USDA United States Department of Agriculture

Food and Nutrition Service, Office of Policy Support **October 2014**

Evaluation of the Healthy Incentives Pilot (HIP): Spatial Analysis

Authors:
Todd Grindal
Gabe Schwartz
Jacob Klerman
Susan Bartlett
Abt Associates

Parke Wilde
Friedman School of Nutrition Science and Policy, Tufts University

Submitted by:
Abt Associates Inc.
55 Wheeler Street
Cambridge MA 02138

Submitted to:
Office of Policy Support
Food and Nutrition Service
3101 Park Center Drive
Alexandria, VA 22302-1500

Project Director:
Susan Bartlett

Project Officer:
Danielle Berman

This study was conducted under Contract number AG-3198-D-10-0044 with the Food and Nutrition Service, United States Department of Agriculture.

This report is available on the Food and Nutrition Service website: http://www.fns.usda.gov

Suggested Citation:
Grindal, Todd, Gabe Schwartz, Jacob Klerman, et al. *Evaluation of the Healthy Incentives Pilot (HIP): Spatial Analysis*. Prepared by Abt Associates for the U.S. Department of Agriculture, Food and Nutrition Service, October 2014.

Contents

Executive Summary

The Food, Conservation, and Energy Act of 2008 authorized funds for pilot projects to determine if financial incentives provided at the point of sale to Supplemental Nutrition Assistance Program (SNAP) participants would increase their consumption of fruits, vegetables, and other healthful foods. On the basis of this legislative authority, USDA's Food and Nutrition Service (FNS) designed the Healthy Incentives Pilot (HIP) to investigate the impact of making fruits and vegetables more affordable for SNAP participants.

Under HIP, SNAP participants received on their SNAP Electronic Benefit Transfer (EBT) card an incentive of 30 cents for every dollar of SNAP benefits that they spent on targeted fruits and vegetables (TFVs) in participating retailers. TFVs included fresh, canned, frozen, and dried fruits and vegetables without added sugars, fats, oils or salt (with some exceptions), but excluded white potatoes, mature legumes, and 100% fruit juice (the same set of fruits and vegetables eligible for the WIC Fruit and Vegetable Cash Value Voucher). The incentive was capped at $60 per household per month to prevent misuse and ensure that total incentive payments would not exceed $2 million.

The impacts of HIP were evaluated using a random assignment research design in which approximately 15 percent of Hampden County SNAP households were randomly selected to be eligible to earn incentives.

Retailer participation in HIP was voluntary and not all area retailers participated. Participating retailers accounted for approximately 60 percent of total Hampden county SNAP redemptions. HIP participants only earned incentives for fruit and vegetable expenditures in participating stores and only for purchases using SNAP. Fruit and vegetable purchases made using other forms of payment, such as cash or WIC vouchers did not earn incentives.

The pilot was implemented by the Massachusetts Department of Transitional Assistance (DTA) in Hampden County beginning in late 2011 and continuing through the end of 2012. Abt Associates, Inc. and its partners, Westat and MAXIMUS, conducted the evaluation of HIP for the USDA Food and Nutrition Service. The results of this evaluation (detailed in Bartlett, Klerman, et al., 2014), indicate that those SNAP participants randomly assigned to be eligible to earn HIP incentives purchased and consumed significantly more fruits and vegetables than did SNAP participants who were not selected for the pilot.

The *HIP Spatial Analysis* is a follow-on study that investigates the extent to which the limited availability of HIP retailers might have suppressed qualifying purchases and therefore the impact of the incentive on consumption. This study also examines how the food retail environment more broadly defined might have influenced Hampden County SNAP participants' fruit and vegetable purchases in general and the HIP impact estimates in particular.

Evaluation Objectives

The *HIP Spatial Analysis* addresses two research questions:

1. How did relative physical access to HIP participating stores affect household shopping patterns, HIP incentive earnings, and targeted fruit and vegetable (TFV) consumption?

2. Are there "neighborhood effects" with respect to consumer responses to the incentive? Did HIP participants in close proximity to one another, and thereby with shared food retail environment characteristics, exhibit similarities in response to the pilot?

Design and Data Collection

Evaluation Design

To compare food intake and other outcomes for HIP participants relative to outcomes for otherwise similar non-participants, the HIP evaluation used a random assignment research design, which is widely viewed as providing the strongest evidence of causal impact. Of the 55,095 SNAP households in Hampden County, 7,500 households (the "HIP group") were randomly assigned to participate in HIP, while the remaining 47,595 SNAP participating households (the "non-HIP group") continued to receive SNAP benefits as usual. Thus, comparisons between participating and non-participating households provide a reliable estimate of the impact of HIP on households receiving SNAP benefits.

A random subsample of approximately 5,000 households, equally divided between the HIP and non-HIP groups, was selected to participate in survey data collection.

Data Collection

Several types of data collected for the HIP evaluation were used in the spatial analysis. To collect the necessary data on dietary intake, trained telephone interviewers conducted 24-hour dietary recall interviews, a widely used, reliable methodology. Respondents were also asked about their attitudes and preferences for fruits and vegetables, shopping patterns, food expenditures, and household characteristics, including address of their primary residence.

The data collection for the HIP evaluation included three rounds of participant surveys. Round 1 was conducted before HIP was implemented. Rounds 2 and 3 were conducted during HIP, one fielded 4 to 6 months after implementation, and the other fielded 9 to11 months after implementation. Both Round 2 and Round 3 surveys collected information on dietary intake using 24-hour dietary recall interviews.

Several types of extant data were also used for the spatial analysis. SNAP caseload data provided information on household characteristics and residential address. EBT transaction data provided detailed information on households' SNAP EBT purchases, including HIP-eligible purchases and the incentive amounts earned. EBT data have a number of restrictions that affect the analyses that can be performed. First, only purchases made with SNAP benefits are included; purchases made with other forms of payment, such as cash or WIC vouchers are not captured. Second, only HIP purchases at HIP participating stores earned the incentive; purchases of HIP-eligible items at non-participating retailers did not earn the incentive. Finally, at HIP participating stores that did not use integrated electronic cash registers (IECRs), HIP households needed to identify themselves as HIP participants and HIP-eligible items needed to be separated from other items. HIP incentives were only earned for properly identified purchases.

DTA Retailer EBT Data Exchange (REDE) files provided information on retailer location and type. Public use files from the 2010 U.S. Census and the American Community Survey provided information on community demographics. Finally, the U.S. Census Bureau's Topologically Integrated Geographic Encoding and Referencing files and state-specific files found on the Commonwealth of Massachusetts' MassGIS database provided information on roads and geographic boundaries.

Spatial Analysis

This study presents three types of analysis. First, we use maps and tabulations to describe the food retailer environment in Hampden County, noting the most heavily used retailers and identifying areas with limited food retailer access. Second, we use regression analyses to estimate the relationship between the distance to food retailers on food spending and food intake outcomes for SNAP participants. Third, we use spatial autocorrelation and spatial regression models to determine whether ‘neighborhood effects’ may have influenced key spending and dietary intake outcomes.

Findings

SNAP Household Access to Food Retailers

Nearly all (94 percent) Hampden County SNAP participants lived within one mile of a SNAP retailer. Approximately two-thirds (65 percent) lived within one mile of a supermarket, and roughly one quarter (26 percent) lived within one mile of a supermarket that participated in HIP. Access to different types of food retailers varied across the county, with a retailer density higher in the more urban central region than in the less densely populated eastern and western portions. The analysis did not find a consistent relationship between neighborhood poverty and access to supermarkets.

The analyses explored both the type of retailers where SNAP participants used their benefits and the distances between SNAP participants’ homes and those retailers. Hampden County SNAP households spent the bulk of their benefits (78 percent) at supermarkets and superstores and typically did all of their SNAP shopping at between one and three stores per month. This preference for supermarkets was consistent regardless of whether or not the families lived within a mile of a supermarket, but it differed systematically based on other household characteristics, such as income and race/ethnicity.

Households typically did not shop in the supermarket that was closest to their residence. Of all the benefits SNAP households spent in supermarkets, only 16 percent were spent in the supermarket that was closest to their home. Although Hampden County SNAP participants lived an average of 1.02 miles from the nearest supermarket, they spent the largest portion of their monthly household benefits at retailers that were an average of 3.22 miles from where they lived. Moreover, the majority of SNAP benefits were spent in just 13 individual high-volume retailer locations, which were patronized by SNAP participants residing both nearby and farther away.

The Role of Retailer Access in Fruit and Vegetable Spending and Intake

We found no evidence that distance to food retailers affected the *impact* of HIP on fruit and vegetable spending in HIP participating retailers or on the consumption of fruits and vegetables among SNAP households. This finding was true regardless of the measurement approach or definitions of food retailer access. We did find some evidence that a SNAP household’s distance to participating supermarkets was associated with its *level* of TFV spending in participating supermarkets. Households that lived farther away from a HIP participating supermarket, on average, spent fewer benefit dollars on targeted fruits and vegetables in participating supermarkets than did households that lived in closer proximity to these retailers. All else being equal, every additional mile that a SNAP household lived from the nearest HIP participating supermarket was associated with $0.69 less spending on targeted fruits and vegetables. This finding is limited to purchases made in participating retailers. We do not have data on SNAP participants’ purchases in non-participating retailers or purchases made using some other form of payment.

Neighborhood Effects

We did find evidence of "neighborhood effects" on households' spending on fruits and vegetables. Households that lived nearer to one another were more similar in their purchases of fruits and vegetables than they were to households that lived farther away. The ordinary least squares regression analyses used in the main HIP report accounted for some but not all of these neighborhood effects. That said, correcting for this shared variation using spatial regression analyses did not substantively change the estimated impact of HIP on spending, the estimated precision of that impact, or the inferences about statistical significance.

Conclusions

We find no evidence that distance to food retailers moderated the impact of the HIP incentive. That is, the difference in both spending on and intake of targeted fruits and vegetables for the HIP group compared to the non-HIP group was not significantly larger among those households that lived closer to retailers where it was possible to earn incentives. Although our analysis does show some relation between food retailer access food expenditure patterns, we also found that SNAP households in Hampden County usually traveled beyond their nearest supermarket to spend their SNAP benefits.

1. Introduction

The Healthy Incentives Pilot (HIP) investigated the impact of making fruits and vegetables more affordable for participants in the Supplemental Nutrition Assistance Program (SNAP). The Food, Conservation, and Energy Act of 2008 authorized funds for pilot projects to determine if financial incentives provided to SNAP participants at the point of sale increase the consumption of fruits, vegetables, or other healthy foods. On the basis of this legislative authority, the Food and Nutrition Service (FNS) of the U.S. Department of Agriculture (USDA) designed HIP.

Under HIP, SNAP participants received on their SNAP Electronic Benefit Transfer (EBT) card an incentive of 30 cents for every dollar of SNAP benefits that they spent on targeted fruits and vegetables (TFVs) at participating retailers. TFVs included fresh, canned, frozen, and dried fruits and vegetables without added sugars, fats, oils or salt, but excluded white potatoes, 100% fruit juice, and mature legumes (the same set of fruits and vegetables eligible for the Women, Infants, and Children (WIC) Fruit and Vegetable Cash Value Voucher). The incentive was capped at $60 per household per month to prevent misuse and ensure that total incentive payments would not exceed $2 million. The cap did not appear to constrain households, as very few households reached it.

The pilot was implemented by the Massachusetts Department of Transitional Assistance (DTA) in Hampden County. Located in western Massachusetts, the county is a mix of urban, rural, and suburban areas with 55,095 SNAP households in July 2011. Hampden County has the lowest median household income in the State. Massachusetts, like the rest of the country, is in the midst of an obesity epidemic, and residents in the western region have the highest rates of obesity and related chronic illnesses in the State.

1.1 HIP Evaluation

HIP was evaluated using a rigorous research design. The 55,095 SNAP households in Hampden County that were active in July 2011 were randomly assigned to the HIP group (7,500 households) and the non-HIP group (47,595 households). The HIP households were divided into three waves of 2,500 households each to begin the pilot successively during the first three months of operation. The pilot operated from November 2011 to December 2012. HIP participants were eligible to earn incentives for 12 months.

The HIP evaluation used participant surveys to measure HIP’s impact on targeted fruit and vegetable intake (the main confirmatory outcome). HIP participants’ average daily TFV intake was 0.24 of a cup-equivalent higher (26 percent higher) than the control group (Bartlett et al., 2014). The HIP impact partly reduced the difference between current consumption and the *Healthy People 2020* objectives for total fruit and vegetable intake—narrowing the gap by approximately 18 percent. Thus, HIP had a significant positive impact on fruit and vegetable consumption.

In addition to the participant surveys, the HIP evaluation also used EBT transaction data to explore HIP’s impact on food spending in participating retailers. One would expect the overall effect of HIP on food spending to be larger if participants responded more strongly to the incentive, purchased more fruits and vegetables, and earned more incentives. On average, during March to October 2012, HIP households spent just over $12 on targeted fruits and vegetables in participating stores (representing 5 percent of their SNAP benefits), earning an average incentive of $3.65 each month. HIP households spent more than non-HIP households on TFVs in participating supermarkets—non-

HIP households spent $10.86 each month using their EBT card on TFVs and HIP households spent $12.05. This represented an increase in monthly TFV spending of $1.19 or 11 percent. Although this impact was positive and statistically significant, it was smaller than expected at the start of the pilot. This raises the possibility that limited availability of HIP retailers—in general or of a specific type — might have suppressed qualifying purchases and therefore the impact of the incentives.

The *HIP Spatial Analysis* is a follow-on study that further investigates how the food retail environment influenced Hampden County SNAP participants' fruit and vegetable purchases in general and the HIP impact estimates in particular. This study connects two lines of research in contemporary U.S. nutrition policy: (1) research on the use of price incentives to promote intake of healthful foods and (2) research on how the quality of the food retail environment in low-income neighborhoods influences food choices. This report addresses local conditions that may have modified the HIP impact and more broadly offers new insight into the features of the local food environment that most strongly affect food choices for low-income Americans.

1.2 Research Questions and Objectives

The HIP Spatial Analysis addresses two research questions posed by USDA's Food and Nutrition Service (FNS):

1. How did relative physical access to HIP participating stores affect household shopping patterns, HIP incentive earnings, and targeted fruit and vegetable (TFV) consumption?
2. Are there "neighborhood effects" with respect to consumer responses to the incentive? Did HIP participants in close proximity to one another (and thereby with shared food retail environment characteristics) exhibit similarities in response to the pilot?

To answer these questions, the analysis does the following:

- Characterizes the relative physical access to nearby HIP participating and non-participating stores, including stores of various types, for all block groups in Hampden County.
- Characterizes the food retailers actually patronized by SNAP participants, whether they were nearby or farther away and whether they participated in HIP or not.
- Measures the effect of food retailer access on eligible TFV purchases and on HIP/non-HIP differences in eligible TFV purchases.
- Estimates the extent to which the impact of HIP varied with proximity to HIP participating stores (and the relative proximity of HIP participating and non-participating stores).
- Investigates to what extent TFV intake was correlated for SNAP participants who lived near each other and to what extent that correlation was induced by the (measured) common food retail environment.

1.3 Previous Research

A large body of literature has investigated the lack of access to healthy foods in low-income neighborhoods. In principle, lack of food retail access could be cause for concern. Indeed, as USDA noted in a 2012 report:

> [F]or some Americans and in some communities, access to affordable healthy foods may be limited. If healthy foods can be obtained only with great effort, those affected by poor access may have poorer diets and higher rates of diet-related disease, such as obesity and diabetes (Ver Ploeg et al., 2012).

