Institute for International Economic Policy Working Paper Series
Elliott School of International Affairs
The George Washington University

Amenity, Diversity and Obesity: Unobserved Heretogeneity in Cities

IIEP-WP-2015-12

Stephen Popick
Anthony Yezer
George Washington University

June 2015

Institute for International Economic Policy
1957 E St. NW, Suite 502
Voice: (202) 994-5320
Fax: (202) 994-5477
Email: **iiep@gwu.edu**
Web: **www.gwu.edu/~iiep**

AMENITY, DIVERSITY AND OBESITY: UNOBSERVED HERETOGENEITY IN CITIES

STEPHEN POPICK[#], AND ANTHONY M. YEZER[+]

June, 2015

ABSTRACT. Some sources of heterogeneity among cities, i.e. age, gender, race, income, and education, have been the object of substantial inquiry. The reasons are obvious. These differences are easily observed and may have important implications for economic activity. This study considers another potentially important population characteristic, obesity. Descriptive statistics reveal that the intercity variance in obesity rates is substantial. Empirical results demonstrate that demographic and regional amenity variables all have a relation to intercity differences in obesity. Because obesity is important for preferences, performance, and productivity, its omission from previous studies and its correlation with amenity and demographic characteristics, could create problems for empirical research.

JEL Codes: I12, J10, R23,

[#] Ph.D. Candidate, Department of Economics, George Washington University 505 A E Windsor Ave Alexandria VA 22301 (571) 224-5114 <spopick@gwu.edu>

[+] Corresponding Author, Professor, Department of Economics, George Washington University, 2100 G Street NW Washington DC 20006 (202) 994-6755 <yezer@gwu.edu>

The authors declare that they have no relevant or material financial interests that relate to the research described in this paper. None of the data are proprietary.

AMENITY, DIVERSITY AND OBESITY: UNOBSERVED HERETOGENEITY IN CITIES

I. INTRODUCTION

It is well known that obesity is negatively associated with income and education. [1] Age, ethnicity, race, and gender effects are more complex but certainly significant. Because cities differ in composition along income, education, age, race, ethnicity, and gender dimensions, they should naturally differ in obesity rates. However, a simple comparison of obesity rates across cities suggests that other factors may be influential. Approximately 42.7% of the population of the San Francisco-Oakland, CA MSA have a non-obese Body Mass Index (BMI < 25) while 20.7% are obese (BMI > 30). In contrast, the Detroit, MI MSA percentages are 29.2% non-obese and 35.6% obese. Could such differences in obesity rates be due to observable age, race, ethnicity, gender, education, and income characteristics of individuals in the population of these cities or are specific city characteristics differentially attractive to the obese?

The question addressed in this paper is whether, in addition to observable personal characteristics of their inhabitants, city amenities are associated with differences in obesity. The literature reviewed here suggests a number of reasons to believe that differences in climate, topography, and of course food prices, make cities differentially attractive to individuals with

1 Much of this knowledge is based on the questions in the National Longitudinal Survey of Youth. This sample is inadequate to test the hypotheses regarding intercity differences being examined here. It is quite adequate to demonstrate differences in BMI associated with personal characteristics. The other source of BMI data is the National Health Inventory Survey, which is annual, but only identifies 33 MSAs. This number is too small for the analysis of the effects of differences in city characteristics undertaken here.

high BMI. This differential attraction may then cause individuals to select into these cities through migration. To the extent that differences in BMI are hereditary and city amenities are permanent, past migration may also select those prone to obesity into certain cities. Thus, there may be a "BMI selection effect" of city characteristics. [2] Alternatively, there may be a "BMI adaptation effect" as individuals adjust their BMI to city characteristics. Regardless of the process, both selection and adaptation effects, if they are important, may contribute significantly to the large spatial differences in BMI documented in this paper.[3]

Why is the possibility that spatial characteristics play a role in determining BMI differences in the resident population important? First, general interest in obesity is high because of its strong connection with development of type 2 diabetes mellitus (DM) and other ailments that impose substantial costs on society. Recent estimates of these costs are as large as 1.1 trillion dollars per year for the U.S. economy. [4] Second, there is an important issue for empirical research in economics. BMI is generally unobservable to the econometrician. Nevertheless, BMI is an important determinant of individual preferences and worker productivity partly due to direct effects and indirectly through the connection between obesity and DM. Therefore, because BMI is correlated with variables such as income, education, and demographic

2 The BMI selection effect has created problems in the literature on the effects of "sprawl" on BMI. There is substantial disagreement in this literature. See, for example, Eid, Overman, Puga, and Turner (2008) versus Zhao and Kaestner (2010). This paper is not concerned with the distribution of BMI within cities.

3 It may be that individual expectations for "optimal" BMI are based on community standards and that behavioral economics could explain local variation in diet and exercise. Following Brennan's (2014) recent suggestion, the analysis can be viewed as testing a rational choice model.

4 Brookings Institution Study. See: http://www.brookings.edu/blogs/brookings-now/posts/2015/05/societal-costs-of-obesity

characteristics that are important in empirical research, this raises the possibility for omitted variable bias in estimates of the effects of these personal characteristics on measures of wage differentials for constant quality workers, and in a variety of other empirical research. Alternatively, if BMI is related to city amenity characteristics, this unobserved association could confound inferences about causes of spatial wage or productivity differences across cities.[5] Could it be that a significant portion of the wage differential between observationally similar workers in San Francisco and Detroit is due to unobserved differences in their BMIs?[6]

Recent availability of large scale individual survey data on BMI for a representative sample of city populations allows testing of the hypothesis that, holding income, education, and demographic characteristics constant, selected city characteristics have a significant relation to their obesity rates.[7] The object of this study is to test the hypothesis that the city characteristics that are expected, based on physiological effects of obesity, to make areas differentially attractive to those with high BMI, have an influence on the average body mass index and obesity rate in the city.

The next section of this paper develops the theoretical rational for believing that there is a BMI selection effect in which city characteristics have a differential attraction for obese

---

5 There is a substantial quality of life following Roback (1982) that relates wage differentials to city amenities under the assumption that amenities have no effect on obesity or other unobservable population characteristics.

