it as one of the five categories: "Excellent", "Very Good", "Good", "Fair" and "Poor". SRHS has been found to be a very good indicator or future illness and mortality (Idler and Benyamini (1997)). Two dichotomized variables HLTPOV and HLTRCH are constructed using the SRHS data. HLTPOV takes a value 1 if an individual reports "fair" or "poor" health, 0 otherwise. HLTRCH takes value takes a value 1 if an individual reports "excellent" or "very good" health, 0 otherwise.

AR is a dichotomized variable that takes a value if an individual reported to have diagnosed with Arthritis, 0 otherwise. HA is a dichotomized variable that takes a value if an individual reported to have diagnosed with Heart Attack, 0 otherwise. DI is a dichotomized variable that takes a value if an individual reported to have diagnosed with Diabetes, 0 otherwise.

The details about the construction of control variables are provided in the appendix.

---

[4] We look at many measures of health to make sure that our results are not driven by any specific measure of health.

To minimize bias due to missing data, we deleted observation with unknown BMI, unknown marital status, refused to give or unknown SRHS, unknown home tenure status, unknown health insurance coverage status, unknown educational achievement or people who refused to report their grade attainment, and unknown family income.

Our final sample had 20,180 observations. After using NHIS01 recommended observation relevant weights, 20,180 observations, represented over 124 million adults in the age group of 18-65 in the US population in 2001.

The likelihood of having any health state or condition is analyzed using logistic regression. Results report the odds ratios from the logistic regressions. The logistic regressions with AR, HA and DI as dependent variables are all controlled for obesity since obesity is known to be a risk factor for these conditions.

To control for the selection bias discussed above, logistic regressions are done for five population groups: all adults, adults living in families having income $75,000 or more, adults with at least a high school or equivalent education, and adults who have at least or equivalent education and living in families with income of $75,000 or more.

III. Results and Discussion

Table 1 presents the prevalence rates for the variables used in this paper. Note that larger cities have appreciably lower prevalence for all the bad health conditions explored in this paper. Smaller cities perform worse on all health measures.

It is striking that irrespective of the city size, there is a high fraction of people who can be considered as above “healthy” weight as characterized by high BMI. But smaller cities also have comparatively more poor, more people who smoke, more people who are at the upper parts of the age distribution, and larger mass without health insurance.

Logistic regressions for individual health are presented in Table 2. We present results for all adults and also for different population sub-groups.

Controlling for a large number of variables that can considerably affect health, city size still seem to play a very crucial role in determining one’s health status. Individuals are 10-15% less likely to suffer from adverse health outcomes like AR, DI, HA, HLTPOV or OB if they live in a larger urban areas. People are also 11% more likely to report excellent or very good health if they live in larger urban areas.

It seems that people who are rich, well educated or both all have a better likelihood of being more healthy if they live in larger cities.

It is seen from Table 2 that people who are quite rich (i.e. living in families with income $75,000 or more) are 13-30% more likely to report better health if they live in larger cities. Also, well educated people (i.e. having at least high school diploma or equivalent or more education) living in larger urban clusters are 11-15% more likely to have better health compared to those in the same educational group but living in smaller cities.

Table 2 also shows that people who are very rich and well educated (i.e. having at least a high school diploma or equivalent & living in families with family income of $75,000 or more) are 12-30% more likely to report better health if they live in larger urban areas.

Results presented in Table 2 strongly suggests that the size of the city in which an individual lives may have a non-trivial impact on one's health. Note that all the models are controlled for age, gender, race, education, marital status, smoking, enrollment in health insurance, asset (house) ownership, income etc.; all relevant variables that immediately come to mind as potentially explanatory for individual health outcomes.

IV. Conclusion and Policy Directions

The basic policy related message that should come out of this paper is that urban health policy should take into account city size as an important explanatory variable impacting individual health. It may be hypothesized that a large component of urban health disparity could be reduced if aggressive policies target people especially living in smaller urban areas and try to improve their health.

Since, the results of this paper are entirely based on American data, the general applicability of the results in a global context remains an interesting issue. Additional research is necessary to ascertain if similar gradients exist across differentially sized cities in other nations. It may be possible that American cities are different from similar cities in other nations. For example, mega-cities in Latin America (like Sao Paulo or Rio) or Asia (like Mumbai, Kolkata, New Delhi etc.) often contain widespread urban slums offering very difficult and unhygienic living conditions. Such urban sprawls are usually absent in American cities.

APPENDIX: Construction of Control Variables

OB is a dichotomized variable derived from the Body Mass Index (BMI) that is calculated according to the formula: BMI =[Weight (kg)/[Height (meter) squared]]. Obesity will take a value 1 if BMI $\geq$ 25. In the NHIS01 documentation, people with BMI > 25 are called overweight and people with BMI > 30 are called obese. Hence, the dichotomous variable created in this paper will include both overweight & obese people.

We looked at sets of variables that are known to have influence on individual health. They are demographic (age, gender, race, and marital status), behavioral (smoking), poverty/wealth/income (family income and house ownership), education (highest level of education completed) and access to health care (health insurance enrollment).

