Obesity in Urban Food Markets
Evidence from georeferenced micro data

Susan E. Chen\textsuperscript{1}, Raymond J.G.M. Florax\textsuperscript{1,2} and Samantha D. Snyder\textsuperscript{1}

\textsuperscript{1} Dept. of Agricultural Economics, Purdue University
Space, Health and Population Economics (SHAPE) Program
West Lafayette, USA

\textsuperscript{2} Dept. of Spatial Economics, VU University Amsterdam
Amsterdam, The Netherlands

rflorax@purdue.edu
http://web.ics.purdue.edu/~rflorax/
Outline

- Introduction
  - obesity epidemic
  - potential causes

- Theory and data
  - assessment empirical literature
  - health production function
  - data, descriptive statistics, and maps

- Methods and regression results
  - OLS and diagnostics
  - random neighborhood effects
  - spatial ARAR model

- Policy experiments

- Conclusions
Introduction
Obesity epidemic in the US

- BMI = kg/m²

<table>
<thead>
<tr>
<th>BMI</th>
<th>&lt; 18.5</th>
<th>18.5–25</th>
<th>25–30</th>
<th>&gt; 30</th>
</tr>
</thead>
<tbody>
<tr>
<td>kg (1.75m)</td>
<td>&lt; 57</td>
<td>57–77</td>
<td>77–92</td>
<td>&gt; 92</td>
</tr>
</tbody>
</table>

- United States, between 1980 and 2000
  - obesity rates +27% for men, +48% for women
  - rates of overweight +166% for men, +62% for women
  - disparities in prevalence across groups

- Indiana, in 2005
  - 35.1% overweight,
  - 27.2% obese
  - 10th in country

- Implications
  - health care costs
Spatial topology of urban food markets

• Potential causes for epidemic
  – value of time
  – sedentary lifestyle
  – lower cost food away from home
  – better transportation
  – # per capita restaurants
  – ...
  – stricter non-smoking laws
  – lower savings

• Many issues left
  – inconsistencies in a new field
  – data issues

• Focus of this research
  – obesity in urban food markets
  – is there a causal relationship between (lack of) access to fast food outlets and large grocery stores and obesity?
Assessment state-of-the-art

- Main findings literature
  - minority and low income populations differential access to grocers
  - availability of grocery stores associated with healthier diets, lower rates of obesity
  - effect of access to fast food less clear

- Selection effects complicate causal interpretation
  - people select into neighborhoods based on neighborhood characteristics (spatial sorting)
  - stores and food outlets locate in neighborhoods where they will make a profit (location and spatial competition)

- Spatial dependence should be incorporated
  - (unobserved) social and environmental factors affect people’s health behavior
  - spatial data!

- Danger of ecological fallacy
  - due to use of aggregate data
Behavioral model
Health production function

- **Micro-economic setup**
  - major role for technological change (Cutler et al. 2003)
  - implications for real value of time
  - and hence for preference food away from home and pre-prepared highly processed high-caloric foods at home

- **Formally**
  - $h = f(F,C)$
  - $\max U(h,C)$ subject to $pF + C = I$
  - assuming interior solution, first order conditions
    - $\lambda_1 p - \lambda_2 f_F = 0$
    - $U_C + \lambda_1 - \lambda_2 f_C = 0$
    - $-I + pF + C = 0$
    - $-f(F,C) + h = 0$
  - substitution leads to
    - $-f(F,I - pF) + f(F,I - pF) = 0$
  - optimization results in the reduced form demand equations
    - $F^* = g(I,p)$; $C^* = h(I,p)$; and $h^* = j(I,p)$
  - note, vector $p$ is function of price of food and travel cost
Empirical framework

- Health (BMI) function of prices (incl. travel cost)
  \[ h_{ij} = P_{ij}y_1 + \varepsilon_{ij} \]

- Selection effect
  - spatial sorting into neighborhoods
  \[ h_{ij} = P_{ij}y_1 + X_{ij}\beta + N_{ij}y_2 + \varepsilon_{ij} \]
  - unobservable individual and neighborhood effects still create bias
  - so, instrument \( P_{ij} \)

- Account for spatial effects
  - neighborhood heterogeneity with random effects, or
  - (unobserved) social and environmental factors causing spatial correlation
  - spatial ARAR model
  \[
  h = \lambda Wh + P\gamma_1 + X\beta + N\gamma_2 + \varepsilon, \quad \varepsilon = \rho W\varepsilon + \mu
  \]
Data
Case study for Indianapolis, Marion County
Census tracts
Median family income
Data

- Marion County Obesity Needs Assessment Survey
  - information on demographics, behavior, health
  - 3605 adults, including (x,y) coordinates, in 2005

- Marion County Health Department sanitation inspection records
  - selected and classified large grocery stores and fast food restaurants, with (x,y) coordinates of retailers, in 2005

- Indiana Spatial Data Portal
  - zoning maps, for 2004

- Indianapolis Metropolitan Police Department
  - Crime Reporting Data for 2007, with (x,y) coordinates of crimes
## Descriptive stats