6 Observational equivalence in this case refers to research that does not observe worker body mass index (BMI).

7 The Centers for Disease Control Behavioral Risk Factor Surveillance System surveys of individual BMI used in this study have been conducted for many years but, over the past 15 years the sample size increased significantly so that reliable estimates of BMI differences across a range of cities are possible.

individuals. Then the available literature that relates BMI to preferences for climate, topography, and other city characteristics is reviewed. The data section discusses the construction of variables designed to measure these city differences. Finally, empirical results show general agreement between prior expectations and the obesity rate of cities.

II. THEORY: BMI AND CHOICE OF LOCATION

Assume that there are multiple households differentiated by a single scalar characteristic, *B*, which is an "inherited" property of individuals. [8] They must choose a location among areas indexed by *j* that are differentiated by wages, w_j, transportable goods, *x*, whose price everywhere is p, non-transportable goods, *h*, for "housing" whose price, r_j, varies spatially, and local amenity whose implicit price, q_j, varies spatially. The indirect utility of a particular household , i, in location *j* can be written as:

$$U_{ij} = V(w_j, p, r_j, q_j; B_i) \qquad (1)$$

Taking the total differential of indirect utility under the assumption of constant utility across cities and solving for dw_j, gives:

$$d w_j = -\lambda_{xw} dp - \lambda_{hw} dr_j - \lambda_{qw} d q_j \qquad (2)$$

8 For purposes of this model, it does not matter whether B has a genetic origin or if it is learned in childhood.

where λ_{yz} is the ratio of the partial derivative *V(.)* with respect to *y* divided by the partial of *V(.)* with respect to *z*.[9] Applying Roy's identity, the relation in equation (2) can be solved for the total derivative of earnings:

$$d\,w_j = x_j\,dp + h_j\,d\,r_j - a_j\,d\,q_j \qquad (3)$$

which, assuming $dp = 0$, implies that:

$$d\,w_j = h_j\,d\,r_j - a_j\,d\,q_j$$

(4) Equation (4) states that the equilibrium tradeoff between wages and rents depends on the quantity of amenity consumed by the individual. It follows from the effect of *B* on indirect utility of the amenity that $da_j/dB > 0$ and high *B* households will require a smaller compensating differential in wages to live in areas where the price of the amenity in question is lower. Therefore, high *B* households are differentially attracted into areas with low q_j.

Obviously, the *B* factor relevant for this research is BMI and the hypothesis is that the relative concentration of high BMI individuals will rise in areas where the prices of amenity factors that are differentially attractive to obese are low. This spatial sorting of population by BMI could arise through migration and/or genetic selection. Ford (2005) has noted that migration is one possible sorting mechanism under the hypothesis that the tendency to be obese varies significantly in the population. Alternatively, Piziak (2010) and Andersson (2011) contend that heredity is an important determinant of BMI. This suggests that spatial differences in BMI could be the result of prior migration by those with a genetic predisposition to obesity. Finally, Chen

---

9 Note that (1) can be written in implicit form and provide a clear statement of the spatial iso-utility condition.

(2013) has observed that standards of diet and exercise could vary spatially based on the interaction of preferences, which vary with BMI, and the relative proportion of the obese in the population.

This paper does not test the exact mechanism that accomplishes the sorting, although estimation results indicate that the effects of climate and topography on BMI for individuals under 25 years of age are identical to those for individuals 25 or older. To the extent that migration occurs at ages greater than 24, this suggests that differences in BMI are not due to recent migration but rather to differences in the resident populations of areas that could be the result of past migration.[10]

III. EFFECTS OF INDIVIDUAL AND CITY AMENITY CHARACTERISTICS ON OBESITY

The four most prominent individual characteristics recognized in the economics literature as having a possible relation to BMI are income, education, gender, and age. The empirical evidence strongly suggests that obesity varies inversely with both income and education (Baum 2004). The underlying reasons for this relation are not clear and may be quite complex but the relation holds within as well as among cities. Females have lower BMI.[11] Age effects are non-linear because there is a tendency for BMI to be highest in middle age (Gallup 2012). To the extent that income, education, gender, and age are distributed unequally across cities, they may explain a significant portion of the variation in spatial obesity rates. In addition, race and

10 Limited sample size and the discrete categories of age did not permit testing of BMI effects for younger age cohorts.

11 CDC/NCHS, Health, United States, 2014, Table 64. Data from the National Health and Nutrition Examination Survey (NHANES)

ethnicity are unequally distributed across cities and may have an independent relation to BMI. The purpose of this research is not to sort out the causal relation between these factors and BMI but rather to test whether spatial differences in their distribution can explain the large differences in BMI and obesity across cities or if other factors involving population selection based on amenities analyzed in the theory section are important.

Much of the literature on amenity factors, whose attractiveness might vary with individual BMI, lies outside of economics because physiology is the basis for differences in preferences. Simply put, endomorphs react differently than ectomorphs to the same environmental conditions. Examination of the literature reveals a number of area characteristics that should relate to obesity because the preferences of the obese are observed to differ from the average and thin populations. These preferences are the result of physiological effects of obesity. A substantial literature, stemming from seminal work by Roback (1982), classifies these factors as local amenities. First are opportunities for outdoor recreation. City characteristics including access to water and parkland should be valued less by the obese. Secondly, the obese have difficulty dealing with certain topographic characteristics. Voss (2013) has found that elevation and elevation change are more physically demanding for the obese. Accordingly, individuals with high BMI will avoid mountainous locations and seek relatively flat coastal locations because of both topography and oxygen availability. BMI has a significant effect on preferences for climate. Cold winters are more uncomfortable for those with low BMI. Hot summers make outdoor recreation more difficult. Lin (2007) has argued that, given that exercise is associated with lower BMI, individuals who exercise try to avoid areas where summers are extremely hot. Dehghan et. al. (2013) found obese workers suffered more cardiac strain than their non-obese colleagues did in hot and humid conditions. Accordingly, those with low BMI have a relatively stronger preference for areas with mild winters and mild summers. Causality can go in either direction here. Cold winters and hot summers may make outdoor exercising difficult and the lack of exercise may lead to higher BMI. As noted above, for purposes of this study, the direction of causality is not material.

One limitation of the explicit measures of environmental amenity discussed above is that measures of parkland, water bodies, etc., are not adjusted for quality of the recreational experience that they provide. To the extent that the quality dimension of these local amenity variables is missing, there is measurement error that causes attenuation bias in estimates of the amenity effect.