The evidence is mixed on these issues. As this section describes, some low-income neighborhoods do lack easy access to a supermarket, while many other low-income neighborhoods have a supermarket nearby. Similarly, some previous research found that distance from a supermarket harms dietary quality (IOM and NRC, 2013; Larson, Story, and Nelson, 2009; Rose and Richards, 2004), while other research concluded that the larger problem is an abundance of comparatively unhealthy food sources rather than a lack of healthy food sources (Block et al., 2011; Currie et al, 2010). Thus, it seems appropriate to approach the analysis of the local food environment as an open empirical question, with few advance assumptions.

The nature of the food retail environment in the United States varies greatly across neighborhoods and across households. At the national level, USDA research identified and counted low-income and higher-income areas that were located more than one mile from a supermarket in 2010. They found:

- In low-income areas, 30.8 percent of the population (29.7 million people) lived more than one mile from a supermarket.
- In moderate- and high-income areas, 45.9 percent of the population (96.6 million people) lived more than one mile from a supermarket (Ver Ploeg et al., 2012).[1]

It is initially surprising—yet with some reflection understandable—that low-income neighborhoods were more likely than other neighborhoods to have a nearby supermarket. Low-income neighborhoods in the United States have higher-than-average population density, and higher-density neighborhoods in turn are more likely to have a nearby supermarket (Wilde, Llobrera, and Ver Ploeg, 2014). In addition, most low-income people in the United States shop at supermarkets and are reasonably satisfied with their level of food retail access (USDA, 2009).

At the local level, within a particular neighborhood, food retail access varies depending on whether households have a car. Lack of a vehicle may compromise a household's ability to access healthy food, especially if the household does not have a nearby supermarket. USDA estimated that just 1.8 percent of U.S. households in 2010 both lacked a vehicle and resided more than one mile from a supermarket (Ver Ploeg et al., 2012).

In the public health and nutrition literatures, a large body of research has reported an association between better food retailer access and improved dietary quality or lower rates of overweight and obesity (reviewed in Larson, Story, and Nelson, 2009), but some research finds otherwise, and it is difficult to determine cause and effect from most of these studies (Wilde, Llobrera, and Valpiani, 2013). In a sample of SNAP participants, Rose and Richards (2004) studied the association between fruit and vegetable consumption and distance from the retailer where a respondent purchased most food. Greater distance was associated with significantly lower fruit consumption in the sample (and

---

[1] Low-income areas were defined as those where more than 40 percent of the population had income at or below 200 percent of the Federal poverty threshold; these areas contained 31.4 percent of the U.S. population.

with lower vegetable consumption but the vegetable results were not statistically significant). Other research has found an association between the presence of quick-service restaurants and body mass index (BMI) or weight status (Block et al., 2011), and some of this research argues for a cause-and-effect interpretation of this association (Currie et al., 2010).

Because HIP is directly connected to SNAP, we are particularly interested in previous research on changes to nutrition assistance programs that could encourage food retailers serving low-income areas to supply more fruits and vegetables. With data from Connecticut, before and after the implementation of new WIC packages that included a voucher for fruit and vegetable purchase, Andreyeva et al. (2012) found increases in a basket of healthy foods (heavily weighted toward fruits and vegetables) supplied by food retailers in low-income neighborhoods.

1.4 Previous HIP Findings on the Retail Environment

The HIP Spatial Analysis provides a detailed study of the food retail environment in Hampden County during the Healthy Incentives Pilot. This report builds on related but more preliminary information and insight than the information included in the Final Report from the HIP evaluation (Bartlett et al., 2014).

The conceptual framework for the HIP evaluation recognized the role of the food retail environment (including inventory, prices, and promotion) and retailer redemptions (including electronic benefit transfer spending in general and TFV purchases in particular). Food expenditures and shopping patterns were described as an important intermediate pathway by which HIP might influence actual food consumption.

The random assignment of SNAP participants to HIP and non-HIP groups and the random sampling for the accompanying participant survey both included a stratifying variable that distinguished three sections of the county: Springfield (the largest city), Chicopee/Holyoke (the second- and third-largest cities, which are adjacent to each other), and the balance of Hampden County. Mean monthly household SNAP purchases, mean HIP purchases of qualifying TFVs in participating retailers, and mean incentives earned all appeared fairly similar across the three sections of the county.

The HIP evaluation considered whether participating retailers offered adequate fruit and vegetable inventories. Not surprisingly, the study found that almost all supermarkets carried a wide variety of fresh, canned, and frozen fruits and vegetables. Grocery store inventories were less extensive, though over half of the grocery stores carried fairly extensive selections of fresh and canned fruits and vegetables. Frozen fruits and vegetables were much less common, however. Also, as expected, convenience stores carried relatively few fruits and vegetables; when they did carry fruits and vegetables, they were primarily fresh fruits.

Some results in the HIP evaluation raised the possibility that lack of access to participating retailers could have reduced the impact of the pilot. The retailer recruitment strategy recognized the value of having as many SNAP authorized retailers as possible participate, but recruitment turned out to be one of the two biggest implementation challenges. In November 2011, as the pilot began, 18 out of 45 eligible supermarket and superstore locations participated in HIP. By October 2012, as the pilot was ending, three additional supermarkets were participating. Using EBT data pooled from March to October, 2012—the same period during which the main post-implementation surveys were fielded—the control group spent only 51 percent of all SNAP benefits in HIP participating retailers.

Because of this evidence, the HIP Final Report investigated the hypothesis that HIP impacts might be higher for those SNAP participants who appeared to have better access to the participating retailers even before implementation. The additional investigation distinguished two groups of households: those that had predominantly shopped in participating retailers before HIP implementation and those that had predominantly shopped in non-participating retailers before implementation. HIP caused increased monthly TFV spending in participating supermarkets for both groups, and the difference in the increases was not statistically significant ($1.24 for the first group and $1.70 for the second). Hence, contrary to our hypothesis, lack of access to the participating retailers did not appear to suppress the HIP impact on TFV spending.

The HIP evaluation included other results that could reflect lack of access to participating retailers. EBT transaction data for March to October 2012 showed that 34.4 percent of HIP participants had no qualifying HIP purchases in a given month and therefore earned no incentives. Approximately half of these participants had no SNAP purchases in participating retailers during the specific month[2] (a pattern that could be consistent with lack of retail access); the other half of these participants made some SNAP purchases in participating retailers, just not qualifying fruits and vegetables (a pattern that is inconsistent with lack of retail access). These results leave open the possibility that lack of retail access could have suppressed HIP purchases for some households.

In the HIP evaluation, survey respondents also were asked for their own perceptions of their food retail environment. Reflecting on their reasons for choosing the food retailer that they patronized most frequently, about half of respondents (48.3 percent of HIP participants and 51.2 percent of non-HIP participants) said low cost was a reason, while under a third of respondents (30.9 percent of HIP participants and 32.4 percent of non-HIP participants) said they shopped at a retailer close to home. Listing barriers to consumption of fruits and vegetables, comparatively more respondents mentioned high cost, while comparatively fewer mentioned difficulty finding fruits and vegetables where they shopped.

This evidence from the HIP evaluation suggests that both cost and geographic proximity were likely factors affecting households' choices of food retailers and the acquisition of fruits and vegetables. Based on the previous literature and on the partial evidence from the HIP evaluation's Final Report, the HIP Spatial Analysis Report addresses empirical questions about the food retail environment and its potential effect on HIP that have not previously been answered.

1.5 Organization of the Report

This report is organized as follows:

- Chapter 2 presents data and methods for the spatial analysis.
- Chapter 3 describes the food retailer environment in Hampden County.
- Chapter 4 focuses on SNAP participants' proximity to retailers, including both the nearest retailers and the retailers where they conducted most of their grocery shopping.
- Chapter 5 investigates the associations between food retailer access in different parts of the county and potentially related food spending and food intake outcomes in HIP.

---

2 Few HIP households (4.5 percent) did not earn any incentives at all during the entire pilot.

- Chapter 6 provides a spatial econometric analysis of the effects of the food environment on these food spending outcomes, including attention to any neighborhood effects.
- Chapter 7 offers discussion and conclusions.

2. Design, Data, and Methods

The HIP evaluation used a random assignment research design, widely viewed as providing the strongest evidence of causal impact. The study compared food intake, expenditures, and other outcomes for SNAP households participating in HIP relative to outcomes for otherwise similar SNAP participants who were not eligible for the incentive. Specifically, 7,500 Hampden County SNAP households were randomly assigned to participate in HIP, while the remaining 47,595 households continued to receive SNAP benefits as usual. Thus, comparisons between HIP and non-HIP SNAP households made from survey and administrative data provide reliable estimates of the impact of HIP.

This first section of this chapter provides an overview of the survey and administrative data collected for the evaluation, which was also used in the spatial analysis. More detail on data collection can be found in the HIP Final Report. The spatial analysis required acquiring and processing some additional data; the second section discusses these activities. The third section provides an overview of the spatial analyses conducted for this report.

2.1 Data Collection

A random subsample of approximately 5,000 households, equally divided between the HIP and non-HIP groups, was selected to participate in three rounds of survey data collection. The first round was conducted before HIP implementation (Round 1). Two rounds of participant surveys were conducted during HIP, one survey 4 to 6 months after implementation (Round 2), and the other participant survey 9 to11 months after implementation (Round 3).

Determining impacts on fruit and vegetable consumption, our main outcome of interest, required collection of dietary intake data from participants. The evaluation collected this information using 24-hour dietary recall interviews, a widely used, reliable methodology conducted by trained telephone interviewers to obtain detailed information on all foods and beverages consumed in the previous 24-hour period. Respondents also provided information about their shopping patterns, food expenditures, attitudes, and preferences for fruits and vegetables, and household characteristics. Both the Round 2 and Round 3 surveys collected information on dietary intake using 24-hour dietary recall interviews.

The evaluation also acquired 17 months of EBT transaction data for all 55,095 SNAP households in Hampden County beginning at the end of July 2011 when random assignment to the pilot occurred. These data provided detailed information on households' SNAP EBT transactions, including HIP-eligible purchase amounts and the amount of incentives earned. Limited information on the retailers where purchases were made was also included.

The way HIP worked—where and how targeted fruits and vegetables (TFVs) could be purchased to earn the incentive—and the use of EBT transaction data to capture TFV purchases restricted the analyses in several ways, which are outlined below.

- Only purchases made with SNAP benefits are included; purchases made with other forms of payment, such as cash or WIC vouchers are not captured.
- Only HIP purchases at HIP participating stores earned the incentive; purchases of HIP-eligible items at non-participating retailers did not earn the incentive.

- At HIP participating stores *without* integrated electronic cash registers (IECRs), HIP households needed to identify themselves as HIP participants and HIP-eligible items needed to be separated from other items. HIP incentives were only earned for properly identified purchases.

Finally, DTA provided 17 months of caseload data from its SNAP eligibility system. Data covered all 55,095 SNAP households in Hampden County and included information on household demographic and economic characteristics and household residential location.

2.2 Spatial Analysis Data

The HIP Spatial Analysis required the following three main data acquisition and processing efforts:

- Compiling and organizing data for map construction
- Creating geocoded location files, which recorded the location of each SNAP-authorized retailer and SNAP household
- Connecting the locations of retailers with the locations of the SNAP households, so that travel distances could be computed.

Data for Map Construction

Data for map construction were acquired from Federal, State, and local sources.

Federal

From American Community Survey (ACS) 2007-2011 Five-Year Estimates, we extracted census block group and tract-level data. Data included median household income (for the entire population and for SNAP participants), household vehicle ownership, and the proportion of the population living below the Federal Poverty Line. We used these five-year ACS data, because five-year estimates are more robust than single-year estimates for the small census block groups and tracts that are the focus of this spatial analysis, and the 2008–2012 five-year estimates were not yet available.

State

From the MassGIS database publicly available on mass.gov, we downloaded county, census block group, census tract, and town/city boundaries. These files included information on land area, total population, and total household count, as per the 2010 Census. From this information, we calculated population density, as well as SNAP participant density (when combined with case file location data). In addition to political and statistical area boundary files, we also obtained the Census Bureau's Topologically Integrated Geographic Encoding and Referencing (TIGER) roads data and Massachusetts hydrography shapefiles (such as rivers and lakes) from mass.gov. The TIGER roads dataset provided a major highway shapefile, which we used on reference maps in lieu of a more comprehensive, but less informative, roadway file.

Local

We acquired data on public transit for Hampden and Hampshire Counties from regional planning commissions. These included bus route and bus stop shapefiles.

Exhibit 2.1 describes each data layer and its source.

Exhibit 2.1: Data Layers Available for Maps

Data layer	Description	Source
County Boundaries	MassGIS derived the boundaries of the 14 counties in Massachusetts	MassGIS
Census Blocks, Block Groups, Tracts	Spatial extracts from the Census Bureau's MAF/TIGER database, containing features such as roads, railroads, rivers, as well as political and statistical geographic areas	2010 U.S. Census
TIGER Roads and associated tables	Census geographic areas, or units, are used by the Bureau to collect and tabulate decennial census data	2010 U.S. Census
MassDOT Roads	The official state-maintained street transportation dataset available from MassGIS	MassDOT's Road Inventory
Hampden and Hampshire County Bus Routes and Stops	Bus routes and stops within public transit systems of Hampden and Hampshire Counties	Regional planning commissions

Geocoded Location Files

We built several geocoded participant and retailer location files for the months from August 2011 to December 2012, inclusive (see below). This time period sufficed to allow more than three months before and 12 months after the beginning of HIP implementation for each of the three waves of households. For most analyses, a small subset of these months was used. As noted below, we used selected months as appropriate for each of the analyses.

We created three geocoded location files, which are described below.

- **Retailer location file.** This file has one observation for each SNAP-authorized retailer in Hampden County and in surrounding counties that include at least one HIP retailer (three additional counties in Massachusetts and two adjacent counties in Connecticut) for each of the 17 months of data. The variables include the retailer ID, geocoded location, retailer type, an indicator variable showing whether the retailer was a SNAP authorized retailer during each study month, and an indicator variable showing whether the retailer participated in HIP during each study month. The files collectively contain 941 unique SNAP retailer locations across the observation period in Hampden County or within 10 miles of the Hampden County border.[3]
- **Household-address location file based on case file address.** In this file, each SNAP participant household in Hampden County as of July 2011 has one observation for each address where the household lived during the 17-month study period. For example, this file has three observations for a household that lived in three locations. The variables include the household ID, an address index number (first address, second address, and so forth),

[3] To increase geocoding efficiency, we geocoded all unique FNS ID-address combinations at once. As retailers became SNAP-authorized and occasionally moved locations throughout the pilot, this number is higher than the total number of SNAP-authorized retailers in and around (within 10 miles of) Hampden County at any given time, which varied from month to month (e.g., 750 in July 2012 and 761 in October 2012).

geocoded location for the address, Census block group ID for the address, a HIP indicator, household size, and variables indicating the time period (month/year) for which each address was the household's residence. If the household moved during a month, we used the household address in which the household resided for the most days. If the household had no address (for example, because of homelessness), the address field is missing and the observation was excluded from analyses that use the household address (such as distance from household address to a retailer address). We successfully geocoded 98.7 percent of Hampden County SNAP participant household addresses. The file contains 69,510 SNAP participant address locations within Hampden County.[4]

- **Household-address location file based on survey file address.** For each household that completed a participant survey, this file has address information current as of the survey administration. We geocoded the address and the Census block group ID for each survey round completed. Using the household ID, this information can be linked to other household-level files. Overall, participant survey geocoding returned latitude and longitude data for 98.9 percent of records. The file contains 6,302 participant addresses across all three waves of data collection.

