

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Geographic Scale Matters in Detecting the Relationship between Neighborhood Food Environments and Obesity Risk: An Analysis of Driver License Records in Salt Lake County, Utah
AUTHORS	Fan, Jessie; Hanson, Heidi; Zick, Cathleen; Brown, Barbara; Kowaleski-Jones, Lori; Smith, Ken

VERSION 1 - REVIEW

REVIEWER	Tomoya Hanibuchi School of International Liberal Studies, Chukyo University, Japan
REVIEW RETURNED	15-May-2014

GENERAL COMMENTS	<ul style="list-style-type: none">• Please provide a rationale for using only a 1km radial buffer in this study.• Please elaborate the rationale for using "median age of neighborhood housing" as an index of "land use diversity" in more detail.• The interpretation of the statistical significance was not clear. The authors regarded the p-value of 0.08 as being marginally statistically significant (p.10, line 6). However, it seems that some associations with p-values larger than 0.1 were also interpreted as being statistically significant in the Abstract, Result, Discussion, and Table 5 [limited-service restaurant at the ZIP-code level for women (p=0.1054), and full service restaurant at the block level for women (p=0.1008)]. Despite the authors mentioning the p-value problem in large samples, it seems that the interpretations were less conservatively made than usual. It is not clear what "large p-value" means (in p.11, line 48). Is this something like conservative p-value?• In table 5, the authors wrote that "ZIP-code or larger" (for men) and "larger than ZIP-code"(for women) are the areas of significance for the association between obesity risk and large grocery stores. However, as the authors themselves mentioned in the note, they are only hypotheses rather than the summary of the results. The expressions "larger" or "smaller" should be removed from the table and the results should be summarized more precisely.• In Table 2, the SD of the average area size at the ZIP-code level is missing.
-------------------------	---

REVIEWER	Diego Rose School of Public Health and Tropical Medicine
-----------------	---

	Tulane University USA
REVIEW RETURNED	13-Jun-2014

GENERAL COMMENTS	<p>Abstract needs minor modification to be balanced (see comment #4 on attached review).</p> <p>References need to be expanded to be more complete (see points #2, 5, and 13 on attached review).</p> <p>This is an interesting and well-written paper on the role of geographic scale in assessments of the food environment and its relationship to health outcomes, specifically obesity. The study is based in Salt Lake County, Utah, and draws on a novel source of data – driver's licenses – which in addition to height, weight, and basic demographics, provides address information, allowing for the extensive exploration of different-sized neighborhoods that forms the core of the study. The paper would definitely provide an important contribution to the literature.</p> <p>There are a number of shortcomings to the paper as currently written. All of these could be addressed. Below I suggest main areas that would benefit from more work. Following that I suggest some minor changes that could improve the paper.</p> <p><u>1. Lack of consistency with published research.</u> There is a notable lack of consistency between the results presented here and what has been previously published. This is true for a number of specific issues, and while this is not a critical problem in and of itself, when combined with a lack of acknowledgment and discussion, it does become a problem. In addition to some brief discussion for each of the issues below, it might be useful to put Salt Lake County in the title of the paper, as some of these differences might be specific to this sample.</p> <p style="padding-left: 40px;"><u>A. Obesity prevalence.</u> Since this is the key outcome variable it will be important to at least mention the observed rate and how it compares to other sources. For example BRFSS 2012 data, which is based on self-reported heights and weights, suggest that Utah's obesity rate is 24% as opposed to the 16 to 18% reported in this study. The authors do a good job of explaining why this is probably not a problem for this study (i.e. reporting bias not associated with geography), but the article would benefit from a couple more key sentences (and a citation or two) in Results and Discussion about the prevalence rates differing from other estimates, and perhaps why.</p>
-------------------------	---

B. Wrong sign on convenience stores. The author's acknowledge this problem, but it flies in the face of everything we know about the mechanism for which the store environment influences the risk of obesity. These stores stock a large supply of energy-dense snack foods and beverages and, as the name implies, provides them conveniently throughout the city. They also contain very few low-calorie options, like fruits and vegetables, and typically provide them at high cost. The authors make the argument that having these stores in the neighborhood promotes more exercise in getting to the stores, but why would this be different than other store types, or for that matter other cities? If you think people are more likely to walk to convenience stores (as opposed to other stores), can you try interacting convenience stores with one of the walkability measures (e.g. intersection density) to see if this makes a difference?