Research on obesity has isolated a number of other non-amenity factors that tend to repel the obese and/or attract ectomorphs. Edwards (2008) identified the availability or use of public mass transit, Booth (2005) cited housing density, and Grossman (2013) among others pointed to the cost of food as factors that may relate to obesity rates in an area. These additional factors are added to the empirical analysis.

IV. DATA ON BMI AND URBAN AMENITY FACTORS

In relating cross section variation in body mass index (BMI) to the city characteristics identified above, availability of data has previously been a major constraint. The source of BMI data for this study is the 2010 Centers for Disease Control (CDC) Behavioral Risk Factor Surveillance System survey of BMI at the individual and 2010 CDC SMART data (derived from the 2010 CDC BRFSS) at the MSA level. [12] The CDC survey includes calculated BMI and self-reported income, education, age and various demographic variables.

The CDC survey overcomes the greatest challenge to research on the spatial distribution of BMI. This is a telephone survey covering all 50 states and including over 400,000 individual responses. The data includes a calculated BMI for each respondent (bmi) that is based on responses regarding height, weight, age, and gender, The CDC relates BMI to healthy weight status. Individuals with $bmi \geq 30$ are considered "obese", $25 \leq bmi < 30$ is classified as "overweight", and individuals with $18 \leq bmi < 25$ are "normal" or "healthy weight". Finally, "underweight" individuals have $bmi < 18$. One difficulty with the BMI computation is that it does not consider muscle mass. In addition to the endomorphs, there are mesomorphs who have rather high BMI. Given that mesomorphs may be at least as physically fit as ectomorphs, they likely have similar preferences. Accordingly, the inability to adjust BMI for muscle mass likely works against finding differences in the spatial distribution of BMI based on amenity factors.

Based on the 2010 CDC SMART data summaries by MSA, Honolulu has the highest positive difference between the percent of individuals who are normal weight (43.9%) and the percent who are obese (43.9% – 21.2% = 22.7%). On the other end of the spectrum McAllen has the most substantial negative difference between the percent of individuals who are normal weight (22.7%) and the percent who are obese (22.7% – 41.7% = –18.9%). This illustrates the

12 In its technical documentation, the CDC defines geographies using MMSAs (http://www.cdc.gov/brfss/smart/smart_data.htm). However, MMSAs and CBSAs are equivalent (http://www.census.gov/population/metro/) and this study focuses on metropolitan areas (MSAs). MSAs are defined using a state county to MSA crosswalk published by the US Census (http://www.census.gov/population/metro/files/lists/2009/List1.txt). Using the most recent crosswalk file from 2013 produced similar results. (http://www.census.gov/population/metro/data/def.html)

potential for spatial variation in BMI to have a substantial influence on differences in preferences and productivity of the population of cities.

Recalling the example noted at the start of this paper, San Francisco – Oakland MSA(SF) includes 42.7% with normal weight BMI and 20.7% who are obese for a BIM difference of $(42.7\% - 20.7\% = 22\%)$. Its average January temperature is 50 degrees while average July temperature is 58 degrees. In contrast, the Detroit MSA(DT) population ratios are 29.2% normal and 35.6% obese for a BMI difference of $(29.2\% - 35.6\% = -6.4\%)$. Detroit's average January and July temperatures are 25 and 74 degrees respectively. Recalling that cold winters repel ectomorphs due to their greater physiological response and hot summers make outdoor recreation more difficult, the difference between the BMI ratios of SF (mild winter and mild summer) and DT (cold winter and warm summer) appears to be consistent with expected climatic effects. Thus, differences in climate that are pure amenity effects may have dramatic implications for unobserved heterogeneity, BMI in this case, of city population.

The CDC BRFSS reports individual level information on household income, education, gender, and age. Respondents self-report age as a continuous measure but report household income and education by selecting the appropriate range or value from a categorical list. The statistical analysis uses these categories. Therefore, this paper models BMI as a step function of income and education along with continues variables for age and indicator variables for race/ethnicity and gender.

Table 1 provides descriptive statistics for income, education, age, race, ethnicity, gender, intra-MSA location, and BMI from the CDC BRFSS individual level micro data. As indicated by the descriptive statistics, 191,215 persons reside in one of the 110 MSAs with NOAA weather data (January and July temperatures) which is a primary amenity variable in this study. In this 110 MSA sample, 1% of individuals surveyed are under-weight, 36% are normal-weight, 36% are overweight, and 27% are obese.

Table 2 details descriptive statistics for the city characteristics (amenity and non-amenity) that are potentially related to BMI for individuals in the 110 MSAs where temperature data is available. Non-amenity factors include 2010 U.S. Census estimates of the total population (*Pop(100k))* expressed in hundreds of thousands, percent of the population that regularly uses mass transit, (*Transit(%))*, and the density of housing, (*Density)*. Data from the ACCRA food costs index, (*GroceryCOLA),* measures the cost of groceries.

Three topographic factors commonly used to reflect local amenity that relate to BMI, are considered. The percentage of total area MSA covered by water, (*Waterarea)*, is taken from the 2010 U.S. Census. A second topographic variable, (*Coastal),* is a binary indicator variable for MSA coastal location, (Atlantic, Pacific, Gulf of Mexico or Great Lakes) using NOAA coastal county definitions.[13] Coastal locations give more opportunity for outdoor activity but they are also generally flatter and near sea level. [14] Recreation activity attracts those with lower BMI but flat terrain and low altitude make mobility easier for the obese. The percent of the MSA composed of park and recreation space, (*Parkland(%))*, is available from the 2010 American Fitness Index, but only for a subset of cities. [15] Clearly, this local public good should be valued less by the obese. Elevation in thousands of feet above sea level, (*Elev)*, tends to repel the obese for two reasons. First, it is associated with uneven terrain and second with lower oxygen

13 This variable equals 1 if the MSA's primary city is located in a NOAA Coastal County.

content. City elevation is based on the elevation of the local weather station whose selection is discussed below.

Climate characteristics are potentially very important in determining differential amenity for the obese. The climate variables are measured using observations from NOAA weather stations; with preference given to major city airports as these typically report all climatic data of importance.[16] In cases where there is such airport weather station, this paper uses a weather station that reports all climatic data and is located near the city center. The climate variables listed in Table 2 include the average recorded temperature in January (*January*) and July (*July)*, annual precipitation (*Precip)*, and average sunshine for the months of January and July *(JanuarySun* and *JulySun)*. Using averages over the 1981-2010 period smooths idiosyncratic variation in temperature and sunshine assuming that individuals locate based on expectations of past climate. Individuals with low BMI should have a more negative physiological response to cold winters and prefer for summer climate that is not very hot or rainy so that outdoor activity, particularly exercise, is pleasant.