Further detail about the construction of these three files is provided in Appendix A.

Spending and Dietary Intake Data Matched with Location Data

We created household analysis files based on the EBT transaction data and the participant survey data.

- **EBT household-month-retailer analysis file.** This file contains one observation for each SNAP authorized retailer patronized in each month by each household in the EBT data. The file includes all of the retailers in which SNAP participants conducted transactions, even if the dollar amounts of the transactions at that retailer were small. The file contains unique identifiers for household and retailer, and the month in which data were collected. Information on SNAP participants includes in which wave the household began participating in the intervention, whether the household was assigned to the treatment or control group, household size, monthly SNAP benefit, the amount of benefits used in HIP participating retailers as well as non-participating retailers, and the amount of purchases that were eligible TFV. Retailer information includes the retailer category (supermarket/superstore, grocery store, convenience store, farmers' market, or other) and an indicator for whether the retailer participated in HIP.
- **Survey household-round analysis file.** We appended SNAP household location to the survey analysis files used for the HIP Final Report. We also appended several variables measuring distance from the household residence to the nearest supermarket/superstore and the nearest HIP supermarket/superstore.

4 All households lived in Hampden County at the time of random assignment to HIP and non-HIP groups. Approximately 8.5 percent of the 54,861 households in the evaluation sample lived outside Hampden County for one or more months during the evaluation period. They remained eligible to earn HIP incentives. These households were, however, excluded from the spatial analysis for the months they resided outside Hampden County.

We computed the distances between SNAP households and SNAP retailers using ArcGIS. Specifically, we computed both the distance to the nearest retailer of a given type and the distance to retailers at which SNAP participants redeemed their benefits. We computed three versions of each distance, which are described below.

- Euclidean distance, as the crow flies
- Driving distance, using a roadway network
- Walking distance, using a roadway network

Euclidean and driving distances differ substantially from one another, with driving distances being longer than Euclidean distances. The difference between Euclidean and driving distances varied across Hampden County. This difference was particularly striking in the densely populated north central portion of the county located to the immediate east and west of the Connecticut River. For some residents of this region the closest supermarket, as measured by Euclidean distance, was located on the opposite side of the river. Since the river can only be traversed using one of three bridges, Euclidean distances overestimated access to supermarkets for many Hampden county SNAP participants (See Appendix A, Exhibit A.5 for a visual illustration of this phenomenon). In this report we therefore use only network-derived (i.e., walking or driving) measures of distance to different types of food retailers.

Walking and driving distances are calculated differently based on road-specific travel restrictions. For instance, pedestrians are not allowed on highways, while service roads and walking paths are not open to privately owned vehicle traffic. We compared walking and driving distances and found that they were nearly identical to one another for Hampden County SNAP households. We therefore used driving distance to represent the two types of network distances.

2.3 Spatial Analysis

We performed the following four types of spatial analyses:

- Prepared maps and tabulations to describe the food retail environment in Hampden County, noting the most heavily used retailers and identifying areas with limited food retail access.
- Measured the effect of the distance to food retailers on food spending and food intake outcomes for SNAP participants.
- Explored "neighborhood effects" by estimating the spatial correlation in explanatory variables and unobserved factors that may influence key spending and dietary intake outcomes.
- Re-estimated our main HIP impact models on TFV expenditures and consumption, adjusting for spatial autocorrelation.

Description of Proximity to Retailers

We examined SNAP households' proximity to retailers by calculating the distance between household residence, as recorded in EBT files, and six types of retailers: any retailer; any HIP participating retailer; any supermarket or superstore; any HIP participating supermarket or superstore, high-volume supermarkets or superstores and HIP participating high-volume supermarkets or superstores.[5] We considered how these distances differed for neighborhoods in which relatively large or relatively small percentages of families were living below the poverty line. We use this distance information to construct maps as well as to generate covariates for inclusion in regression analyses of the relationship between distance to food retailers and HIP outcomes.

Effects of Food Retailer Access

The HIP Final Report describes the impact of the HIP intervention on participants' SNAP spending on, and their consumption of, targeted fruits and vegetables. These impacts were estimated controlling for a set of explanatory covariates. Controlling for covariates improved the comparability of the treatment and control groups and increased the precision of our estimators.

In this report, we took this analysis two steps further. First, we used variables describing the distance to food retailers as additional covariates to see if they influenced spending and consumption outcomes. Second, we investigated the interaction of HIP participation and the distance to food retailers, which allows us to measure whether the HIP impact differed for households that experienced different food retailer access conditions.

The models take the following form:

$$(1a) \quad Y = \beta_0 + \beta_1\, HIP + \beta_2\, ControlVars + u$$

$$(1b) \quad Y = \beta_0 + \beta_1\, HIP + \beta_2\, Distance + \beta_3\, ControlVars + u$$

$$(1c) \quad Y = \beta_0 + \beta_1\, HIP + \beta_2\, Distance + \beta_3\, HIP{*}\, Distance + \beta_4\, ControlVars + u$$

where *Y* represents an outcome of interest; *HIP* is a binary variable that identifies the treatment group; *Distance* represents a continuous or categorical measure of distance to the nearest supermarket; *HIP* Distance* represents the interaction of the distance variable and the treatment indicator; and *ControlVars* is a vector of characteristics measured as of the Round 1 (baseline) survey from SNAP case file data. Thus, Model (1a) is the conventional model for the HIP impact, as estimated in the HIP Final Report. Model (1b) augments the conventional model with a continuous or categorical measure of distance to show how the distance to food retailers affected the outcome. Model (1c) further augments the conventional model with the interaction of distance with the HIP dummy variable to show how food retailer access shifted the magnitude of the HIP impact on the outcome.

Note that the findings about the relation of distance to purchases—in Model (1b)—are observational analyses. They are not supported by the random assignment underlying the main HIP analysis and should be viewed as exploratory rather than causal. By contrast, findings about the relation of distance to HIP impacts—in Model (1c)—make use of the random assignment research design. The HIP impact estimates can reasonably be given a causal interpretation, so long as we can assume that HIP did not affect the proximity to stores (which seems to be a reasonable assumption).

---

[5] See Chapter 3 for a description of how high-volume retailers were identified.

We used the following two approaches to defining the distance to retailers.

- Defined the variable *Distance* as a continuous variable showing the distance to the supermarket in miles.
- Used categorical descriptions of food retailer access.

In each approach, we analyzed nine outcome variables. The following three outcome variables came from EBT data:

- SNAP purchases in all retailers
- SNAP purchases in HIP participating retailers
- Eligible TFV purchases in HIP participating supermarkets and superstores

The following six outcome variables came from participant survey data:

- Targeted fruit and vegetable intake
- Targeted vegetable intake only
- Targeted fruit intake only
- All fruits and vegetables
- Monthly SNAP expenditures
- Monthly fruit and vegetable spending

We defer detailed discussion of specific explanatory variables describing the distance to food retailers until the chapter in which specific results are presented.

Econometric Measures of the Magnitude of "Neighborhood Effects"

The regression-adjusted findings presented in the HIP Final Report, using models such as (1a) above, provided unbiased estimates of the impact of HIP on SNAP participants' purchase and consumption of fruits and vegetables. The error term (u) in such regression models refers to all the other factors—not explicitly in the model—that may have influenced food spending and food intake outcomes.

For each SNAP participant, these "other factors" may be more similar to those of nearby neighbors than to those who live farther away. Hence, the error terms in such regression models may be spatially correlated, meaning that there is more correlation in the errors for people who live closer to each other.

In this report, we investigate neighborhood effects by measuring the spatial correlation in these error terms. In some neighborhoods, there may be a cluster of households that have comparatively high (or low) values of the error term, which means higher (or lower) values of an outcome variable than one would expect based on the variables explicitly controlled in the model. For example, one cluster of neighbors might have surprisingly high (or low) fruit and vegetable intake.

The degree of spatial correlation in the error term depends on what variables are controlled in the model. In this analysis, we measure three types of spatial correlations, based on three possible sets of explanatory variables in the regression model.

- **No control variables.** The simple spatial correlation tells us about neighborhood effects that are attributable to a wide range of variables, including the characteristics of households in the neighborhood, economic variables, food retailer access, and unobservable factors.
- **Variables describing household characteristics.** In this analysis, the control variables are the same explanatory variables used in regression models in the HIP Final Report, as in model (1a) above. The spatial correlation in the error term (u) then tells us about neighborhood effects that are attributable to a narrower range of variables, including the distance to food retailers and unobservable factors.
- **Variables describing household characteristics and food retailer access.** In these analyses, we include additional explanatory variables describing the food retail environment, as in model (1b) above. We do this in two ways. The first incorporates a continuous measure of the distance between a SNAP household's residence and the nearest HIP participating supermarket or superstore, and the second uses a categorical description of food retailer access. The spatial correlation in the error term (u) then tells us just about neighborhood effects that are attributable to unobservable factors.

We expect that the measured amount of spatial correlation will be highest with no control variables and lowest with all of the available control variables.

In each case, we measure "neighborhood effects" using both global (*Moran's*) and local (*Getis-Ord Gi**) indicators of spatial autocorrelation. Both the Moran's I and the Getis-Ord Gi* statistics estimate the presence of clustering of particularly high or low values of a given outcome. Both the global and local measures estimate the degree to which households were more like their neighbors on key outcomes measures than would be expected given a random distribution of HIP outcome levels.

We calculated both indices for two HIP-related outcomes: (1) total SNAP purchases in HIP participating supermarkets and (2) eligible fruit and vegetable purchases in HIP participating supermarkets.

Econometric Estimates of the Impact of HIP, Controlling for "Neighborhood Effects"

Neighborhood effects and the induced spatial autocorrelation have the potential to induce bias in the precision (i.e., standard error) of the impact estimates. In addition, controlling for spatial autocorrelation has the potential to improve the estimates of impact.

To assess the impact of failing to control for spatial autocorrelation, we re-estimated the basic impact models, correcting for spatial autocorrelation. Estimates were performed in R. Computational considerations caused us to limit the sample for these analyses to households that lived within 0.25 mile of at least eight other SNAP households.

3. Geography, Demographics, and Food Retailer Environment in Hampden County

This chapter presents information on the geography, demographics, and food retail environment of Hampden County. The chapter has two sections. The first section examines geography, demographics, poverty, and concentration of SNAP participants across the county. The second section examines the locations of food retailers. Unless otherwise specified, data in this chapter that are related to SNAP households are drawn from the July 2012 EBT data.

3.1 Geography and Demographics of Hampden County, by Block Group

Located in southwestern Massachusetts, Hampden County encompasses 23 different municipalities within a 634-square-mile area. The county is demographically and economically diverse and contains a mix of urban, rural, and suburban areas. Although Hampden County has the lowest median household income of the 14 counties in Massachusetts, communities within the county vary widely in their population density, median household income, percentage of households for which a vehicle is not readily available, and percentage of the population living in poverty. Information on these block group characteristics for Hampden County is displayed in Exhibit 3.1.

Exhibit 3.1: Demographic Characteristics of Hampden County Block Groups (N=355)

	Mean[a]	Standard deviation	Minimum	Maximum
Number of persons per square mile	5,143	5,140	15	34,031
Median household income	$50,581	$24,620	$7,763	$154,261
Percentage of population living below the poverty line	19	18	0	80
Percentage of population with no vehicle available	15	15	0	72

[a]Mean across all block groups.
Sources: 2010 U.S. Census (population density); American Community Survey 2007–2011 Five-Year Estimates (all other variables).

Hampden County differs from the United States as a whole in a number of ways (see Exhibit 3.2). Although some individual block groups within Hampden County are comparatively densely populated compared with the average U.S. county, Hampden County as a whole is much less dense. In terms of median household income and the percentage of the population living below the poverty line, Hampden County is roughly comparable to the rest of the country, but a higher percentage of Hampden County residents have no access to a vehicle.[6]

[6] The mean *block group* population density (in Exhibit 3.1) is much greater than the *county-wide* population density (in Exhibit 3.2), because there are many more block groups in densely populated sections of the county than in rural sections. By design, census block groups have approximately equal populations, so the block groups have a much smaller land area in urban places than in rural places.

Exhibit 3.2: Demographic Characteristics: Hampden County Compared With All U.S. Counties

Variable	Hampden County[a]	All U.S. counties[b]			
		Mean	Standard deviation	Min.	Max.
Number of persons per square mile	751	2096	6586	0	69468
Median household income	$49,729	$55,115	$14,507	$19,624	$122,844
Percentage of population living below the Federal Poverty Line	17	15	5	0	50
Percentage of population with no access to a vehicle	13	9	9	0	82

[a]Mean for the entire county (not averaged across block groups as in Exhibit 3.1).
[b]Means weighted by total population of the county.
Source: American Community Survey 2007–2011 Five-Year Estimates. N = 3,143 counties (all counties within the 50 States).

Exhibit 3.3 displays the variation in the demographic characteristics for Hampden County block groups using four maps. In these maps, we shade the block groups to correspond to their values for population density, median household income, percentage of households without a car available for use, and percentage of the population whose income is below the poverty line.

As Exhibit 3.3 indicates, the more densely populated areas of the county are located in the north central region, which is bisected by Interstate 90 (the Massachusetts Turnpike, the State's main east-west road) and by the Connecticut River (flowing north-south). This area—including the communities of Holyoke, Chicopee, Springfield, West Springfield, Agawam, Long Meadow and East Longmeadow—is home to approximately three-quarters of Hampden County residents.

In order to better observe block-level differences for this important region of the county, we redisplay the maps of demographic characteristics from Exhibit 3.3 zoomed to this north central region (Exhibit 3.4) and provide summary statistics in Exhibit 3.5. These communities are, on average, more densely populated and have lower median household incomes, higher percentages of their populations living in poverty, and lower rates of vehicle access than the rest of the county.