C. African-Americans and obesity. African-Americans typically have a higher obesity rate than whites, but your results show the opposite for men. There are very small percentages of African-Americans in your sample, so I wouldn't make a big deal about this. But it is a notable difference that could be mentioned with a citation to the literature.

2. Limitations on using broad-scale business directories. There is no acknowledgement of the limitations of using Dun and Bradstreet, or other general business directories to describe the food environment. These directories have significant error rates with documented problems in listing stores that have closed, not listing stores that have opened, and improper classification of store types. Moreover, food store typologies can be limited in that availability and price of foods can vary dramatically within a store type. This doesn't mean these business directories cannot be used. But rather, some acknowledgement of their limitations is needed with reference to some of the research that has been done in this area (see citations below). Consider also listing this as a limitation of the study (on page 3).

Gustafson AA, Lewis S, Wilson C, Jilcott-Pitts S. Validation of food store environment secondary data source and the role of neighborhood deprivation in Appalachia, Kentucky. *BMC Public Health*. 2012;12:688-688.

Liese AD, Colabianchi N, Lamichhane AP, et al. Validation of 3 food outlet databases: completeness and geospatial accuracy in rural and

urban food environments. *American Journal Of Epidemiology*. 2010;172(11):1324-1333.

Powell LM, Han E, Zenk SN, et al. Field validation of secondary commercial data sources on the retail food outlet environment in the U.S. *Health & Place*. 2011;17(5):1122-1131.

Rose D, Bodor JN, Swalm C, Rice JC, Farley TA, Hutchinson PL. Deserts in New Orleans? Illustrations of urban food access and implications for policy. *Understanding the Economic Concepts and Characteristics of Food Access*. UM and USDA-ERS: Washington, DC; January, 2009. Available at: http://www.npc.umich.edu/news/events/food-access/rose_et_al.pdf .

3. Consider citing novel use of driver's license method as a strength of the study. If you list the problem cited in #2 of using business directories as a weakness, consider citing an additional strength for balance. The use of driver's license data is relatively novel and allows for a very large sample size for analysis of geographic determinants of obesity. Also, in the name of balance, since I'm suggesting more discussion of limitations, you could remove current discussion of limitations, that's relatively minor. For example on page 14, your third limitation is about the restricted age of the sample. I don't see this as a big deal. That's just the age of your sample, which is adequately described in your methods. So you could remove this point from the discussion on page 14.

4. Lack of significance on large groceries. You state in the abstract that large groceries are significant at the zip code level. Most biomedical journals use an alpha level of .05, so your result for men would not be statistically significant, and it's not even close for women. Obviously the .05-level is arbitrary, but it is the convention. It might make sense to bend this when you have a small sample size that leads to somewhat imprecise estimates. But you have 400,000 observations. So, I would delete this statement from the abstract, as this result does not deserve top billing. And you will probably want to reframe how you address this in the rest of the paper.

5. Citing of previous research on geographic scale. There are a number of authors who have looked at geographic scale in terms of the food environment and diet or weight status outcomes. You should discuss this work, compare it to yours, and cite them in your paper. These include, for example:

Hutchinson PL, Bodor JN, Swalm CM, Rice JC, Rose D.
"Neighbourhood food environments and obesity in southeast Louisiana," *Health & Place* 2012;18:854-860.

Rose D, Hutchinson PL, Bodor JN, Swalm CM, Farley TA, Cohen DA, Rice JC. "Neighborhood Food Environments and Body Mass Index: The Importance of In-Store Contents," *American Journal of Preventive Medicine*, 2009;37:214-219.

The second of these actually looked at 5 different scales, so you might want to modify your key strength number #1.

6. ID cards versus driver's licenses for additional modifying information about scale. On page 14, line 34 you refer to a driver privilege card for the first time. This might be relevant if you also have in your database individuals who have ID cards, rather than driver's licenses. This would indicate that they are not drivers, so would likely affect the scale of their food environment. If you have such a variable, it would be interesting to use this as an independent variable in your regressions, or, with enough cases, perhaps stratify by drivers and non-drivers. If you don't have this information, you should either describe earlier the difference between a license and privilege card. Or, if not relevant at all, just drop the usage on page 14.

7. Consider including a table of correlations between outlet types at each scale. I'm surprised you didn't have more collinearity problems when you put all retail outlet variables into the same model (assuming you did – see point #11 below). You might better convince the reader of this by including a table of correlations between counts of different store types at each scale. This could also be an appendix table, if you don't think it is central enough to your flow.