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonwhite</td>
<td>0.30</td>
<td>0.46</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Female</td>
<td>0.58</td>
<td>0.49</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Age</td>
<td>46.99</td>
<td>14.22</td>
<td>21.00</td>
<td>75.00</td>
</tr>
<tr>
<td>Less than 200% of the FPL</td>
<td>0.20</td>
<td>0.40</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>More than high school</td>
<td>0.65</td>
<td>0.48</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Vigorous physical activity per week</td>
<td>2.68</td>
<td>2.34</td>
<td>0.00</td>
<td>7.00</td>
</tr>
<tr>
<td>Physically demanding job</td>
<td>0.41</td>
<td>0.49</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>smoker</td>
<td>0.26</td>
<td>0.44</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Body Mass Index</td>
<td>27.63</td>
<td>6.11</td>
<td>10.02</td>
<td>61.99</td>
</tr>
<tr>
<td>Obese</td>
<td>0.27</td>
<td>0.44</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>No. large groceries (1/2 mile radius)</td>
<td>0.34</td>
<td>0.66</td>
<td>0.00</td>
<td>4.00</td>
</tr>
<tr>
<td>No. fast food (1/2 mile radius)</td>
<td>1.77</td>
<td>2.69</td>
<td>0.00</td>
<td>18.00</td>
</tr>
<tr>
<td>Percent zoned non-residential</td>
<td>0.32</td>
<td>0.21</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>No. serious crimes (1/2 mile radius)</td>
<td>27.11</td>
<td>30.50</td>
<td>0.00</td>
<td>222.00</td>
</tr>
</tbody>
</table>

N = 3605
Healthy, overweight, obese

BMI
- healthy
- overweight
- obese

Median family income
- 18625 - 33750
- 33751 - 45769
- 45770 - 58684
- 58685 - 78990
- 78991 - 140217
BMI, Thiessen polygons
Hot spots of low and high BMI individuals
Local food environment, high-income suburb

BMI
- healthy
- overweight
- obese
- Chain store
- Fast food outlet

Median family income
- 18625 - 33750
- 33751 - 45769
- 45770 - 58684
- 58685 - 78990
- 78991 - 140217
Local food environment, down town neighborhood
Methods and regression results
## OLS results

<table>
<thead>
<tr>
<th></th>
<th>OLS(1)</th>
<th>OLS(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. fast food (1/2 mile radius)</td>
<td>0.055</td>
<td>0.052</td>
</tr>
<tr>
<td>No. large groceries (1/2 mile radius)</td>
<td>-0.303*</td>
<td>-0.159</td>
</tr>
<tr>
<td>Nonwhite</td>
<td></td>
<td>0.972**</td>
</tr>
<tr>
<td>Female</td>
<td>-0.343*</td>
<td>0.404**</td>
</tr>
<tr>
<td>Age</td>
<td>0.404**</td>
<td>0.004**</td>
</tr>
<tr>
<td>Age$^2$</td>
<td></td>
<td>-0.004**</td>
</tr>
<tr>
<td>Less than 200% of the FPL</td>
<td></td>
<td>1.158**</td>
</tr>
<tr>
<td>More than high school</td>
<td></td>
<td>-0.854**</td>
</tr>
<tr>
<td>Vigorous physical activity work</td>
<td></td>
<td>-0.320**</td>
</tr>
<tr>
<td>Physically demanding job</td>
<td></td>
<td>-0.603**</td>
</tr>
<tr>
<td>Smoker</td>
<td></td>
<td>-1.339**</td>
</tr>
<tr>
<td>No. serious crimes (1/2 mile radius)</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>27.669**</td>
<td>19.777**</td>
</tr>
<tr>
<td>R$^2$</td>
<td>0.0002</td>
<td>0.064</td>
</tr>
<tr>
<td>N</td>
<td>3550</td>
<td>3550</td>
</tr>
</tbody>
</table>

* 0.05 < p < 0.10, ** p < 0.05, White-adjusted standard errors
Modeling setup

- **“Standard” approach**
  - use random effects at the census tract level
  - non-zero covariances in variance-covariance matrix
  - does not model spatial dependence

- **Selection effects**
  - spatial sorting across neighborhoods of individuals
  - location choice and spatial competition food retailers

- **Instrumental variables**
  - potentially endogenous variables
    - fast food access and grocery store access
  - grocery stores: people drive, low marginal cost distance
  - fast food outlets: local consumption, high marginal cost of distance, selection criterion for location

- **Instrument for fast food**
  - percent of non residential zoning within 1/2 mile radius
Spatial heterogeneity and dependence

- Spatial ARAR model allowing for heteroskedasticity
  - recall, \( h = \lambda Wh + \gamma_1 + X\beta + N_{\gamma_2} + \epsilon, \epsilon = \rho \epsilon + \mu \)
  - instruments needed for \( Wh \) and fast food outlets
  - spatial generalized two stage least squares estimator
  - multi-round IV/GM approach (Kelejian and Prucha, 2007)