Exhibit 3.3: Demographic Characteristics of Hampden County, by Block Group

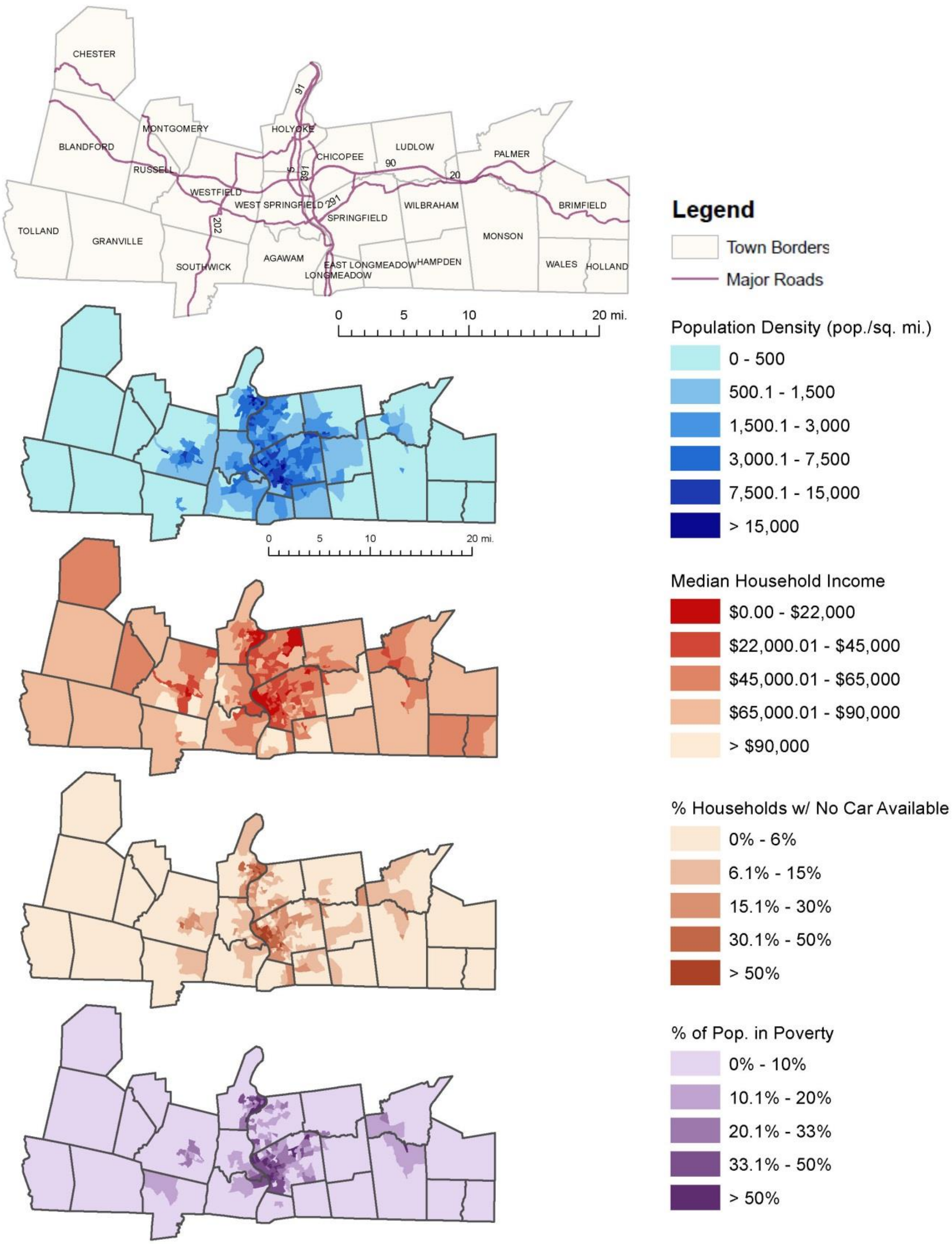


Exhibit 3.4: Demographic Characteristics of North Central Region Hampden County, by Block Group

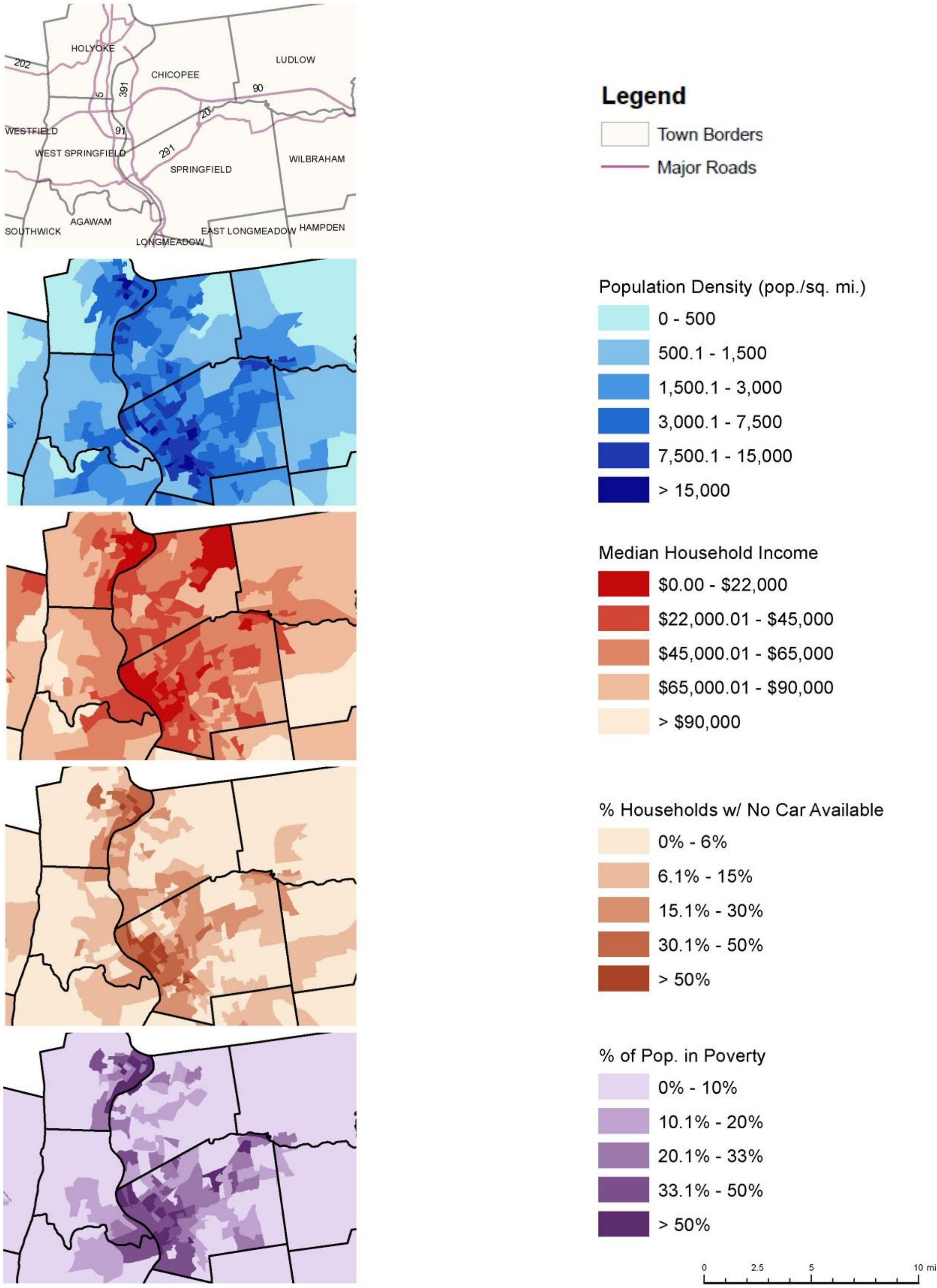


Exhibit 3.5: Demographic Characteristics of North Central Region Hampden County Block Groups (N=242)

	All Hampden County	North Central Region			
		Mean	Standard deviation	Minimum	Maximum
Number of persons per square mile	5,143	6,590	5,247	429	34,031
Median household income	$50,581	$43,200	$20,000	$7,763	$108,958
Percentage of population living below the poverty line	18	22	20	0	80
Percentage of population with no access to a vehicle	14	18	16	0	72

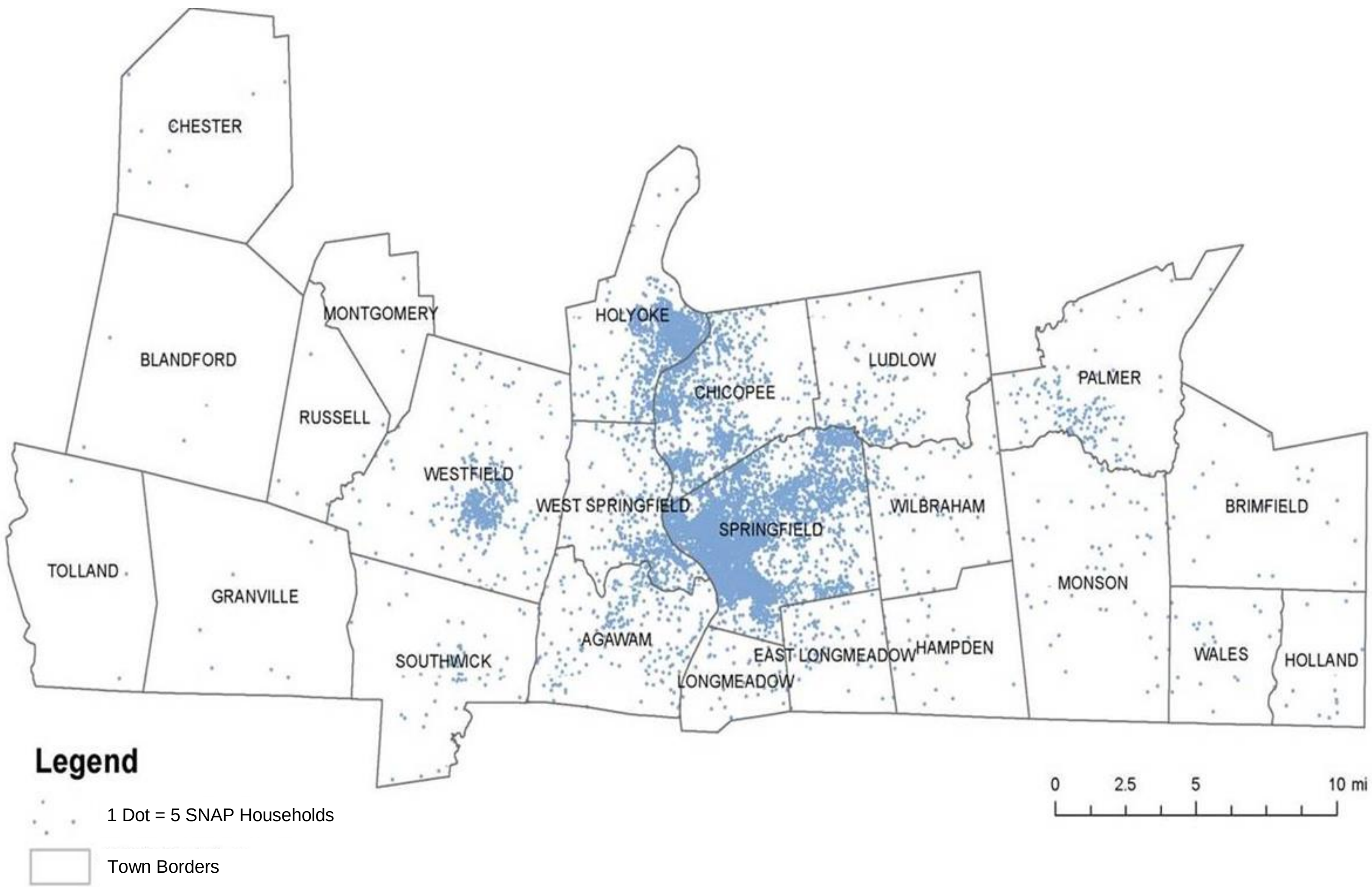
Represents 68 percent of the County's block groups and 89 percent of SNAP participant households.
Source: American Community Survey 2007–2011 Five-Year Estimates.

SNAP participants were strongly concentrated in the low-income and urban areas in the north central region of the county, as would be expected. Approximately 89 percent of SNAP participant households lived in the north central region in July 2012. The map in Exhibit 3.6 displays the distribution of SNAP participants across the county. In this map, each dot represents five SNAP households.[7]

Exhibit 3.7 displays bus routes in Hampden County and nearby Hampshire County. This map indicates that there are more bus routes in the high population density north central area of the county. There is no public transportation in the more sparsely populated western part of the county and there are only a small number of bus routes in the eastern portion of the county. The areas of the county with no bus routes are also the areas of the county with the highest levels of vehicle ownership/access.

[7] Dots are placed randomly within a SNAP participant's block-group so as to not reveal the actual location of any individual SNAP household.

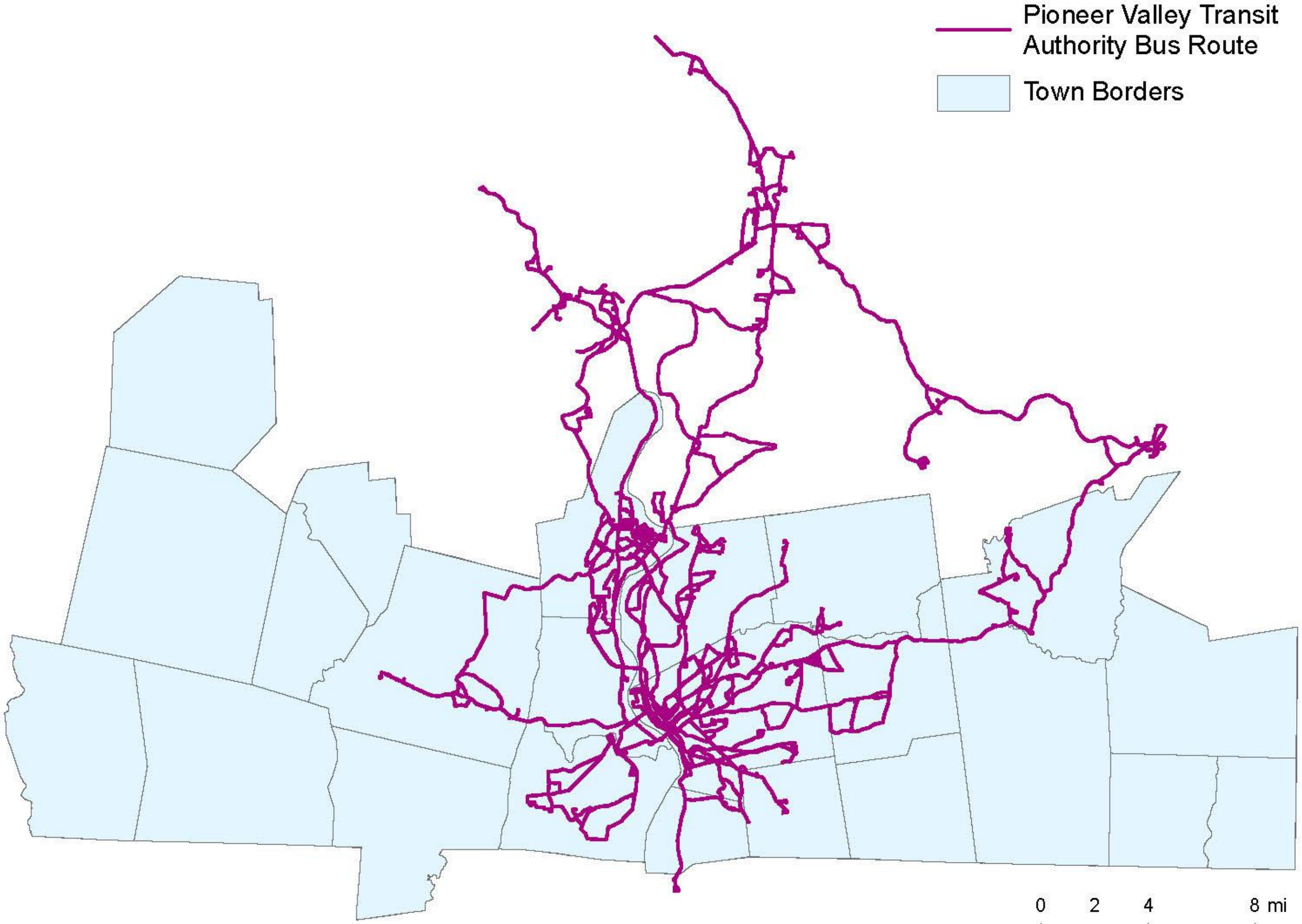
Exhibit 3.6: SNAP Household Locations in Hampden County, July 2012[8]



---

8 Dots are randomly placed within each block group to provide an illustration of the block group level distribution of Hampden County SNAP participants. They do not represent the actual latitude and longitude of any individual SNAP household's address.