Minor points:

8. Page 3, Line 50. Include "weight and" in the phrase "self-reported height data".

9. Page 7, Line 32 and 37. It's not clear how "mean age of

neighborhood housing" is a measure of "land use diversity". This seems a bit of a stretch. Either explain how it is a measure of diversity, or drop the variable from analysis, or simply edit the sentence to eliminate reference to the "3Ds" (though you can still cite the same reference #29) and refer instead to neighborhood characteristics including those associated with walkability.

10. Page 8, Line 46. Since you don't present results on the pooled sample (both genders together) and there is no reason to present these, just delete mention of it (i.e. delete "the pooled sample of both genders and also stratified models for"). Otherwise it confuses the reader.

11. Page 8, last paragraph. It is not clear from the description of your regression models whether you included variables for all four retail types (groceries, restaurants, etc) in one model at each scale, or whether you ran separate models for each retail type at each scale. I assumed you included all retail outlet variables in the same model (for each geog scale), because you talk about collinearity diagnostics. But either way, you should make this explicit.

12. Page 10, 11. You have some odd findings – older individuals associated with higher obesity risk, but older median age associated with lower risk. Also higher intersection density is associated with higher level of risk, but higher population density is associated with lower. Have you tried dropping some of these redundant variables? For example, why have median neighborhood age, if you have individual's age? Is my weight status affected by the age of my neighbors? Also, have you tried dropping the "older housing stock variable" or one of the walkability variables to see if you would get more consistent results, i.e. increased walkability leads to lower risk of obesity?

13. Page 12, Lines 30-40. You should include reference in your discussion to the following citation, since it is relevant to this section.

Bodor JN, Farley TA, Rice JC, Swalm C, Rose D. "The Association between Obesity and Urban Food Environments," *Journal of Urban Health* 2010;87:771-781.

14. Page 13, Lines 15-22. Same comment as above.

	<p>15. Figure 1 – Needs a bit more clarity. On the legend, when printed in black and white, the 1km buffer and census tract look the same. Consider making the 1km buffer a dotted line.</p> <p>16. Table 2 – Why present population-weighted counts if it this wasn't used in the regression? Tract delineation is roughly determined by population size, so it's not really necessary.</p> <p>17. Table 5 – Grocery stores not really significant for men at zip-code level, so make first row cell contents for men the same as for women.</p>
--	---

VERSION 1 – AUTHOR RESPONSE

Reviewer Name Diego Rose
 Institution and Country School of Public Health and Tropical Medicine
 Tulane University
 USA
 Please state any competing interests or state 'None declared': None declared

This is an interesting and well-written paper on the role of geographic scale in assessments of the food environment and its relationship to health outcomes, specifically obesity. The study is based in Salt Lake County, Utah, and draws on a novel source of data – driver's licenses – which in addition to height, weight, and basic demographics, provides address information, allowing for the extensive exploration of different-sized neighborhoods that forms the core of the study. The paper would definitely provide an important contribution to the literature.

There are a number of shortcomings to the paper as currently written. All of these could be addressed. Below I suggest main areas that would benefit from more work. Following that I suggest some minor changes that could improve the paper.

1. Lack of consistency with published research. There is a notable lack of consistency between the results presented here and what has been previously published. This is true for a number of specific issues, and while this is not a critical problem in and of itself, when combined with a lack of acknowledgment and discussion, it does become a problem. In addition to some brief discussion for each of the issues below, it might be useful to put Salt Lake County in the title of the paper, as some of these differences might be specific to this sample.

R: We have added "Salt Lake County, Utah" to the title of the paper. We also added a paragraph in the discussion section on the issue of lack of consistency with published research. See details below.

A. Obesity prevalence. Since this is the key outcome variable it will be important to at least mention the observed rate and how it compares to other sources. For example BRFSS 2012 data, which is based on self-reported heights and weights, suggest that Utah's obesity rate is 24% as opposed to the 16 to 18% reported in this study. The authors do a good job of explaining why this is probably not a problem for this study (i.e. reporting bias not associated with geography), but the article would benefit from a couple more key sentences (and a citation or two) in Results and Discussion about the prevalence rates differing from other estimates, and perhaps why.