V. STOCHASTIC SPECIFICATION AND EMPIRICAL RESULTS

---

15 Parkland data for 49 MSAs comes from the 2010 American Fitness Index (AFI). The AFI adjusts Trust for Public Land (TPL) data to exclude parkland such as wildlife refuges (considered outside the built area of a city) such as those found in New Orleans. To supplement this data, we also use 2010 TPL data that allows us to add parkland data for 13 more MSAs. For these cities, only El Paso appears to have significant wildlife parkland that the AFI may have removed had El Paso been one of the 50 largest MSAs. Importantly, the estimation results are not sensitive whether this adjustment is to El Paso's data or not.

16 Of course, for some cities, the major airport is located in the center of the city (Washington, DC or San Diego, CA) or for others it is located several miles further out in the suburbs (San Francisco, CA or Chicago, IL).

The hypothesis tested here is that the variation in BMI across cities is due to both variation in individual characteristics and amenities among cities. Specific hypotheses regarding the effects of individual amenities were developed because of specific differences in the effects of BMI on preferences including physiological responses to amenities. The resulting BMI equation can be written as:

$$B_i = \alpha + \Sigma_h \lambda_h X_{hi} + \Sigma_j \beta_j Y_{ji} + \Sigma_k \theta_k Z_{ki} + \varepsilon_i \qquad (5)$$

where: B_i is the BMI of individual i, X_{hi} is a matrix of observations of income, education, race, ethnicity, gender and age variables, Y_{ji} is a matrix of observations of j area characteristics that might influence BMI, Z_{ki} is a matrix of amenity variables for which preferences may depend on BMI as discussed in the theory section. α, λ, β, and θ are parameters to be estimated, and ε is a residual error term. Lastly, measurement errors in the BRFSS survey may be correlated at the MSA level due to sampling factors. Accordingly, this paper clusters standard errors (therefore ε is not treated as iid in the estimation) at the MSA level and utilizes respondent survey weights from the CDC BRFSS data file.

There are strong prior expectations for the signs of most of the estimated parameters. Table 3 provides estimation results for BMI (*bmi)* across a number of different MSA samples due to the differential availability of sunshine and parkland data. Model 1 in the table displays individual level estimation results across the 110 MSAs with temperature and amenity data. Model 2 adds parkland data to the specification in Model 1, and accordingly restricts the analysis to the smaller sample of MSAs where this data is available. Model 3 contains estimation results for individuals across 78 MSAs with temperature, sunshine and amenity data. Model 4 in Table

3 adds in AFI/TPL parkland data to the specification from Model 3, which restricts the analysis to a smaller sample of MSAs where this data is available.

Estimation results across Models 1-4 in Table 3 yield exceptionally consistent results. BMI first rises and then falls with age. Women have slightly lower BMI than men. [17] Race and ethnicity have a substantial relation to BMI. Individuals with more education and household income have lower BMI. These are all personal characteristics of the resident population that, based on non-spatial analysis of the determinants of BMI were expected to be important. Among the non-amenity variables, higher food costs, as expected, are associated with lower BMI.

The expected relation between climate or topography amenity variables and BMI that was based on physiological factors tends to hold in the Table 3 results. Milder January temperatures (*January*) and/or increased sunshine result (*JanuarySun*) in lower individual BMI. Hotter July temperatures (*July*) discourage exercise and raise BMI. Of the topographic variables, only elevation (*elev*) is consistently statistically significant, and, as expected, it has a negative relation to BMI.

Apart from statistical significance, the effects of climatic and topographic effects can be appreciated by computing the difference in income or education needed to offset the effects of a one standard deviation shift in these variables. These tradeoffs are displayed in Table 4. Some are quite dramatic. First comparing amenity effects with income, the decline in BMI from a one standard deviation shift in elevation (*elev*) is equivalent to 2/3rds of the change due to a shift in household earnings from between 15 and 20 thousand dollars (*Income 15k - 20k*) to between 35

17 In the 2010 CDC BRFSS individual level dataset among the 110 MSAs studied, the average male BMI is 28.0. For women, their average BMI is 27.2. Therefore, the statistically significant and negative effect of gender on BMI is consistent with the underlying data.

and 50 thousand dollars (*Income 35k - 50k*). The expected change in BMI is similar whether July temperatures (*July*) shift up by a 1/3 standard deviation or household income rises from between 20,000 and 25,000 (*Income 20k - 25k*) to 50,000 to 75,000 (*Income 50k - 75k*). Concerning comparable effects of education on BMI, a 2/3 standard deviation shift upwards in grocery costs (*GroceryCOLA*) yields an equivalent reduction in BMI to education improving from having completed elementary school (*Some HS*) at most to having graduated from high school (*completed HS*). Further, a full standard deviation decline in July temperatures (*July*) yields an equivalent reduction in BMI to education improving from completing high school (*completed HS*) to having attended some college (*some college*).

These results not only provide evidence that local amenity and BMI are related they also indicate that, compared to important factors like income and education, the amenity effects are quantitatively important.

VI. ROBUSTNESS

As noted above, mesomorphs tend to have higher BMI due to greater muscle mass. The presence of mesomorphs works against finding differences in the spatial distribution of BMI based on amenity factors because their amenity preferences are similar to ectomorphs. In an attempt to minimize the influence of the mesomorphs, the first robustness check estimates equation (5) using a logit model where B_i is now a binary indicator variable equal to 1 if the person is obese or 0 if the person has normal weight. This estimation excludes underweight or overweight individuals, attempting to keep only ectomorphs and endomorphs in the estimation. The estimation results are reported in Table 5. Demographic, educational, and income characteristics are strongly related to incidence of obesity. Grocery costs (*GroceryCOLA)*, as expected, is negatively related to BMI.

City amenities, such as higher elevation and mild January (*Jan*) and July temperatures (*July*) are strongly statistically significantly explanatory variables to BMI and act in the expected direction based on theory and previous research. In addition, there is some evidence that housing density (*Density*) is negatively related to obesity, as perhaps urban cores with high density may simply be more walk-able. Overall, the results for the test attempting to remove the influence of mesomorphs reported in Table 5 are quite similar to those found for BMI in Table 3.