Exhibit 3.7: Bus Routes in Hampden and Hampshire Counties



3.2 Food Retailer Environment

The food retailer environment of Hampden County SNAP participants includes stores located in Hampden County as well as those located just outside the county borders. Findings from the food access literature suggest that including stores within 10 miles of the Hampden County border represents an appropriate measure of the food retailer environment of SNAP participants who lived in the most rural areas on the county border and would thus be more than adequate for residents in the densely populated interior of Hampden County.[9]

In the analysis that follows in this chapter and the rest of the report, we categorize food retailers as belonging to one of five types. We generated these categories by collapsing the 14 FNS-defined food retailer categories listed in the EBT data files into the following groups: (1) supermarkets (supermarkets and superstores); (2) grocery stores (small, medium, and large grocery stores and fruit/vegetable, seafood, and meat specialty stores); (3) convenience stores (convenience store and combination grocery/other); (4) farmers markets (including direct marketing farmers); and (5) other retailers (not eligible for HIP: bakery specialty stores and food cooperatives).

Food retailer density varies widely across the county. There were few food retail establishments of any type in the low-population density, western portions of the county (Exhibit 3.8). Grocery stores were concentrated in the central portion of the county, while supermarkets were primarily located in the central and eastern-central portions of the county. Notably, some portions of the densely populated, largely low-income western half of the City of Springfield, as well as the most densely populated areas of Holyoke, had relatively few supermarkets. Food retail stores in these areas were primarily smaller grocery stores and convenience stores. Conversely, farmers' markets were relatively evenly distributed throughout the central third of the county, covering the area from its southern border with Connecticut up into Hampshire County.

As an alternate visualization, Exhibit 3.9 displays the density of retailers for each census block group by showing the number of SNAP-authorized retailers of all types located within 0.5 mile, one mile, and 5 miles from each block group's SNAP population-weighted centroid.[10] The north central region of the county had the highest retailer density with some of the block groups that border the Connecticut River having more than 10 retailers within a mile of the block group center. Density was lowest in the eastern- and western-most areas of the county.

---

9 Rose and Richards (2004) reported that 27 percent of a nationally representative sample of SNAP participants purchased most of their food at a retailer further than five miles from their homes, which set a lower bound for inclusion distance. A 2012 report by the USDA Economic Research Service showed that 93 percent of rural residents in the United States lived within 10 miles of a supermarket, suggesting that a 10-mile buffer would be adequate.

10 A "centroid" is the geographic center of a given area. Here, we generated "population-weighted" centroids, that is, the geographic center of each block group relative to the people who lived in that block group, which can be operationalized in a number of ways. Specifically, the population-weighted centroids we calculated were the points within each block group with the lowest possible mean distance to the homes of the SNAP participants who lived there.

Exhibit 3.8: SNAP Retailers Within Hampden County or Within 10 Miles of Hampden County (July 2012)

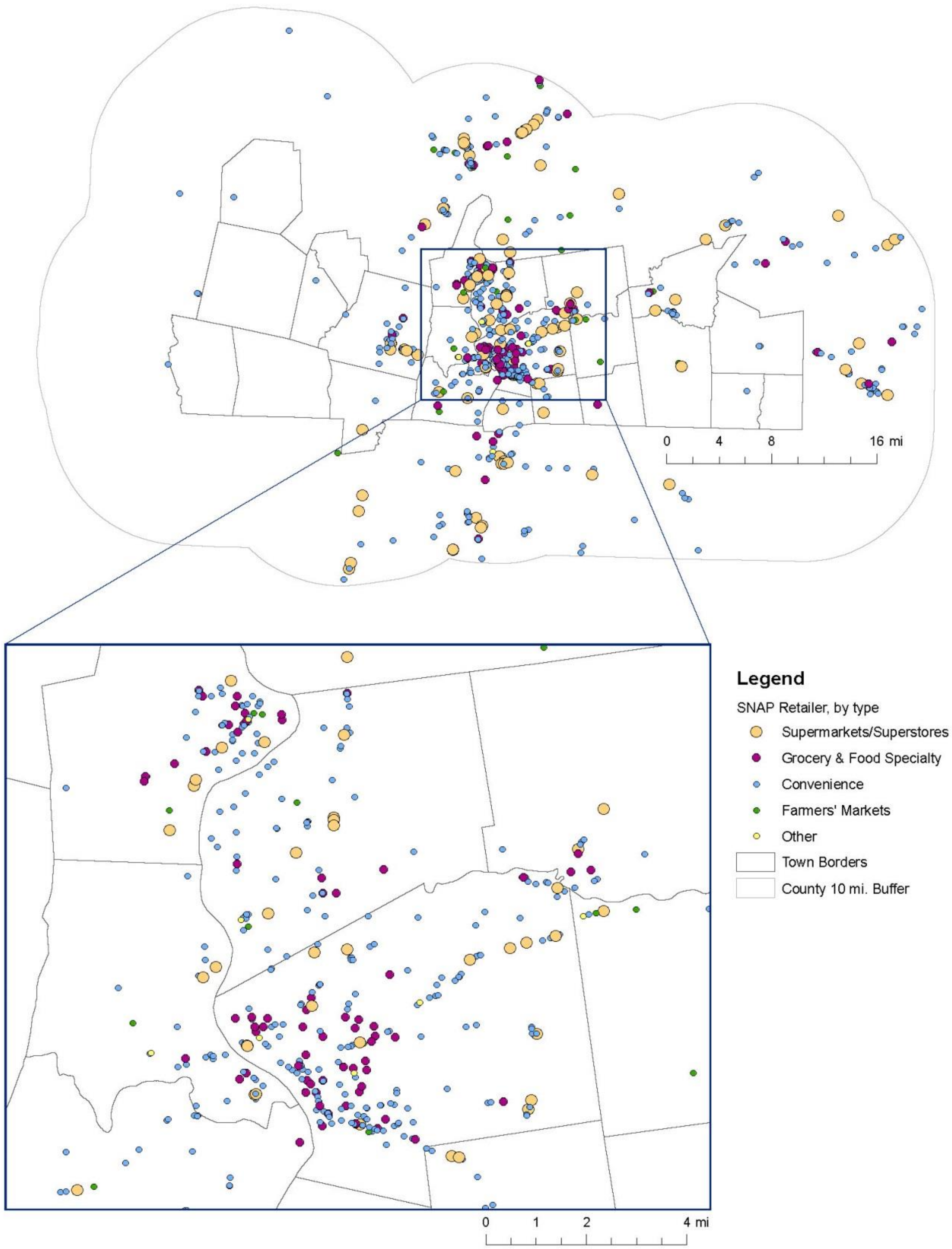
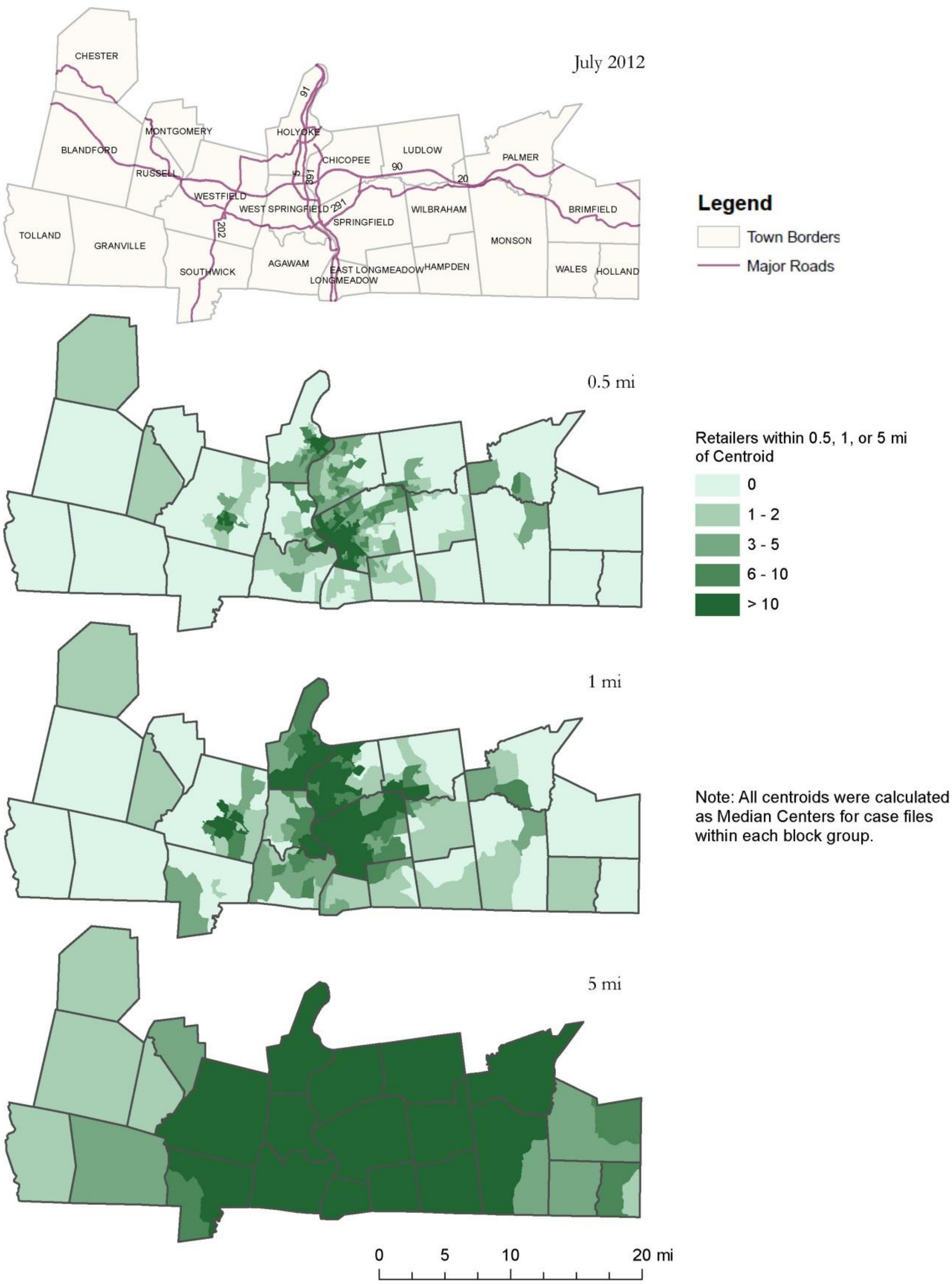


Exhibit 3.9: Density of SNAP Retailers, by Block Group



Of the 750 SNAP retailers in our analysis area in July 2012, 94 (12 percent) were HIP participating retailers, though as noted below, SNAP redemptions in these stores represented about 50 percent of all SNAP redemptions.[11] Exhibit 3.10 displays the locations of all five types of HIP participating retailers. It shows that, similar to the findings for all SNAP retailers, HIP participating food retailers were more plentiful in the north central portion of the county than in the eastern and western portions. That said, although most north central-region block groups contained at least one HIP participating retailer, there were some communities in which there were no HIP participating supermarkets located within one mile of the center of the block group (Exhibit 3.11).

[11] The analysis area includes retailers located in Hampden County or within 10 miles of the county border. This analysis pertains to July 2012. Several additional stores began participating in HIP in October 2012; this had only a slight effect on the overall retailer environment as presented in this chapter. However, several of these additional retailers were supermarkets, which substantially increased access to HIP supermarkets for households in some parts of Hampden County, as shown in the maps in Chapter 4.

Exhibit 3.10: HIP Participating Retailers Within Hampden County or Within 10 Miles of Hampden County (July 2012)

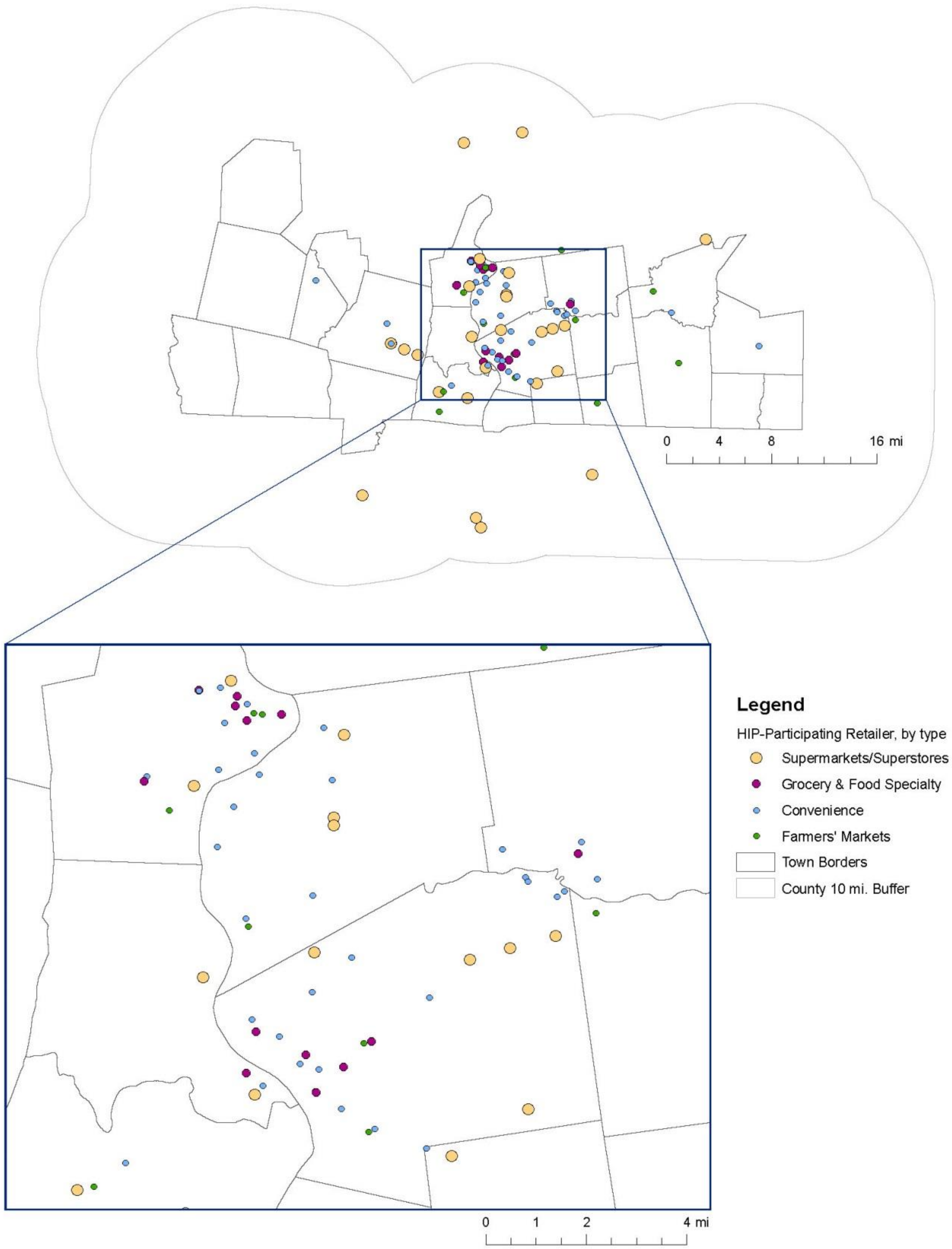


Exhibit 3.11: Density of HIP and Non-HIP Retailers in the North Central Region (July 2012)

Food Retailer Classification

The analysis of the food retailer environment discussed above informed the choice of measures we used in the spatial analysis, presented in Chapters 4 and 5, to characterize the food retailer access of SNAP participants during the pilot. We developed three measures to characterize retailer access. The first measure simply looked at whether a household's closest supermarket participated in HIP or not. The second measure incorporated distance to the nearest HIP participating and non-participating supermarket. The third measure took into account that there were some supermarkets/superstores that were heavily patronized by SNAP households. Our approach to developing this third measure is discussed below.