R: The relatively low obesity rate estimated in this paper is likely because we selected the DLD record closest to 2000 to synchronize with the 2000 Census for any given individual. According to historical obesity statistics on CDC's website (<http://www.cdc.gov/obesity/data/adult.html>), in 2000, adult obesity rate in Utah was between 15%-19% while in 2010 the obesity rate in Utah increased to 20-24%, based on BRFSS data. Given the timing of records, our obesity rate is in line with BRFSS estimates. We have added a sentence in the results section to note this with a CDC reference citation at the beginning of the "Results" section:

The estimated obesity rate was 17.16% for adults 25-64 in Salt Lake County around year 2000, with 18.28% for men and 15.94% for women. This is in line with obesity rate estimates from the Behavioral Risk Factor Surveillance System (BRFSS) of between 15% to 19% for Utah in 2000 [38].

B. Wrong sign on convenience stores. The author's acknowledge this problem, but it flies in the face of everything we know about the mechanism for which the store environment influences the risk of obesity. These stores stock a large supply of energy-dense snack foods and beverages and, as the name implies, provides them conveniently throughout the city. They also contain very few low-calorie options, like fruits and vegetables, and typically provide them at high cost. The authors make the argument that having these stores in the neighborhood promotes more exercise in getting to the stores, but why would this be different than other store types, or for that matter other cities? If you think people are more likely to walk to convenience stores (as opposed to other stores), can you try interacting convenience stores with one of the walkability measures (e.g. intersection density) to see if this makes a difference?

R. We have consistently found this relationship between convenience stores and obesity/BMI using Salt Lake County data. As per your suggestion, we interacted median housing age, as an indicator for walkability, with convenience store counts. The interaction term was significant only in one out of eight models with a small positive coefficient (1km buffer measure for women). The convenience store coefficients remained almost the same in all models. We will continue to look into possible explanations for this counterintuitive relationship in the future.

African-Americans and obesity. African-Americans typically have a higher obesity rate than whites, but your results show the opposite for men. There are very small percentages of African-Americans in your sample, so I wouldn't make a big deal about this. But it is a notable difference that could be mentioned with a citation to the literature.

R: We have noted this with two references in the "Discussion" section:

In addition, an increase in the proportion of blacks in the neighborhood was related to lower odds of obesity, which is counter to the picture provided by national statistics [43]. In the 2000 Census for Utah, blacks constituted fewer than 1% of the residents and had somewhat higher incomes than their national counterparts [44]. As such, our results may not generalize to neighborhoods in other parts of the country that have higher concentration of black residents with lower income.

2. Limitations on using broad-scale business directories. There is no acknowledgement of the limitations of using Dun and Bradstreet, or other general business directories to describe the food environment. These directories have significant error rates with documented problems in listing stores that have closed, not listing stores that have opened, and improper classification of store types. Moreover, food store typologies can be limited in that availability and price of foods can vary dramatically within a store type. This doesn't mean these business directories cannot be used. But rather, some acknowledgement of their limitations is needed with reference to some of the research that has been done in this area (see citations below). Consider also listing this as a limitation of the study (on page 3).

Gustafson AA, Lewis S, Wilson C, Jilcott-Pitts S. Validation of food store environment secondary data source and the role of neighborhood deprivation in Appalachia, Kentucky. *BMC Public Health*. 2012;12:688-688.

Liese AD, Colabianchi N, Lamichhane AP, et al. Validation of 3 food outlet databases: completeness and geospatial accuracy in rural and urban food environments. *American Journal Of Epidemiology*. 2010;172(11):1324-1333.

Powell LM, Han E, Zenk SN, et al. Field validation of secondary commercial data sources on the retail food outlet environment in the U.S. *Health & Place*. 2011;17(5):1122-1131.

Rose D, Bodor JN, Swalm C, Rice JC, Farley TA, Hutchinson PL. Deserts in New Orleans? Illustrations of urban food access and implications for policy. *Understanding the Economic Concepts and Characteristics of Food Access*. UM and USDA-ERS: Washington, DC; January, 2009. Available at: http://www.npc.umich.edu/news/events/food-access/rose_et_al.pdf .

R: We have added these two limitations in our "Discussion" section, and cited the above references. Thank you for reminding us of these issues as important limitations.

Third, commercial business directories such as the D&B are subject to significant error rates in listing and in classification [47-49]. Fourth, availability and price of foods can vary dramatically within a store type, and as such, food store typologies can be limited in capturing the impact of food environment on consumer decisions regarding food consumption [22, 50].