The theoretical relation between amenity characteristics and BMI is consistent with two possible mechanisms for achieving spatial differences in BMI. One possibility is a selection effect, in which individuals with a natural tendency to obesity migrate to areas with high amenity for the obese. Second is an adaptation effect in which the population adapts to the amenities in the surrounding area. The adaptation argument suggests that individuals born in an area will adjust to conditions thorough childhood and that differences in age-adjusted BMI should not vary with amenity factors. The selection effect suggests that individuals who have reached an age where migration is likely will shift locations based on amenity.

One partial test of these two possibilities is to determine if the differences associated with the amenity variables in Table 3 are robust to differences based on age. This test was performed by interacting the vector Z_k amenity variables in equation (5) with dummy variables for age less than 25 in order to determine if the amenity effects were different among individuals who likely were born in the city and adapted to amenities there rather than being selected to move there as adults. Rather than clutter the paper with further tables of results, the findings can be simply stated. The results reported in Table 3 are robust to differentiation by age < 25 in that none of the terms with amenity interacted with age < 25 were statistically significant. This suggests that, even without adult migration, differences in BMI associated with amenity variables are substantial. Put another way, the large climatic and topographic effects on BMI are not due to migration of adults selecting areas but either to past migration or to environmental adaptation. The results may be responses to these amenity factors that begin in youth or genetic selection from the past. Heterogeneity is related to differences in the resident population rather than based on differential migration of adults.

VII. CITY-PAIR EXAMPLES

This section explores the city-pairs mentioned previously to determine the relative importance of city amenities or city population characteristics in accounting for the large differences in BMI. The estimation results found in Table 3 allow these relative effects of amenity and non-amenity factors to be compared. The comparisons focus on the contributions of statistically significant variables in either Model 1 or Model 4 of Table 3. The city pairs are selected based on the observation of significant differences in average BMI.

San Francisco and Detroit in the first panel of Table 6 have very different BMI and population characteristics. Nevertheless, amenity differences, related to climate, are more important than demographic, education, and income differences in explaining the difference in BMI.

St. Louis and Honolulu differ significantly in both amenity and population characteristics. This results in a very large BMI difference. In this case, demographic differences are more influential but city amenities still explain one-third of the difference in

BMI. The third city pair, Pittsburgh and Denver, has differences in BMI that are similar to St. Louis and Honolulu but their population characteristics are not so dissimilar. As a result, the city amenity effects on BMI are far more important than the non-amenity variables in explaining the large BMI difference.

The preceding city-pair examples demonstrate that city-amenities, both topographic, and climatic, can explain as much or more of the differences in BMI among cities as differences in individual characteristics such as age, income, education, ethnicity, and race.

VIII. CONCLUSIONS

The initial research question was to determine whether the substantial differences in obesity across cities are due to effects associated with observable population characteristics known to explain the substantial BMI differences among individuals or if city amenities are also important in selecting individuals based on BMI. The answer is clear. While differences in individual income, education, age, race, ethnicity, and gender play a role in intercity variation in BMI, specific city amenity characteristics also matter.

Furthermore, the influence of these amenity characteristics agrees well with prior expectations based on physiological effects of various city amenities. The empirical results confirm the differential attraction for persons of healthy or obese weight to local area temperatures. Topography and elevation are also influential. Food prices even play the expected role. The influence of these factors is not only statistically significant, it is of practical significance in comparison to factors such and income and education. For example, the estimates imply that a small shift in July average temperatures can affect BMI as much as doubling (or perhaps tripling) of income along a certain income range.

An attempt to determine if differential migration is an important in determining the relation between city characteristics and BMI found that there were no differential effects of these characteristics on those under 25, i.e. those less likely to have migrated in response to city differences. This is not a very strong test and leaves open the question of the influence of past migration by those with a heredity tendency toward obesity. In sum, the question of whether these results reflect selection of individuals into locations or adaptation of identical individuals to different local conditions will require additional research.

The results raise concerns regarding omitted variables bias in economic research. BMI is generally not observed, but it is correlated with very important individual characteristics that are observed, particularly education and income, and it also varies with a variety of city amenity variables. Both categories of variables are important in empirical studies. For example, studies of intercity wage or productivity differentials include personal characteristics such as education, and city amenity variables. If these variables are correlated with omitted BMI, the effect of BMI on wages may bias estimates of the effects of education and city amenity on wages. City quality of life measures allow wages of residents to vary due to education. However, they assume that local amenity based on climate and topography does not select population based on factors like obesity that influence wages and productivity. It appears that the potential for unobserved differences in BMI to play a role in determining the way wages and productivity vary over space limits the ability to construct measures of wage differentials across cities for individuals who are otherwise observationally equivalent.

References

Andersson, J. and Andrew J. Walley. "The Contribution of Heredity to Clinical Obesity." *Obesity Before Birth. Endocrine Updates*, 30, 2011, 25-52.

Baum, C. and William Ford. "The wage effects of obesity: a longitudinal study." *Health Economics*, 13(9), 2004, 885–899.

Booth, K. M. "Obesity and the built environment." *Journal of the American Dietetic Association*, 105(5S), (2005), 110-117.

Brennan, T.J. "Behavioral economics and policy evaluation." *Journal of Benefit-Cost Analysis*, 5(1), 2014, 89-109.

Brookings Institution Study, accessed May 10,2015, http://www.brookings.edu/blogs/brookings-now/posts/2015/05/societal-costs-of-obesity.

Chen Ji, et. al. (2013). "Spatial variation of salt intake in Britain and association with socioeconomic status." *British Medical Journal*, 31, 2013.

Dehghan, H. et. al. "Cardiac strain comparison between workers with normal weight and overweight in the hot humid weather of the Persian Gulf region." *Journal of Education and Health Promotion,* 2, 2013.

Edwards, R. "Public transit, obesity, and medical costs: assessing the magnitudes." *Preventative Medicine*, 46(1), 2008, 14-21.

Eid, J. et. al. "Fat city: questioning the relation between urban sprawl and obesity." *Journal of Urban Economics*, 63, 2008, 385-404.

Ford, W. et. al. (2005). "Geographic variation in the prevalence of obesity, diabetes, and obesity-related behaviors." *Obesity Research*, 13(1), 2005, 118-22.