We constructed several dichotomous variables to capture these groups. AGE takes a value 1 if age>50, 0 otherwise. GEND is equal to 1 if the individual is female, 0 otherwise. RACE takes a value 1 if the individual is black, 0 otherwise. MARIT takes value 1 if the individual is currently married, 0 otherwise. SMOK takes a value 1 if the individual is currently a smoker, 0 otherwise. POOR is equal to 1 if the individual's family income is less than $20,000, 0 otherwise. HOWN takes a value 1 if the individual lives in a house that is owned or being bought by the individual or somebody in the family, 0 0therwise. HSD takes a value 1 if the person is at least a high school graduate, 0 otherwise. Lastly, NCOV takes a value 1 if the person is not currently covered by any health insurance. SRHS, NCOV, and HOWN are collected from the file reporting data on all persons. Rest of the variables can be constructed using data available in the sample adult file.

References

1. **Blumenthal, S. J.**, (2002), “The Effects of Socioeconomic Status on Health in Rural and Urban America”, Journal of American Medical Association, 287(1), 109.
2. **Cohen, C. I., Magai, C., Yaffee, R., and Walcott-Brown, L.**, (2004), “Racial differences in paranoid ideation and psychoses in an older urban population.”, American Journal of Psychiatry, 161(5), 864-71.
3. **Contoyannis, P. and Forster, M.**, (1999), "The distribution of health and income: a theoretical framework", Journal of Health Economics, 8(5), 603-620.
4. **Idler, E.L. and Benyamini, Y.**, (1997), "Self-rated Health and Mortality: A Review of Twenty Seven Community Studies", Journal of Health and Social Behavior, (38), 21-37.
5. **Kaplan, G., A., Pamuk, E., R., Lynch, J., W., Cohen, J., D., and Belfour, J., L.**, (1996), "Inequality in income and mortality in the United States: Analysis of mortality and potential pathways", British Medical Journal, 312, 999-1003.
6. **Kennedy, B. P., Kawachi, I. and Prothrow-Stith, D.**, (1996), "Income distribution and mortality: cross-sectional ecological study of the Robin Hood index in the United States", British Medical Journal, 312, 1004-1007.
7. **Kennedy, B. P., Kawachi, I., Glass, R. and Prothrow-Stith, D.**, (1998), "Income distribution, socioeconomic status, and self rated health in the United States: multilevel analysis.", British Medical Journal, 317, 917-921.
8. **Lynch, J., W., Kaplan, G., A., Pamuk, E., R., Cohen, R., D., Belfour, J., L., and Yen, I., H.**, (1998), "Income inequality and mortality in metropolitan areas of the United States.", American Journal of Public Health, 88, 1074-1080.
9. **Pincus, T., Callahan, L.F., and Burkhauser, R.V.**, (1987), “Most chronic diseases are reported more frequently by individuals with fewer than 12 years of formal education in the age 18-64 United States population.”, Journal of Chronic Diseases, 40(9), 865-74.
10. **Pincus, T. and Callahan, L.F.**, (1994), “Associations of low formal education level and poor health status: behavioral, in addition to demographic and medical, explanations?”, Journal or Clinical Epidemiology, 47(4), 355-61.
11. **Pritchett, L. and Summers, L. H.**, (1997), "Wealthier is Healthier." Journal of Human Resources 31(4), 841-868.
12. **Rabinowitz, H.K., and Paynter, N.P.**, (2002), “The Rural Vs. Urban Practice Decision”, Journal of American Medical Association, 287(1), 113.
13. **Ray, A.,** (2007), “Are Racial Disparities in Diabetes in the USA Driven by Education Distribution?”, Economics Bulletin, Vol. 9 no. 2 pp. 1-18.
14. **Strauss, J. and Thomas, D.,** (1998), "Health, Nutrition and Economic Development", Journal of Economic Literature, (36), 766-817.
15. **Wilkinson R.G.**, (1996), "Unhealthy Societies: The Afflictions of Inequality", Routledge, London.

Table 1
Prevalence Rates, National and across city sizes. All adults (Age 18-65).
Figures in percentages.

Field	National	BIGCITY=1	BIGCITY=0
HLTPOV	8.58	8.13	9.18
HA	1.69	1.53	1.90
AR	15.61	14.46	17.16
DI	4.49	4.15	4.95
OB	57.44	56.05	59.33
HLTRCH	69.71	70.75	68.31
AGE	21.88	21.49	22.40
GEND	50.03	50.60	49.24
RACE	12.91	15.09	9.96
MARIT	57.63	56.69	58.90
HSD	99.38	99.27	99.52
POOR	15.64	14.28	17.49
NCOV	16.12	15.80	16.57
SMOK	24.51	22.78	26.86
HOWN	66.35	64.09	69.43

Table 2
Effect of City Size on Individual Health*
All figures are statistically significant (p<0.01)

	Dependent Variables					
	HLTPOV	HA	AR	DI	OB	HLTRCH
All Adults (18-65)						
BIGCITY	0.90	0.88	0.85	0.86	0.85	1.11
All adults (Age 18-65) living in families with family income of $75,000 or more						
BIGCITY	0.75	0.70	0.82	0.82	0.84	1.13
All adults (Age 18-65) having at least a high school diploma or equivalent						
BIGCITY	0.89	0.89	0.85	0.85	0.85	1.11
All adults (Age 18-65) having at least a high school diploma or equivalent & living in families with family income of $75,000 or more						
BIGCITY	0.77	0.70	0.83	0.82	0.84	1.12

* **All Results are controlled (Wherever applicable) for AGE, GEND, RACE, MARIT, HSD, POOR, NCOV, SMOK, and HOWN**