3. Consider citing novel use of driver's license method as a strength of the study. If you list the problem cited in #2 of using business directories as a weakness, consider citing an additional strength for balance. The use of driver's license data is relatively novel and allows for a very large sample size for analysis of geographic determinants of obesity. Also, in the name of balance, since I'm suggesting more discussion of limitations, you could remove current discussion of limitations, that's relatively minor. For example on page 14, your third limitation is about the restricted age of the sample. I don't see this as a big deal. That's just the age of your sample, which is adequately described in your methods. So you could remove this point from the discussion on page 14.

R: We have revised our manuscript as suggested. We removed the limitation on age range, and added the utilization of driver license records as a strength in the "Discussion" section:

The utilization of driver license records has the advantage of extensive local coverage with a large sample size, allowing us to test a wide range of definitions of neighborhood, and to estimate models for men and women separately.

4. Lack of significance on large groceries. You state in the abstract that large groceries are significant at the zip code level. Most biomedical journals use an alpha level of .05, so your result for men would not be statistically significant, and it's not even close for women. Obviously the .05-level is arbitrary, but it is the convention. It might make sense to bend this when you have a small sample size that leads to somewhat imprecise estimates. But you have 400,000 observations. So, I would delete this statement from the abstract, as this result does not deserve top billing. And you will probably want to reframe how you address this in the rest of the paper.

R: We now use 0.05 as the critical p-value of significance according to your suggestion. Our estimates have changed somewhat due to dropping the "median age of residents" variable as suggested. The details of our manuscript have been revised accordingly but our key findings remain the same.

5. Citing of previous research on geographic scale. There are a number of authors who have looked at geographic scale in terms of the food environment and diet or weight status outcomes. You should discuss this work, compare it to yours, and cite them in your paper. These include, for example:

Hutchinson PL, Bodor JN, Swalm CM, Rice JC, Rose D. "Neighbourhood food environments and obesity in southeast Louisiana," *Health & Place* 2012;18:854-860.

Rose D, Hutchinson PL, Bodor JN, Swalm CM, Farley TA, Cohen DA, Rice JC. "Neighborhood Food Environments and Body Mass Index: The Importance of In-Store Contents," American Journal of Preventive Medicine, 2009;37:214-219.

The second of these actually looked at 5 different scales, so you might want to modify your key strength number #1.

R: We have included these papers in our background section. We have also modified our key strength accordingly :

When a buffer method was used, there was a wide range of buffer distances, ranging from 100 meters [20] to 5 km [21], with two studies examining multiple distances from what's considered walkable (e.g., 500 meters and 1 km) to needing transportation (e.g., 2 km and 5 km) [21, 22].

5. ID cards versus driver's licenses for additional modifying information about scale. On page 14, line 34 you refer to a driver privilege card for the first time. This might be relevant if you also have in your database individuals who have ID cards, rather than driver's licenses. This would indicate that they are not drivers, so would likely affect the scale of their food environment. If you have such a variable, it would be interesting to use this as an independent variable in your regressions, or, with enough cases, perhaps stratify by drivers and non-drivers. If you don't have this information, you should either describe earlier the difference between a license and privilege card. Or, if not relevant at all, just drop the usage on page 14.

R: We have deleted the reference to driver's privilege cards as suggested.

6. Consider including a table of correlations between outlet types at each scale. I'm surprised you didn't have more collinearity problems when you put all retail outlet variables into the same model (assuming you did – see point #11 below). You might better convince the reader of this by including a table of correlations between counts of different store types at each scale. This could also be an appendix table, if you don't think it is central enough to your flow.

R: We thought that adding more details of our collinearity tests might address the concern of multicollinearity among the independent variables more adequately than correlation tables. Here is what we added at the end of the "Methods" section:

Using the tract-level model as an example, the highest Variance Inflation Factor (VIF) was 2.57 for percent of Hispanics in the neighborhoods, and the highest condition index was 3.58, well within the acceptable range for multicollinearity diagnosis. For the four food outlet variables, the VIF ranged from 1.169 for healthy grocery stores to 2.50 for full service restaurants.