Gallup-Healthways Well-Being Index, accessed August 6, 2012, http://www.gallup.com/poll/156440/middle-aged-linked-having-higher-bmi.aspx.

Grossman, M. et. al. "Food Prices and Body Fatness among Youths." *NBER Working Paper No. 19143*, 2013.

Lin, G. "Climate Amenity and BMI." *Obesity*, 15(8), 2007, 2120-2127.

National Health and Nutrition Examination Survey (NHANES), accessed May 15, 2015, http://www.cdc.gov/nchs/pressroom/calendar/2014_schedule.htm#Subject.

Piziak, et. al. "Elevated body mass index in Texas Head Start children: a result of heredity and economics." *Southern Medical Journal*, 103(12), 2010, 1219-1222.

Roback, J. "Wages, Rents, and the Quality of Life". *The Journal of Political Economy*, 90(6), 1982, 1257-1278.

Voss, J.D., et. al. "Association of elevation, urbanization and ambient temperature with obesity prevalence in the United States." *International Journal of Obesity*, 37(10), 2013, 1407-1412.

Zhao, Z. and R. Kaestner. “Effects of urban sprawl on obesity.” *Journal of Health Economics*, 29, 2010, 779-787.

Table 1:

Descriptive Statistics: Individual Characteristics

Variable Name	Observations	Average	Standard Deviation	Minimum	Maximum
Body Mass Index					
bmi	191215	27.50	5.91	7.19	96.34
BMI Category					
Under-weight	2051	0.01	0.10	0.00	1.00
Normal-weight	67993	0.36	0.48	0.00	1.00
Overweight	69657	0.36	0.48	0.00	1.00
Obese	51514	0.27	0.44	0.00	1.00
Demographics					
Age	191215	55.85	16.79	7.00	98.00
Age^2	191215	3401.78	1858.70	49.00	9604.00
Age < 25	6632	0.03	0.18	0.00	1.00
Female	116288	0.61	0.49	0.00	1.00
Male	74927	0.39	0.49	0.00	1.00
White, Non-Hispanic	144255	0.75	0.43	0.00	1.00
Black, Non-Hispanic	19788	0.10	0.30	0.00	1.00
Hispanic	13755	0.07	0.26	0.00	1.00
Other	13417	0.07	0.26	0.00	1.00
Education					
Elem or less	4685	0.02	0.15	0.00	1.00
Some HS	9829	0.05	0.22	0.00	1.00
Completed HS	49045	0.26	0.44	0.00	1.00
Some College	49860	0.26	0.44	0.00	1.00
Completed BA	77486	0.41	0.49	0.00	1.00
Refused to Answer	310	0.00	0.04	0.00	1.00
Income					
Income <15	16459	0.09	0.28	0.00	1.00
Income 15k - 20k	11446	0.06	0.24	0.00	1.00
Income 20k - 25k	14610	0.08	0.27	0.00	1.00
Income 25k - 35k	18141	0.09	0.29	0.00	1.00
Income 35k - 50k	23862	0.12	0.33	0.00	1.00
Income 50k - 75k	26495	0.14	0.35	0.00	1.00
Income > 75k	56194	0.29	0.46	0.00	1.00
Refused to Answer	24008	0.13	0.33	0.00	1.00

Unweighted Survey Responses

Table 2:

Descriptive Statistics: City Characteristics

	Observations	Average	Standard Deviation	Minimum	Maximum	25th %tile	75th %tile
Average BMI	110	27.68	0.57	26.14	29.06	27.36	28.02
Urban Characteristics							
GroceryCOLA	110	100.93	11.07	79.80	160.10	92.7	105.9
Transit(%)	110	0.03	0.04	0.00	0.31	0.01	0.03
Density	110	197.90	196.19	9.00	1125.70	66.9	259.7
Pop(100k)	110	17.70	25.82	0.92	189.20	4.15	20.76
Topography							
Waterarea	110	0.09	0.14	0.00	0.72	0.009	0.08
Coastal	110	0.35	0.48	0.00	1.00	N/A	N/A
Elev	110	1.04	1.40	0.00	6.18	0.132	1.064
Parkland(%)	62	10.14	6.03	1.58	28.14	5.3	14.2
Climate							
January	110	38.07	12.14	9.30	73.20	29.2	46.3
July	110	77.55	6.05	57.70	94.20	73.5	82.1
JanuarySun	78	52.15	12.07	28.00	80.00	45	58
JulySun	78	70.27	8.11	57.00	97.00	64	74
Precip	110	35.77	15.05	4.19	66.15	22.2	46.63

Table 3:

Empirical Results (Dependent Variable – Individual BMI)

	Empirical Model : Individual Level			
	1	2	3	4
Dependent Variable:	BMI			
Independent Variables:				
Individual Characteristics				
Demographics				
Age	0.3187***	0.3186***	0.3151***	0.3160***
	(0.000)	(0.000)	(0.000)	(0.000)
Age^2	-0.0030***	-0.0030***	-0.0029***	-0.0029***
	(0.000)	(0.000)	(0.000)	(0.000)
Female	-1.0037***	-1.0282***	-1.0273***	-1.0520***
	(0.000)	(0.000)	(0.000)	(0.000)
Black, Non-Hispanic	1.9072***	1.9295***	1.8789***	1.9023***
	(0.000)	(0.000)	(0.000)	(0.000)
Hispanic	0.9293***	0.9397***	0.9848***	0.9567***
	(0.000)	(0.000)	(0.000)	(0.000)
Other	-0.7340***	-0.7789***	-0.7711***	-0.8027***
	(0.000)	(0.000)	(0.000)	(0.000)
Education				
Some HS	-0.0482	-0.0215	-0.1518	-0.1314
	(0.844)	(0.936)	(0.587)	(0.658)
Completed HS	-0.1028	-0.1267	-0.1548	-0.1829
	(0.590)	(0.544)	(0.487)	(0.434)
Some College	-0.1911	-0.2405	-0.2318	-0.2968
	(0.350)	(0.280)	(0.327)	(0.227)
Completed BA	-1.2441***	-1.3034***	-1.3063***	-1.3599***
	(0.000)	(0.000)	(0.000)	(0.000)
Education Not Reported	-1.4509***	-1.6326***	-1.3665***	-1.5105***
	(0.001)	(0.000)	(0.007)	(0.004)
Income				
Income 15k - 20k	-0.4540**	-0.4246**	-0.4944***	-0.4949**
	(0.012)	(0.034)	(0.009)	(0.014)
Income 20k - 25k	-0.7795***	-0.7455***	-0.8436***	-0.8105***
	(0.000)	(0.000)	(0.000)	(0.000)
Income 25k - 35k	-0.8059***	-0.8105***	-0.8405***	-0.8578***
	(0.000)	(0.000)	(0.000)	(0.000)
Income 35k - 50k	-0.7704***	-0.7588***	-0.8600***	-0.8455***
	(0.000)	(0.001)	(0.000)	(0.000)
Income 50k - 75k	-0.8162***	-0.7582***	-0.8608***	-0.8099***
	(0.000)	(0.000)	(0.000)	(0.000)
Income > 75k	-1.3917***	-1.3476***	-1.4453***	-1.4075***
	(0.000)	(0.000)	(0.000)	(0.000)
Income Not Reported	-1.3573***	-1.3176***	-1.3646***	-1.3394***
	(0.000)	(0.000)	(0.000)	(0.000)