- Regression results
  - strong social network ties among individuals and
  - shared unobserved neighborhood characteristics across individuals living in proximate neighborhoods
## Results including ARAR specification

<table>
<thead>
<tr>
<th></th>
<th>OLS(1)</th>
<th>OLS(2)</th>
<th>ARAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. fast food (1/2 mile radius)</td>
<td>0.055</td>
<td>0.052</td>
<td>0.201**</td>
</tr>
<tr>
<td>No. large groceries (1/2 mile radius)</td>
<td>-0.303*</td>
<td>-0.159</td>
<td>-0.481*</td>
</tr>
<tr>
<td>Nonwhite</td>
<td>0.972**</td>
<td>0.506**</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-0.343*</td>
<td>-0.336*</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.404**</td>
<td>0.393**</td>
<td></td>
</tr>
<tr>
<td>Age²</td>
<td>-0.004**</td>
<td>-0.004**</td>
<td></td>
</tr>
<tr>
<td>Less than 200% of the FPL</td>
<td>1.158**</td>
<td>1.059**</td>
<td></td>
</tr>
<tr>
<td>More than high school</td>
<td>-0.854**</td>
<td>-0.641**</td>
<td></td>
</tr>
<tr>
<td>Vigorous physical activity work</td>
<td>-0.320**</td>
<td>-0.302**</td>
<td></td>
</tr>
<tr>
<td>Physically demanding job</td>
<td>-0.603**</td>
<td>-0.576**</td>
<td></td>
</tr>
<tr>
<td>Smoker</td>
<td>-1.339**</td>
<td>-1.368**</td>
<td></td>
</tr>
<tr>
<td>No. serious crimes (1/2 mile radius)</td>
<td>0.003</td>
<td>-0.001</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>27.669**</td>
<td>19.777**</td>
<td>-0.821</td>
</tr>
</tbody>
</table>

\[ \chi (\text{Wh}) \]
\[ \rho (\text{Ws}) \]

* 0.05 < p < 0.10, ** p < 0.05, White-adjusted standard errors for OLS
Policy experiments
Good and bad scenario

• Policy experiments
  – based on (1 x 1) km over-layed raster grids
  – with socio-economic census data
  – procedure
    ▪ determine affected individuals
    ▪ recalculate individual local food landscape

• “Bad” scenario
  – fast food
  – reduce access with 1 in areas with >6 fast food outlets per km²
  – 563 fast food restaurants reduced by 15
  – 178 people directly affected

• “Good” scenario
  – grocery stores
  – increase access with 1 in areas with >40% below FPL and >40% with less than high school
  – 94 grocery stores increased by 13
  – 74 people directly affected
Spatial simulation

- Spatial model
  - includes spatial heterogeneity and spatial dependence
  - policy effect becomes location dependent

\[
h = \lambda Wh + X\beta + \gamma_1 + N\gamma_2 + \epsilon, \quad \epsilon = \rho W\epsilon + \mu
\]

\[
= (I-\lambda W)^{-1} [X\beta + \gamma_1 + N\gamma_2 + (I-\rho W)^{-1} \mu]
\]

- Spatial multiplier
  - induces spillover effects

\[
(I-\lambda W)^{-1} = I + \lambda W + \lambda^2 W^2 + \lambda^3 W^3 + ...
\]
## Average policy effects

### Bad scenario

<table>
<thead>
<tr>
<th></th>
<th>Direct</th>
<th>Indirect</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-0.011*</td>
<td>-0.034</td>
<td>-0.045</td>
</tr>
<tr>
<td>Standard error</td>
<td>0.000</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>z-value / 2 mile radius</td>
<td>-44.409</td>
<td>-27.302</td>
<td>-31.875</td>
</tr>
</tbody>
</table>

* Directly affected 178 out of 3,550 by -0.22 BMI (= 0.67 kg)

### Good scenario

<table>
<thead>
<tr>
<th></th>
<th>Direct</th>
<th>Indirect</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-0.012*</td>
<td>-0.042</td>
<td>-0.054</td>
</tr>
<tr>
<td>Standard error</td>
<td>0.001</td>
<td>0.004</td>
<td>0.005</td>
</tr>
<tr>
<td>z-value / 2 mile radius</td>
<td>-17.080</td>
<td>-9.961</td>
<td>-11.483</td>
</tr>
</tbody>
</table>

* Directly affected 74 out of 3,550 by -0.58 BMI (= 1.78 kg)
Total policy effects
Conclusions
Results and implications

- Stronger test of environmental impacts on obesity
  - micro data
  - spatially explicit, network and neighborhood effects
  - selection effect fast food outlets and individual location
  - note, for urban food markets

- Impact
  - positive for fast food
  - negative for grocery stores
  - grocery store effect 2.5 times as big as fast food effect

- Policy implications
  - health zoning
    - LA, two-year moratorium on fast food restaurants
    - increase availability nutritious foods
    - restrictions on store licenses
  - policy design and impact
    - selection target area
    - demographic and neighborhood criteria