Minor points:

8. Page 3, Line 50. Include "weight and" in the phrase "self-reported height data".

R. Revised as suggested.

9. Page 7, Line 32 and 37. It's not clear how "mean age of neighborhood housing" is a measure of "land use diversity". This seems a bit of a stretch. Either explain how it is a measure of diversity, or drop the variable from analysis, or simply edit the sentence to eliminate reference to the "3Ds" (though you can still cite the same reference #29) and refer instead to neighborhood characteristics including those associated with walkability.

R: Instead of stating median age of neighborhood housing as a land use diversity measure, we have revised the relevant sentences to describe this variable as one of the measures of neighborhood build characteristics in the "Methods" section:

"...and median age of neighborhood housing because older neighborhoods have typically been

designed with more walkable features such as tree-shaded sidewalks, narrow streets, and mixed land use [17, 32-34]"

10. Page 8, Line 46. Since you don't present results on the pooled sample (both genders together) and there is no reason to present these, just delete mention of it (i.e. delete "the pooled sample of both genders and also stratified models for"). Otherwise it confuses the reader.

R: We have deleted the "pooled sample" reference from the sentence as suggested.

11. Page 8, last paragraph. It is not clear from the description of your regression models whether you included variables for all four retail types (groceries, restaurants, etc) in one model at each scale, or whether you ran separate models for each retail type at each scale. I assumed you included all retail outlet variables in the same model (for each geog scale), because you talk about collinearity diagnostics. But either way, you should make this explicit.

R: We added the sentence "All four retail outlet types were included in each model at each scale" in the "Methods" section. We also added more details to the collinearity diagnosis as noted above.

12. Page 10, 11. You have some odd findings – older individuals associated with higher obesity risk, but older median age associated with lower risk. Also higher intersection density is associated with higher level of risk, but higher population density is associated with lower. Have you tried dropping some of these redundant variables? For example, why have median neighborhood age, if you have individual's age? Is my weight status affected by the age of my neighbors? Also, have you tried dropping the "older housing stock variable" or one of the walkability variables to see if you would get more consistent results, i.e. increased walkability leads to lower risk of obesity?

R. We have dropped the variable "median residents' age" and re-estimated all models. We added a paragraph in the "Discussion" section to address the street intersection density finding: However, the finding that higher street intersection density was linked to higher obesity risk is counter to much of the existing literature but consistent with other studies using Salt Lake County data [17, 26, 34]. It may be that greater street connectivity reflects greater car traffic, which could discourage individuals from walking or biking. However, it could also be related to Salt Lake County having unusually wide streets and large street blocks that may discourage walking even when street connectivity is high [17, 42].

13. Page 12, Lines 30-40. You should include reference in your discussion to the following citation, since it is relevant to this section.

Bodor JN, Farley TA, Rice JC, Swalm C, Rose D. "The Association between Obesity and Urban Food Environments," *Journal of Urban Health* 2010;87:771-781.

14. Page 13, Lines 15-22. Same comment as above.

R: Thank you for bringing this reference to our attention. We added this reference in what we believe to be the most appropriate place in the "Background" section.

15. Figure 1 – Needs a bit more clarity. On the legend, when printed in black and white, the 1km buffer and census tract look the same. Consider making the 1km buffer a dotted line.

R: We darkened the 1km circular buffer area so it can be differentiated from the tract color when printed in black and white. We have the colored version in the manuscript but the black-white version is shown here:

16. Table 2 – Why present population-weighted counts if it this wasn't used in the regression? Tract delineation is roughly determined by population size, so it's not really necessary.

R: What we did was to estimate the descriptive statistics of these area measures across our sample of individuals. In that sense, all our area measures were population-weighted. We dropped the phrase "population-weighted" to avoid confusion. We believe that the norm in multi-level models is to present

descriptive statistics across the sample of individuals. However, if you feel that we should present these area measures with these areas as units of analysis, we will be happy to do so.

17. Table 5 – Grocery stores not really significant for men at zip-code level, so make first row cell contents for men the same as for women.

R: Revised as suggested. We have updated Table 5 using $p < 0.05$ as the critical p-value with our updated estimates.

Again, we appreciate the opportunity to revise our manuscript. We feel that our revised manuscript addresses the major concerns of the reviewers. We look forward to hearing from you regarding the final decision on this manuscript.

VERSION 2 – REVIEW

REVIEWER	Diego Rose Tulane University, USA
REVIEW RETURNED	18-Jul-2014

GENERAL COMMENTS	The authors have addressed all of my concerns. This paper is ready to be published.
-------------------------	---