Area Characteristics				
Urban				
GroceryCOLA	-0.0112**	-0.0107*	-0.0068	-0.0067
	(0.046)	(0.072)	(0.168)	(0.206)
Transit(%)	-0.2338	-0.4715	-1.0053	-1.8639
	(0.841)	(0.732)	(0.398)	(0.234)
Density	-0.0003	-0.0000	0.0004	0.0008
	(0.424)	(0.997)	(0.359)	(0.191)
Pop(100k)	-0.0000	-0.0010	-0.0023	-0.0029
	(0.981)	(0.623)	(0.328)	(0.214)
Topography				
Waterarea	0.2640	0.3880	0.1251	0.2826
	(0.490)	(0.369)	(0.739)	(0.522)
Coastal	-0.2355	-0.2707	-0.3312*	-0.4008*
	(0.121)	(0.147)	(0.065)	(0.072)
Elev	-0.1460***	-0.1508***	-0.1141***	-0.0984**
	(0.000)	(0.000)	(0.006)	(0.041)
Parkland(%)		0.0012		0.0106
		(0.911)		(0.384)
Climate				
January	-0.0129***	-0.0134***	-0.0067	-0.0061
	(0.000)	(0.001)	(0.106)	(0.205)
July	0.0152**	0.0208**	0.0268***	0.0305***
	(0.049)	(0.028)	(0.001)	(0.001)
JanuarySun			-0.0150***	-0.0169***
			(0.003)	(0.007)
JulySun			-0.0037	-0.0030
			(0.629)	(0.710)
Precip	0.0062*	0.0057	0.0015	0.0016
	(0.062)	(0.139)	(0.711)	(0.704)
Constant	22.2606***	21.7585***	22.0064***	21.5605***
	(0.000)	(0.000)	(0.000)	(0.000)
# Obs	191215	149430	161763	138763
# CBSAs	110	62	78	54
adj. R-sq	0.0776	0.0800	0.0784	0.0805
F	312.4778	386.0102	496.6299	635.5294
p>F	0.0000	0.0000	0.0000	0.0000

Table 4:

Income and Education Equivalents using Table 3 Model 1 Estimated Effects

Change in Income Category		
A change of :	From:	To
4/3 standard deviation increase in Elev	Income 15k - 20k	Income 35k - 50k
1/4 standard deviation increase in GroceryCOLA	Income 20k - 25k	Income 50k - 75k
1/4 standard deviation increase in January (temperature)	Income 20k - 25k	Income 50k - 75k
1/3 standard deviation decrease in July (temperature)	Income 20k - 25k	Income 50k - 75k
1/3 standard deviation decrease in Precip	Income 20k - 25k	Income 50k - 75k
Change in Education Category		
A change of :	From:	To
1 standard deviation increase in Elev	Elementary or less	Some College
2/3 standard deviation increase in Grocery Cola	Elementary or less	Completed HS
1/4 standard deviation increase in January (temperature)	Elementary or less	Completed HS
1/3 standard deviation decrease in July (temperature)	Completed HS	Some College
1/3 standard deviation decrease in Precip	Completed HS	Some College

Table 5:

Empirical Results (Dependent Variable – Binary Obesity Indicator)

	Empirical Model : Individual Level			
	1	2	3	4
Dependent Variable:	Logit (0=Normal-weight, 1=Obese)			
Independent Variables:				
Individual Characteristics				
Demographics				
Age	0.1476***	0.1489***	0.1466***	0.1481***
	(0.000)	(0.000)	(0.000)	(0.000)
Age^2	-0.0014***	-0.0014***	-0.0014***	-0.0014***
	(0.000)	(0.000)	(0.000)	(0.000)
Female	-0.5641***	-0.5717***	-0.5673***	-0.5766***
	(0.000)	(0.000)	(0.000)	(0.000)
Black, Non-Hispanic	0.7766***	0.7801***	0.7606***	0.7673***
	(0.000)	(0.000)	(0.000)	(0.000)
Hispanic	0.4094***	0.4231***	0.4194***	0.4183***
	(0.000)	(0.000)	(0.000)	(0.000)
Other	-0.4060***	-0.4346***	-0.4256***	-0.4469***
	(0.000)	(0.000)	(0.000)	(0.000)
Education				
Some HS	-0.2914***	-0.3063***	-0.3236***	-0.3268***
	(0.001)	(0.002)	(0.001)	(0.002)
Completed HS	-0.3041***	-0.3254***	-0.2996***	-0.3138***
	(0.000)	(0.000)	(0.001)	(0.001)
Some College	-0.3433***	-0.3691***	-0.3402***	-0.3635***
	(0.000)	(0.000)	(0.001)	(0.000)
Completed BA	-0.8574***	-0.8865***	-0.8638***	-0.8807***
	(0.000)	(0.000)	(0.000)	(0.000)
Education Not Reported	-0.8513***	-0.8718***	-0.8072**	-0.8182**
	(0.003)	(0.005)	(0.013)	(0.016)
Income				
Income 15k - 20k	-0.0707	-0.0633	-0.0888	-0.0887
	(0.322)	(0.428)	(0.259)	(0.288)
Income 20k - 25k	-0.1619***	-0.1487***	-0.1772***	-0.1678***
	(0.000)	(0.001)	(0.000)	(0.000)
Income 25k - 35k	-0.1545**	-0.1517**	-0.1561**	-0.1614*
	(0.022)	(0.044)	(0.043)	(0.051)
Income 35k - 50k	-0.1056*	-0.1053*	-0.1304**	-0.1269**
	(0.053)	(0.087)	(0.027)	(0.047)
Income 50k - 75k	-0.1700***	-0.1512***	-0.1762***	-0.1626***
	(0.000)	(0.003)	(0.001)	(0.003)
Income > 75k	-0.3928***	-0.3717***	-0.4068***	-0.3912***
	(0.000)	(0.000)	(0.000)	(0.000)
Income Not Reported	-0.4081***	-0.3982***	-0.4045***	-0.4044***
	(0.000)	(0.000)	(0.000)	(0.000)