Our analyses of the EBT data indicated that a substantial proportion of SNAP redemptions were concentrated in a small number of retailers. There were 4,523 SNAP-authorized retailers (both within Hampden County and across the U.S.) at which at least one SNAP household from Hampden County spent some SNAP benefits in October 2011, the month prior to HIP implementation. Total SNAP redemptions that month for these SNAP households totaled to $12.74 million, but only 4 percent of the retailers (191) had redemptions totaling over $5,000 (Exhibit 3.12). Within that group, 13 individual retailer locations accounted for 51 percent of all SNAP redemptions (Exhibit 3.13). Given that these 13 retailers accounted for a disproportionate share of total SNAP spending, it is possible that they represent a different class of retailer than other stores.[12] Our third measure of food retailer access therefore incorporated distance to what we term "high-volume" retailers.

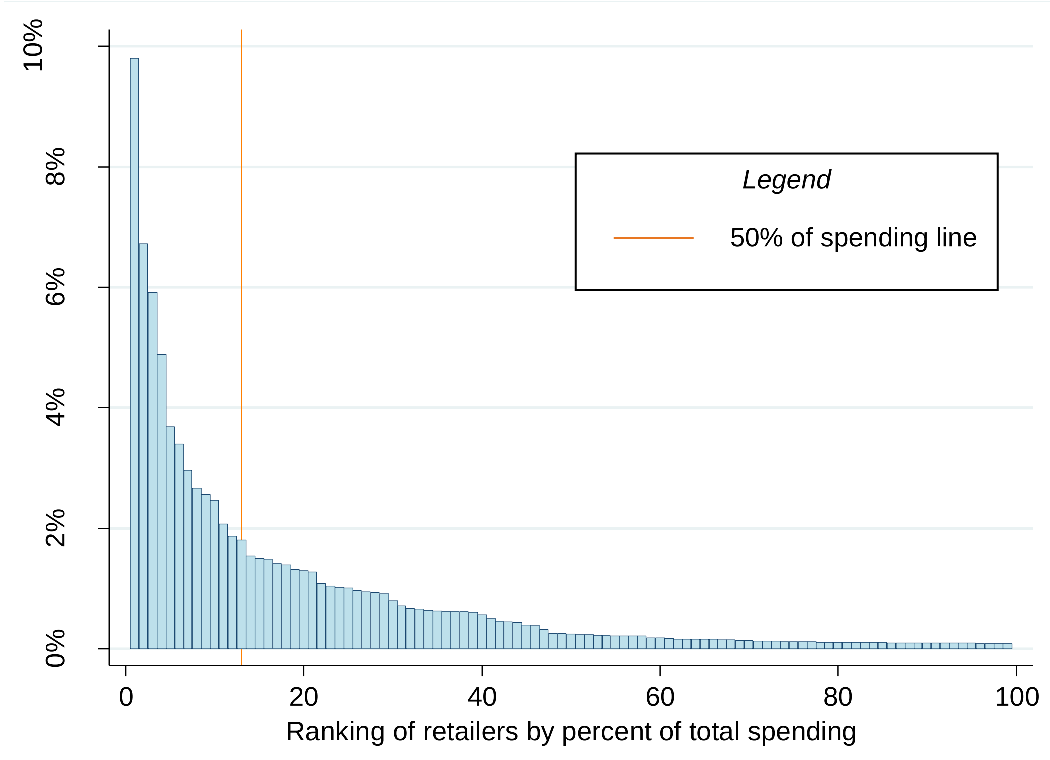
Exhibit 3.12: Distribution of Total SNAP Redemptions Across SNAP Retailers: October 2011

Total SNAP redemptions at a given retailer, October 2011	Number of retailers	Percent	Cumulative Number of retailers	Cumulative Percent	Total $ spent	% of total $ spent	Cum. % of total spent
> $1,000,000	1	0.02	1	0.02	$1,249,045	9.8	9.8
$500,000.01–$1,000,000	3	0.07	4	0.09	$2,231,767	17.5	27.3
$200,000.01–$500,000	9	0.20	13	0.29	$2,990,404	23.5	50.8
$100,000.01–$200,000	17	0.38	30	0.67	$2,537,646	19.9	70.7
$50,000.01–$100,000	15	0.33	45	1.00	$1,087,750	8.5	79.3
$10,000.01–$50,000	57	1.26	102	2.26	$1,058,446	8.3	87.6
$5,000.01–$10,000	89	1.97	191	4.23	$624,154	4.9	92.5
$1,000.01–$5,000	244	5.39	435	9.62	$601,420	4.7	97.2
$500.01–$1,000	127	2.81	562	12.43	$87,620	0.7	97.9
≤ $500	3,961	87.57	4,523	100.00	$268,962	2.1	100.0
Total[a]	4,523	100.00			$12,737,214	100.0	

[a] Includes all SNAP-authorized retailers *across the entire United States* at which at least one SNAP participant from Hampden County made any kind of EBT transaction in October of 2011.
Source: EBT Transaction Data, October 2011 (N=47,381 households).

12 We included these 13 retailers and not others in the high-volume category because: (1) these were the only retailers in which monthly redemptions exceeded $200,000; and (2) we observed a sharp drop-off in store level redemptions after these 13 stores. The 13 retailers remained the top 13 and continued to represent 50 percent of all SNAP redemptions for the control group throughout the duration of the pilot.

Exhibit 3.13: Distribution of Total SNAP Redemptions Across SNAP Retailers (October 2011)



Source: EBT Transaction Data, October 2011 (N=47,381 households).

3.3 Discussion

In summary, there were substantial differences in the demographic characteristics, population density, and food retailer access across Hampden County communities. The urban north central region of the county was home to nearly 9 out of every 10 Hampden County SNAP households. Access to food retailers in this central region was generally better than in the more sparsely populated western and eastern portions of the county. In this north central region, most but not all neighborhoods had a SNAP retailer within a distance of one mile. Households' SNAP spending was not spread evenly across all these neighborhood supermarkets. Instead, most SNAP benefits were spent in just 13 major supermarkets.

4. SNAP Household Access to Food Retailers

In this chapter, we examine SNAP households' access to food retailers in two ways. First, we examine the distance between each SNAP household's residence and the nearest SNAP-authorized retailer—for different types of retailers. Second, we examine the distances between where SNAP households live and the locations where they used their SNAP benefits. Unless otherwise noted, these numbers are drawn from the EBT files for March 2012, July 2012, and October 2012.

4.1 Proximity of SNAP Participants to Nearby Retailers

The median distance between the residences of Hampden County SNAP participants and the closest SNAP-authorized food retailer of any type was approximately 0.22 miles (Exhibit 4.1). The median distance to the closest supermarket[13] was 0.81 miles, and the median distance between residences and the closest HIP participating supermarket was 1.51 miles.

Exhibit 4.1: Average Distances (Miles) Between SNAP Participants and Retailers, by Type of Retailer

	Median	Mean	(SD)	Minimum	25th percentile	75th percentile	Maximum
Any retailer	0.22	0.34	0.51	0.00	0.11	0.38	13.73
Supermarket	0.81	1.02	1.09	0.00	0.56	1.15	20.50
HIP supermarket	1.51	1.79	1.74	0.00	0.96	2.06	31.37
High-volume supermarket	1.48	2.45	2.90	0.00	0.86	2.82	37.09
HIP participating high-volume supermarket	2.30	3.00	2.77	0.03	1.57	3.25	37.09

Source: EBT Transaction Data, pooled across March 2012, July 2012, and October 2012 (average of 42,219 households per month).

Exhibit 4.1 presents county-wide figures. In addition, it is interesting to see differences from one block group to another. Exhibit 4.2 provides these data in terms of the percentage of SNAP participants who lived in communities (block groups) that were within 0.5 miles, 1 mile, and 5 miles from different types of food retailers.[14] Although almost all Hampden County SNAP participants lived within one mile of some type of food retailer, approximately two-thirds of that population lived within one mile of a supermarket, and roughly a quarter lived within one mile of a HIP participating supermarket. Similarly, although approximately one-third of the population lived within one mile of a high-volume supermarket, only 10 percent lived within one mile of a HIP participating high-volume supermarket.

---

[13] In this chapter and throughout the report we use the term "supermarket" to represent supermarkets and superstores.

[14] USDA's Food Access Research Atlas (FARA), an online mapping tool, provides information for threshold distances of 0.5 miles, 1 mile, 10 miles, and 20 miles, with the latter two thresholds designed mainly for applications in rural areas. We added the 5-mile distance in this analysis because Hampden County has a comparatively dense population distribution, so there is no variation in access as measured using the 10- and 20-mile thresholds.

Exhibit 4.2: Percentage of SNAP Participants Who Lived in Areas With Specified Minimum Distance to Retailers

	Within .5 miles	Within 1 mile	Within 5 miles
Any retailer	84.5	94.3	99.8
Supermarket	19.6	65.1	99.0
HIP supermarket	6.7	26.2	96.2
High-volume supermarket	8.6	31.1	89.1
HIP participating high-volume supermarket	2.5	10.0	88.2

Source: EBT Transaction Data, pooled across March 2012, July 2012, and October 2012 (average of 42,219 households per month).

Some previous research has found that low-income communities are more likely to have limited access to food retailers, especially when examining the availability of healthy food options in those neighborhoods (Larson, Story, and Nelson, 2009; Powell et al., 2007). More recent research using nationally representative data has found that low-income neighborhoods are more likely to have access to a supermarket (USDA Economic Research Service, 2013). This is likely because many low-income neighborhoods have comparatively high population density on average, and high-density neighborhoods in turn are more likely to have a supermarket than low-density neighborhoods. In Exhibit 4.3 we display data on access to HIP participating supermarkets by the percentage of persons in the community living below the Federal Poverty Line. This provides a mixed picture of the relationship between community poverty levels and supermarket access. Access to HIP participating supermarkets and superstores was highest among persons living in areas with between 10 and 20 percent poverty. Areas of Hampden County in which a third or more of the population lived below the Federal Poverty Line had higher percentages of SNAP participants with limited access to a HIP supermarket. Areas with less than 10 percent of families living below the Federal Poverty Line were quite similar to those with 20 to 33 percent in terms of access to HIP supermarkets.

Exhibit 4.3: Food Access by Neighborhood Percentage of Persons Living in Poverty

Percent of poverty	Household within one mile of supermarket (%)	Household within one mile of a HIP participating supermarket (%)	Household within one mile of a high-volume supermarket (%)	Household within one mile of a HIP participating high-volume supermarket (%)
Less than 10% poverty	38	21	9	7
10%-20% poverty	63	33	35	17
20%-33% poverty	72	22	25	7
33%-50% poverty	78	13	47	5
More than 50% poverty	72	9	35	3

Source: EBT Transaction Data, July 2012 (42,180 households).

In Exhibits 4.4 and 4.5, block group-level access to different types of retailers is mapped for the north central region of the county. Areas shaded in blue are within one mile of a given category of retailer; areas shaded in red represent block groups in which 20 percent or more of the population was living below the Federal Poverty Line. Areas that appear purple (where blue and red overlap) are those block groups where 20 percent or more of the population lived below the Federal Poverty Line that were also within one mile of the given category of retailer. Exhibit 4.4 uses data from July 2012 and

Exhibit 4.5 uses data from October 2012. Nearly all low-income areas are within one mile of some type of retailer, and most are within one mile of a supermarket. In July, a substantial portion of low-income areas did not have a HIP participating supermarket within one mile. However, this access issue was substantially mitigated in October with the addition of four new HIP participating supermarkets in Springfield, Chicopee, and Holyoke.

Moreover, SNAP participants did not usually spend their benefits in the supermarkets that were nearest to them. As Exhibit 4.6 shows, block group average distance to the stores where people actually shopped were often considerably greater than block group average of household distances to the nearest supermarket. This is not to say that distance to food retailers does not matter. Rather, it shows that SNAP participants in Hampden County did not decide where to shop exclusively based on the proximity of a given retailer to their home; they were often mobile enough to shop for food at stores that were some distance from where they lived.

Exhibit 4.4: Access to SNAP-Authorized Retailers by Block Group, Central Region, July 2012

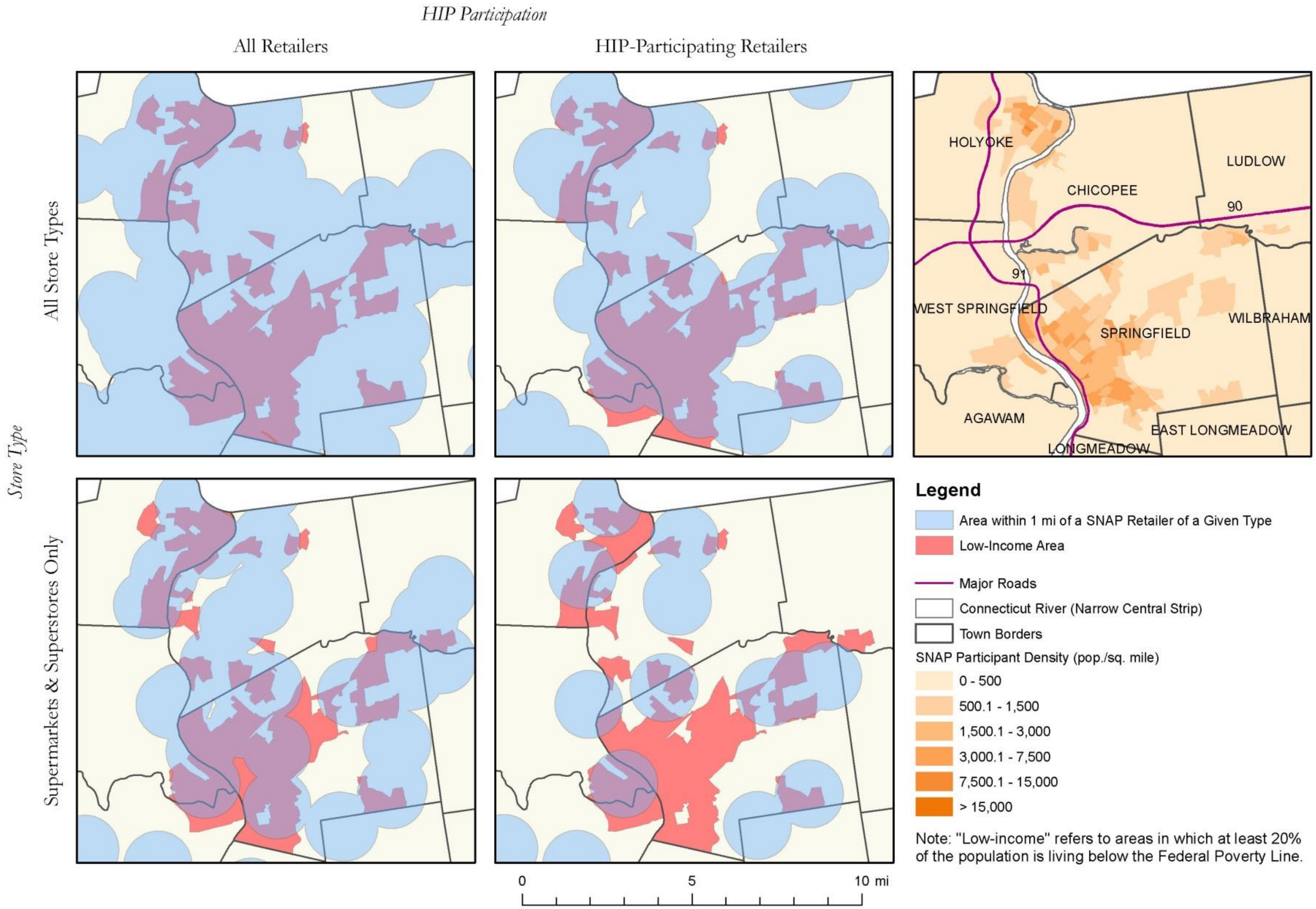


Exhibit 4.5: Access to SNAP-Authorized Retailers by Block Group, North Central Region, October 2012