Area Characteristics				
Urban				
GroceryCOLA	-0.0046**	-0.0045*	-0.0033	-0.0034
	(0.034)	(0.059)	(0.123)	(0.145)
Transit(%)	-0.1942	-0.3378	-0.4178	-0.7403
	(0.693)	(0.556)	(0.423)	(0.262)
Density	-0.0003**	-0.0002	-0.0001	0.0001
	(0.022)	(0.323)	(0.723)	(0.669)
Pop(100k)	0.0007	0.0003	-0.0001	-0.0004
	(0.418)	(0.764)	(0.905)	(0.648)
Topography				
Waterarea	0.2308	0.2546	0.1846	0.2258
	(0.139)	(0.157)	(0.241)	(0.222)
Coastal	-0.0794	-0.0921	-0.1053	-0.1300
	(0.176)	(0.200)	(0.144)	(0.139)
Elev	-0.0673***	-0.0683***	-0.0629***	-0.0565***
	(0.001)	(0.001)	(0.000)	(0.006)
Parkland(%)		0.0012		0.0041
		(0.786)		(0.424)
Climate				
January	-0.0045***	-0.0049***	-0.0029	-0.0028
	(0.003)	(0.003)	(0.112)	(0.184)
July	0.0057*	0.0078**	0.0093***	0.0113***
	(0.060)	(0.034)	(0.003)	(0.001)
JanuarySun			-0.0046**	-0.0056**
			(0.019)	(0.026)
JulySun			-0.0002	0.0001
			(0.950)	(0.978)
Precip	0.0019	0.0017	0.0008	0.0006
	(0.188)	(0.268)	(0.634)	(0.741)
Constant	-2.5472***	-2.7726***	-2.7162***	-2.9307***
	(0.000)	(0.000)	(0.000)	(0.000)
# Obs	119507	93347	101134	86715
# CBSAs	110	62	78	54
pseudo R-sq	0.0892	0.0917	0.0895	0.0917
chi-sq	6483.8102	6424.5463	8298.3207	1.05e+04
p>chi-sq	0.0000	0.0000	0.0000	0.0000

Table 6:

City-Pair Examples (Using Model 1 and 4 Estimates in Table 3)

City Pair: Detroit, MI & San Francisco, CA

	Detroit	San Francisco
Average BMI	28.02	26.08
Using Model 1 Significant Variables	Explains Raw Difference in BMI	Explains % of Difference In BMI
City Amenities		
Topography	-0.09	-4.70%
Climate	0.65	33.42%
Individual Characteristics		
Demographics	0.24	12.18%
Education	0.18	9.34%
Income	0.02	0.89%
Summary		
City Amenities Explain :	0.56	28.71%
Indv. Characteristics Explain:	0.43	22.40%

City Pair: St. Louis MO & Honolulu, HI

	St. Louis	Honolulu
Average BMI	28.07	26.49
Using Model 1 Significant Variables	Explains Raw Difference in BMI	Explains % of Difference In BMI
City Amenities		
Topography	-0.08	-4.86%
Climate	0.66	42.18%
Individual Characteristics		
Demographics	0.82	52.09%
Education	0.04	2.53%
Income	-0.02	-1.33%
Summary		
City Amenities Explain :	0.59	37.32%
Indv. Characteristics Explain:	0.84	53.28%

City Pair: Detroit, MI & San Francisco, CA

	Detroit	San Francisco
Average BMI	28.02	26.08
Using Model 4 Significant Variables	Explains Raw Difference in BMI	Explains % of Difference In BMI
City Amenities		
Topography	-0.06	-3.17%
Climate	0.76	39.07%
Individual Characteristics		
Demographics	0.25	12.78%
Education	0.20	10.21%
Income	0.02	0.80%
Summary		
City Amenities Explain :	0.69	35.90%
Indv. Characteristics Explain:	0.46	23.78%

City Pair: St. Louis MO & Honolulu, HI

	St. Louis	Honolulu
Average BMI	28.07	26.49
Using Model 4 Significant Variables	Explains Raw Difference in BMI	Explains % of Difference In BMI
City Amenities		
Topography	0.35	22.19%
Climate	0.22	13.78%
Individual Characteristics		
Demographics	0.86	54.39%
Education	0.04	2.76%
Income	-0.02	-1.27%
Summary		
City Amenities Explain :	0.57	35.98%
Indv. Characteristics Explain:	0.88	55.88%

City Pair: Pittsburgh, PA & Denver, CO

	Pittsburgh	Denver
Average BMI	28.10	26.43
Using Model 1 Significant Variables	Explains Raw Difference in BMI	Explains % of Difference In BMI
City Amenities		
Topography	0.60	35.77%
Climate	0.16	9.43%
Individual Characteristics		
Demographics	0.02	1.48%
Education	0.12	6.90%
Income	0.06	3.49%
Summary		
City Amenities Explain :	0.75	45.20%
Indv. Characteristics Explain:	0.20	11.87%

City Pair: Pittsburgh, PA & Denver, CO

	Pittsburgh	Denver
Average BMI	28.10	26.43
Using Model 4 Significant Variables	Explains Raw Difference in BMI	Explains % of Difference In BMI
City Amenities		
Topography	0.40	24.11%
Climate	0.63	37.72%
Individual Characteristics		
Demographics	0.06	3.74%
Education	0.13	7.54%
Income	0.06	3.39%
Summary		
City Amenities Explain :	1.03	61.82%
Indv. Characteristics Explain:	0.24	14.68%