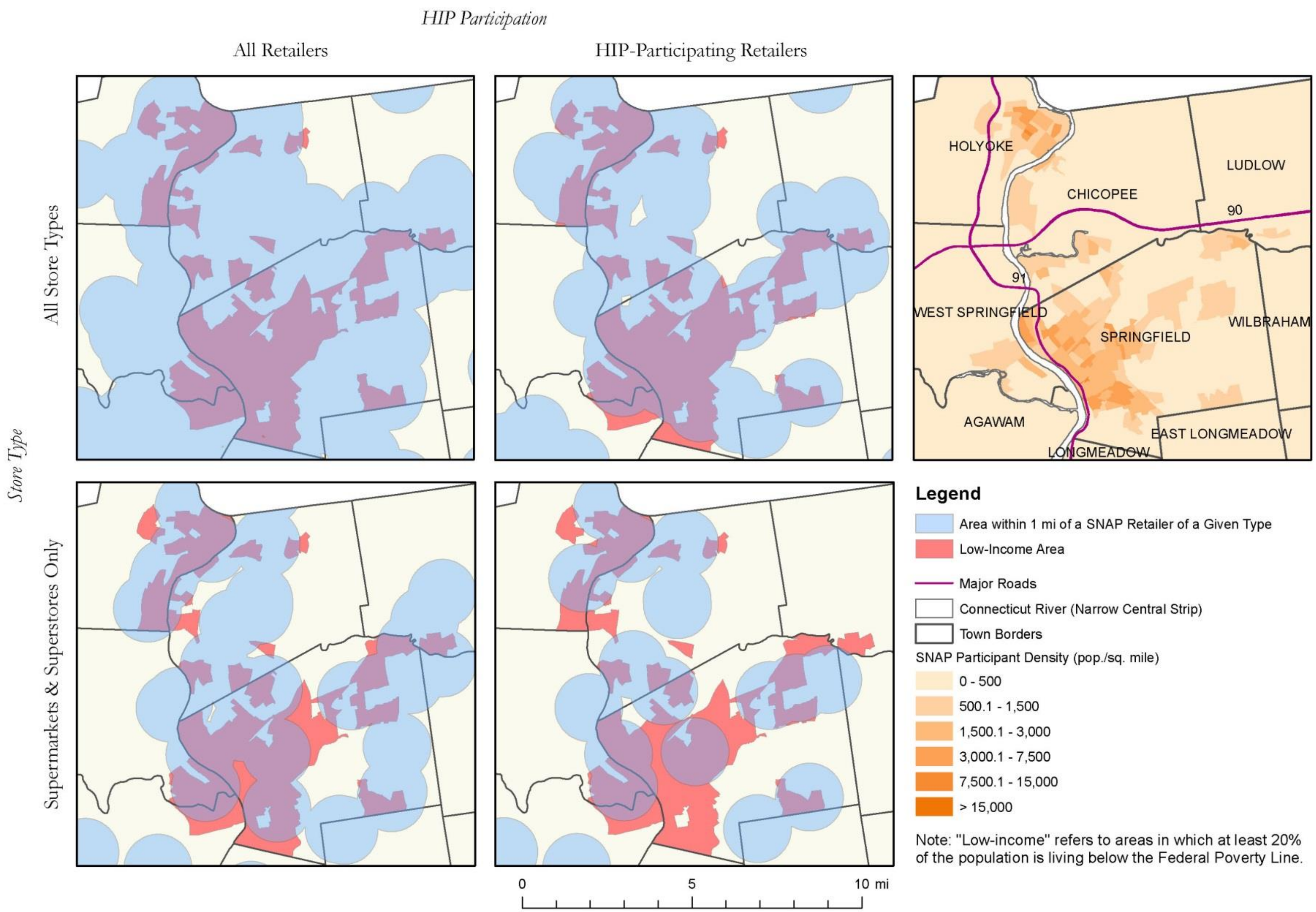
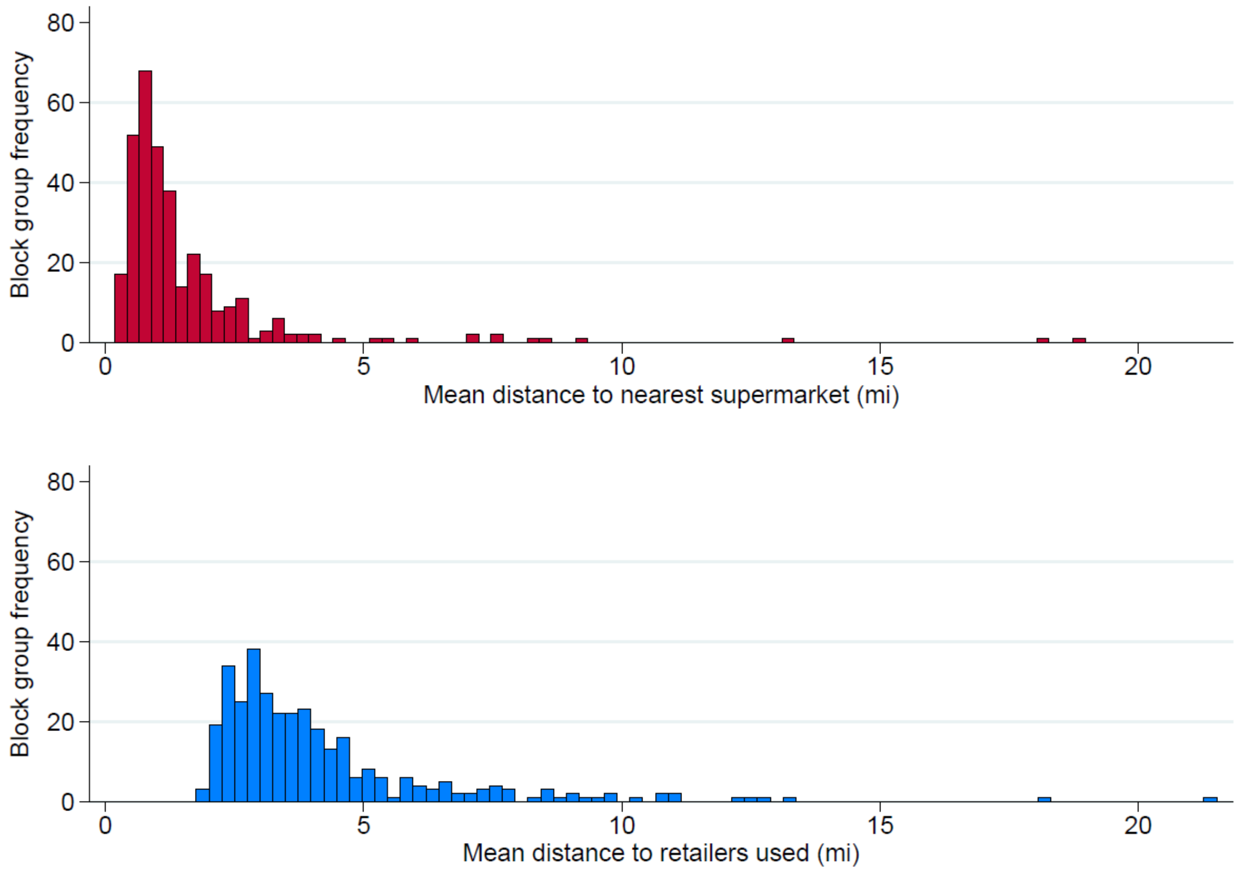


Exhibit 4.6: Mean Distance to Supermarkets and to Retailers Used, by Block Group



Source: EBT Transaction Data, July 2012 (42,180 households).

4.2 Proximity of SNAP Participants to the Retailers Where They Actually Shopped

The previous section considered proximity to various types of retailers—any type, supermarkets, and HIP participating stores. In this section, we show that SNAP households' food spending was concentrated in certain retailers, often at a relatively great distance from their residence. Thus, simple tabulations of distance to the closest retailer or the closest retailer of a given type may not describe actual shopping behavior.

In this section, we examine the distances between SNAP households' residences and the retailers where they spent their benefits. Although SNAP participants shopped in more than one store, the bulk of SNAP dollars were spent in supermarkets/superstores (Exhibit 4.7). Specifically, an average of 78.4 percent of SNAP dollars per month were spent in supermarkets, 9.7 percent in convenience stores, 9.2 percent in grocery stores and food specialty shops, 2.1 percent in other types of retailers, and less than a tenth of a percent in farmers' markets.

Exhibit 4.7: Percentage of SNAP Spending in Different Types of Food Retailers

Supermarkets/super stores	Grocery and food specialty	Convenience	Farmers' markets	Other
78.4	9.2	9.7	0.07	2.1

Source: EBT Transaction Data, pooled across March 2012, July 2012, and October 2012 (average of 42,219 households per month).

As noted in Chapter 3, more than 50 percent of SNAP spending was concentrated in 13 high-volume supermarkets. Use of a high-volume supermarket was, on average, higher among Hispanic and Black households, female-headed households and families that lived in close proximity to these retailers (See Appendix B, Exhibit B.1).

The average SNAP household did the bulk of its shopping (defined as 90 percent of total SNAP dollars spent during the month) at approximately three separate retailers each month during March, July and October of 2012 (Exhibit 4.8). Although a small number of SNAP participants shopped in a large number of stores, most households made 90 percent of their SNAP purchases in between one and five stores.

Exhibit 4.8: Average Number of Retailers in Which Households Made 90 Percent of Their SNAP Purchases

Mean	SD	Min	Max
3.03	1.69	1	20

Source: EBT Transaction Data, pooled across March 2012, July 2012, and October 2012 (average of 42,219 households per month).

Further, although SNAP households shopped in multiple retailers, most tended to make the majority of their SNAP purchases in one store. On average, SNAP participants made more than 62 percent of their SNAP purchases in a single retailer. For approximately 85 percent of SNAP participants, supermarkets were their primary type of retailer (Exhibit 4.9). In addition, 7.3 percent of households made the majority of their SNAP purchases in grocery and food specialty stores, while only 4.8 percent made the majority of their SNAP purchases in convenience stores. Less than 1 percent primarily did their shopping in farmers' markets and approximately 1.7 percent used other retailer types.

Notably, these figures are quite similar for households that did and did not live within one mile of a supermarket. Among households that lived within one mile of a supermarket, 84.6 percent made the majority of their monthly SNAP purchases in a supermarket. Among households that did not live within one mile of a supermarket, a slightly higher 86.3 percent made the majority of their SNAP purchases in a supermarket. This suggests that households' preferences for specific supermarkets were strong enough that they were willing to patronize supermarkets that were not within one mile of where they lived. Households that lived farther away from a supermarket may have made up for the extra travel distance by concentrating their shopping at fewer retailers.[15]

[15] Median = 2.9 retailers for households that were not within one mile of a supermarket and 3.1 retailers for households that were within one mile of a supermarket.

Exhibit 4.9: Type of Retailer Where SNAP Households Made the Majority of Their SNAP Purchases

	All households (%)	Households within one mile of a supermarket (%)	Households not within one mile of a supermarket (%)
Type of retailer			
Supermarkets	85.2	84.6	86.3
Grocery and food specialty	7.3	7.9	6.2
Convenience	4.8	5.1	4.4
HIP participating retailer	33.1	38.3	31.2

Source: EBT Transaction Data, pooled across March 2012, July 2012, and October 2012 (average of 42,219 households per month).

HIP participating retailers were the primary retailer for approximately 33 percent of SNAP households. Among SNAP household that lived within one mile of a HIP supermarket, 38 percent made the majority of their SNAP purchases as at HIP participating retailer. Among SNAP households that did not live within one mile of a supermarket, 31 percent made the majority of their SNAP purchases as at HIP participating retailer.

SNAP participants traveled an average of 3.22 miles from their residences to their primary retailer (Exhibit 4.10). Earlier analysis showed that nearly two-thirds of SNAP participants lived within one mile of a supermarket (Exhibit 4.2) and that the median distance from the residences of SNAP households to their nearest supermarket was 0.81 miles (Exhibit 4.1). Taken together, these findings suggest that many SNAP participants traveled extra distances to shop in stores other than the supermarket closest to their neighborhood. In fact, of all the benefits SNAP households spent in supermarkets, only 16 percent were spent in the supermarket that was closest to their residence.

That said, households that did not live within one mile of a supermarket traveled farther to their primary retailer than did households that lived within a mile of a supermarket (Exhibit 4.10). These findings, when considered in the context of the types of stores in which households shopped (Exhibit 4.9), suggest that whether or not a SNAP household lived within one mile of a supermarket impacted how far they traveled to purchase food but not the types of retailers they frequented.

Exhibit 4.10: Distance to the Primary Retailer Where SNAP Purchases Occurred

	Mean	SD	Min	Max
All households	3.22	2.95	0	41.59
Households within one mile of a supermarket	2.89***	2.60	0	37.48
Households not within one mile of a supermarket	3.85***	3.42	0	41.58

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.
Source: EBT Transaction Data, pooled across March 2012, July 2012, and October 2012 (average of 42,219 households per month).

Some household types were more likely to be in the small group (15 percent of all households) that shopped primarily at a retailer other than a supermarket or superstore. Appendix B presents three sets of models to investigate this issue (see Exhibits B.2 through B.4). Simple models show that shopping primarily in a non-supermarket was comparatively more frequent for households in which Spanish was spoken, households with a non-white head of household, households with no income, households headed by a person with a disability, and male-headed households.

4.3 Discussion

In summary, approximately one-third (34.9 percent) of SNAP households did not live within a mile of a supermarket and nearly three-quarters (73.8 percent) did not live within a mile of a HIP participating supermarket where it was possible to earn the HIP incentive. Although there were some high-poverty communities in which there were no nearby (within one mile) supermarkets, this was also true of lower-poverty communities.

Hampden County SNAP households spent the bulk of their benefits at supermarkets and superstores and typically did all of their household shopping at between one and three stores per month. This preference for supermarkets was true for families regardless of whether or not they lived within a mile of a supermarket, but it did differ systematically based on other family characteristics such as language spoken at home and income. Finally, SNAP households, on average, shopped at supermarkets other than those that were nearest to their homes; more than half of all SNAP benefits were spent at one of 13 high-volume retailers.

5. The Role of Retailer Access in Fruit and Vegetable Spending and Intake

This chapter considers whether the distance to food retailers is related to the purchase and consumption of targeted fruits and vegetables. It addresses the following two questions:

- Is the distance from SNAP households' homes to food retailers related to their spending on and consumption of fruits and vegetables?
- Does the impact of HIP on fruit and vegetable spending and consumption vary by distance to food retailers?

Distance to food retailers could affect a household's food expenditures at particular stores and consumption of particular food items. If, for example, a household lived close to a supermarket that carried a wide variety of fruits and vegetables at competitive prices, the household might purchase more of these items than if there were no such store nearby. Distances to food retailers was not randomly determined in HIP, so these analyses should be viewed as correlational (not causal) and exploratory.

In addition, the impact of HIP might be expected to vary with distance to food retailers. For example, the impact of HIP on fruit and vegetable spending could be smaller in areas where all households had to travel relatively far to supermarkets. These analyses take advantage of HIP's random assignment design and, as discussed in Chapter 2, can be viewed as causal.

In this chapter, we focus on the following three outcome variables:[16]

- Eligible TFV purchases in HIP participating supermarkets and superstores (EBT data)
- Targeted fruit and vegetable intake (survey data)
- Self-reported monthly total fruit and vegetable expenditures (survey data)

For each of these outcomes, we present analyses using two approaches for measuring SNAP households' food retailer access, as described below.

- In Section 5.1, we use a ***continuous measure*** of distance between a SNAP household's residence and the nearest HIP participating supermarket.
- In Section 5.2, we use distance information to generate ***categorical descriptions*** of the distance to food retailers.

The analysis in this chapter provides some evidence that food retailer access was associated with food spending outcomes as measured in the EBT data. We did not find similar associations for food intake and spending outcomes in the participant surveys. Despite investigating multiple different ways of

---

[16] We also analyzed five additional outcomes. Two of these outcomes, SNAP purchases in all retailers and SNAP purchases in HIP participating supermarkets, were from the EBT data. Three outcomes were from the survey data: targeted vegetable intake, targeted fruit intake, and all fruit and vegetable intake. See Appendix B, Exhibits B.5-B.14.

describing the distances to food retailers we found no evidence that the distance to food retailers modified the HIP impact on food spending or food intake outcomes.

5.1 Continuous Measures of Food Retailer Access

This section describes the associations between continuous measures of distance to food retailers and (a) the level of TFV purchases and consumption, and (b) the HIP impact on these TFV purchases and consumption.

Level of TFV Purchases and Consumption

Analyses that used EBT measures of TFV spending showed evidence of the importance of food retailer access. In a simple model that did not control for distance to food retailers, HIP increased household monthly TFV spending in participating supermarkets by $1.10 (Exhibit 5.1, Model 1). Models that included driving distance to nearest HIP supermarket found a strong association with TFV spending. For every additional mile that a household lived from the nearest HIP participating supermarket, the level of TFV spending in participating supermarkets on average fell by $0.69 (Exhibit 5.1, Model 2).

Because the household's residential location was not randomly assigned, it is not possible to give this finding a causal effect interpretation (as in the HIP evaluation's main impact estimates), but this finding does show that distance to food retailers is associated with the spending outcomes measured with EBT data.

To put some perspective on the magnitude of this estimate, average households spent $12.11 on TFVs each month. Examining the distribution of household distance to the nearest HIP participating supermarket shows that between the 25th and 75th percentile, distance changed from about 0.5 miles to about 2.0 miles, a difference of about 1.5 miles. The estimate of $0.69 per mile suggests that over this range from the 25th to the 75th percentile, monthly TFV expenditures in participating supermarkets would decline by about a dollar or about 8 percent of the average TFV expenditure. [17]

[17] Calculation: 1.5 miles x $0.69 per mile = $1.03.

Exhibit 5.1: Targeted Fruit and Vegetable Purchases ($) in HIP Participating Supermarkets by Distance to HIP Participating Supermarket

	Regression coefficient[a] (SE)		
	Model (1)	**Model (2)**	**Model (3)**
Treatment	1.10*** (0.18)	1.07*** (0.18)	1.09*** (0.25)
Driving distance to nearest HIP supermarket		-0.69*** (0.030)	-0.69*** (0.03)
Distance to nearest HIP-participating supermarket by treatment interaction			-0.01 (0.09)
Constant	27.74*** (0.89)	29.80*** (0.91)	29.80*** (0.91)
R-squared	0.11	0.112	0.112

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.
Standard errors adjusted for non-independence by clustering.
[a] All regression models included additional household demographic covariates: household residential location, household size, gender and age of household head, gender and age of respondent, race/ethnicity of respondent. Model 2 also contains a variable representing the driving distance between the SNAP household and the nearest HIP supermarket. Model 3 contains the interaction of treatment status and the driving distance between the SNAP household and the nearest HIP supermarket.
Source: EBT Transaction Data, pooled across March 2012, July 2012, and October 2012 (average of 42,219 households per month).

Similarly structured analyses, using survey-based food intake outcomes yielded smaller and not statistically significant associations with the distance to food retailers. In a simple model that does not control for distance, HIP increased adult respondents' daily TFV intake by 0.226 cup equivalent (Exhibit 5.2, Model 1). A model that includes distance to the nearest HIP supermarket finds no statistically significant association with TFV intake (Exhibit 5.2, Model 2).

Exhibit 5.2: Targeted Fruit and Vegetable Intake, Cup-Equivalents, by Distance to HIP Participating Supermarket

	Regression coefficient[a] (SE)		
	Model (1)	**Model (2)**	**Model (3)**
Treatment	0.232*** (0.054)	0.230*** (0.054)	0.167* (0.091)
Driving distance to nearest HIP supermarket		-0.015 (0.022)	-0.019 (0.024)
Distance to nearest HIP-participating supermarket by treatment interaction			0.035 (0.044)
Constant	-0.363 (0.323)	-0.314 (0.321)	-0.305 (0.321)
R-squared	0.113	0.113	0.113

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.
Standard errors adjusted for non-independence by clustering.
[a] All regression models included additional household demographic covariates: household residential location, household size, gender and age of household head, gender and age of respondent, race/ethnicity of respondent, baseline fruit and vegetable consumption, and baseline composite scales describing the home food environment, barriers to grocery shopping, and attitudes about consumption of fruits and vegetables. Model 2 also contains a variable representing the driving distance between the SNAP household and the nearest HIP supermarket. Model 3 contains the interaction of treatment status and the driving distance between the SNAP household and the nearest HIP supermarket.
Source: Participant Survey (AM/PM dietary recall module), pooled Round 2 and Round 3 responses, including 10 percent second-day subsamples for each round; (unweighted N=3,890 recalls from 1,996 respondents).

Likewise, analysis using survey-based fruit and vegetable spending did not yield evidence of a relationship with distance to food retailers. In a simple model that does not control for food retailer access, HIP increased self-reported monthly fruit and vegetable spending by $8.35 (Exhibit 5.3, Model 1). A model that includes distance to the nearest HIP supermarket finds no association with fruit and vegetable spending (Exhibit 5.3, Model 2).

Exhibit 5.3: Self-Reported Monthly Total Fruit and Vegetable Expenditures ($), by Distance to HIP Participating Supermarket

	Regression coefficient[a] (SE)		
	Model (1)	Model (2)	Model (3)
Treatment	8.35*** (3.03)	8.37*** (3.04)	5.04 (4.21)
Driving distance to nearest HIP supermarket		0.21 (1.06)	-0.03 (1.20)
Distance to nearest HIP-participating supermarket by treatment interaction			1.82 (1.45)
Constant	-21.82 (21.13)	-22.47 (20.93)	-22.05 (20.92)
R-squared	0.177	0.177	0.178

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.
Standard errors adjusted for non-independence by clustering.
[a] All regression models included additional household demographic covariates: household residential location, household size, gender and age of household head, gender and age of respondent, race/ethnicity of respondent, baseline fruit and vegetable consumption, and baseline composite scales describing the home food environment, barriers to grocery shopping, and attitudes about consumption of fruits and vegetables. Model 2 also contains a variable representing the driving distance between the SNAP household and the nearest HIP supermarket. Model 3 contains the interaction of treatment status and the driving distance between the SNAP household and the nearest HIP supermarket.
Source: Participant Survey (AM/PM dietary recall module), pooled Round 2 and Round 3 responses, including 10 percent second-day subsamples for each round; (unweighted N=3,284 recalls from 1,792 respondents).

HIP Impact on TFV Purchases and Consumption

Using continuous measures of food retailer access, we found no evidence that distance to the nearest participating retailer influenced or modified the HIP impact.

First, in the analysis of TFV spending in participating retailers, there was no interaction between the distance variable and HIP participation (Exhibit 5.1, Model 3). This means that the HIP impact on TFV spending in participating retailers appeared unchanged at increasing distances from the nearest participating retailer.

Second, similarly, there was no interaction between the distance variable and HIP participation in the survey-based measure of TFV intake (Exhibit 5.2, Model 3), nor between the distance variable and HIP participation in the survey-based measure of self-reported fruit and vegetable spending (Exhibit 5.3, Model 3). As with the EBT-based measures, we found no evidence that the distance variable modified the HIP impact on TFV intake or fruit and vegetable spending.

5.2 Categorical Measures of Food Retailer Access

This section has a similar structure to the preceding section, but it focuses on categorical measures of food retailer access rather than continuous measures. Once again, this section describes the associations between food retailer access and (a) the level of TFV purchases and consumption and (b) the HIP impact on these TFV purchases and consumption.

To be sure that our results are robust to diverse methods for describing food retailer access, we constructed our categories in these three ways:[18]

1. ***Whether the closest supermarket participated in HIP***, defined as two mutually exclusive categories: (i) the closest supermarket was HIP participating or (ii) the closest supermarket was not HIP participating

2. ***Whether the household lived within one mile of a HIP supermarket***, defined as three mutually exclusive categories: (i) the household lived within one mile of a HIP participating supermarket; (ii) the household did not live within one mile of a HIP participating supermarket, but did live within one mile of a non-HIP participating supermarket; or (iii) there was no HIP participating or non-HIP participating supermarket within one mile of the household

3. ***Whether the household lived within one mile of a HIP supermarket and/or a high-volume supermarket***, defined as five mutually exclusive categories: (i) the household lived within one mile of a high-volume supermarket that participated in HIP, but not within one mile of a high-volume supermarket that did not participate in HIP; (ii) the household lived within one mile of a high-volume supermarket that did not participate in HIP but not within one mile of a high-volume supermarket that did participate in HIP; (iii) the household lived within one mile of both a HIP participating and non-HIP participating high-volume supermarket; (iv) the household did not live within one mile of any high-volume supermarket, but did live within one mile of a non-high-volume supermarket; or (v) the household did not live within one mile of a supermarket of any kind (HIP participating or non-HIP participating, high-volume or non-high-volume)

Level of TFV Purchases and Consumption

Analyses using categorical descriptions of the distance to food retailers also indicated an association with the level of TFV purchases in participating supermarkets and superstores based on EBT data.

Exhibit 5.4 provides information on the TFV purchases in supermarkets for HIP and non-HIP SNAP households in each category of distance to food retailers. Each of these subgroup analyses support the finding that distance to food retailers was related to the level of TFV spending in participating supermarkets. Using a two-category description of the food retailer access, we find households for whom the closest supermarket participated in HIP, on average, spent $2.13 more of their EBT dollars on TFVs than did households for whom the closest retailer did not participate in HIP (Exhibit 5.4, top tier).

Similarly, in our three-category description of food retailer access, TFV spending was significantly higher among those SNAP households that lived within one mile of a supermarket that participated in HIP compared to households that lived within one mile of a supermarket that did not participate in HIP (Exhibit 5.4, middle tier, and also illustrated graphically in Exhibit 5.5). TFV spending at participating supermarkets was lowest for households that had a non-participating supermarket nearby (within one mile) but not a participating supermarket nearby. TFV spending was also significantly

[18] Each of these categorizations focuses on distance to supermarkets and superstores. In constructing these proxies, distance to convenience stores and grocery stores are not considered.

higher among SNAP household that did not live within one mile of any supermarket as compared to SNAP households that lived within one mile of a non-participating supermarket. It is plausible that households that lived within one mile of a non-participating supermarket were encouraged by proximity to shop in the non-participating supermarket, so their TFV purchases in participating supermarkets were suppressed.

We also find significant differences in the levels of TFV spending using a five category description of distance to food retailers. Here, spending was highest among households that lived within one mile of a high-volume HIP participating supermarket and lowest among households that lived within one mile of high-volume non-participating supermarket but not a high-volume HIP participating supermarket.

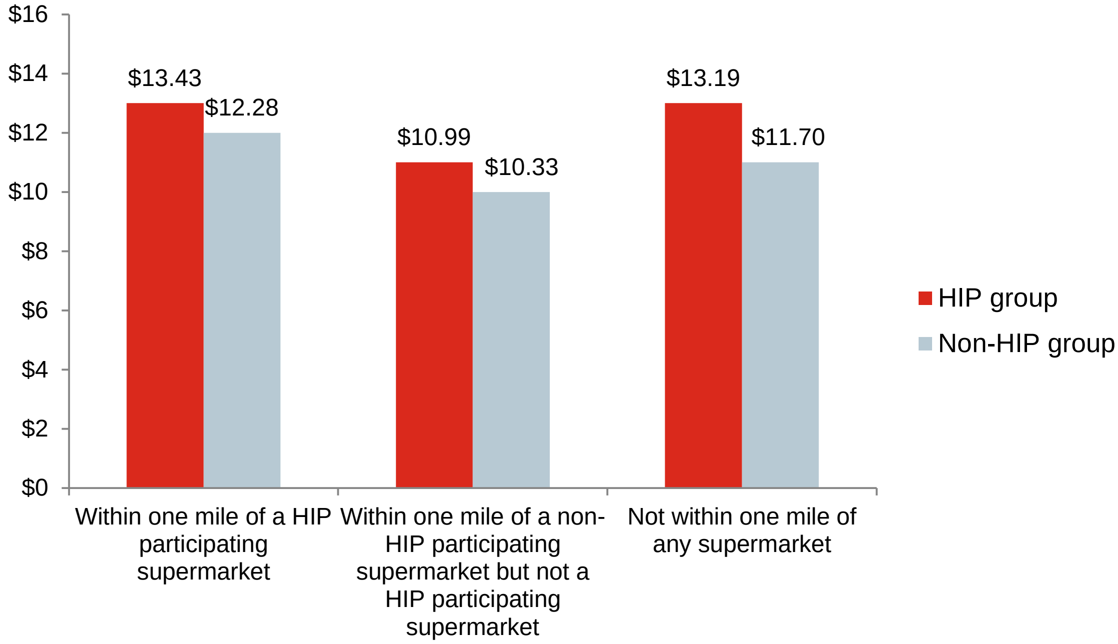
Exhibit 5.4: Impact of HIP on Eligible Fruit and Vegetable Purchases ($) in HIP Participating Supermarkets, by Subgroup

	Regression-adjusted treatment mean (SE)	Regression-adjusted control mean (SE)	Treatment-control impact (P-value)	Difference in impacts (P-value)
Food retailer access: Two-category description (N=46,686 households)				
Closest supermarket is HIP (N=15,451)	13.95 (0.29)	12.74 (0.10)	1.21 (<0.01)***	
Closest supermarket is not HIP (N=31,235)	11.63 (0.19)	10.61 (0.07)	1.02 (<0.01)***	
Difference in levels: Closest HIP, Not Closest HIP (treatment and control combined)				2.13 (<0.01)***
Difference in treatment control impact: Closest HIP, Not Closest HIP				0.19 (0.60)
Food retailer access: Three-category description (N=46,686 households)				
Within one mile of HIP participating supermarket (N=12,122)	13.43 (0.30)	12.28 (0.12)	1.15 (<0.01)***	
Within one mile of non -HIP participating supermarket but not a HIP participating supermarket (N=18,187)	10.99 (0.24)	10.33 (0.09)	0.67 (<0.01)***	
Not within one mile of any supermarket (N=16,377)	13.19 (0.29)	11.70 (0.11)	1.47 (<0.01)***	
P-value for difference in levels across subgroups(treatment and control combined)[a]				(<0.01)***
P-value for difference in treatment control impacts across subgroups [a]				(0.11)
Food retailer access: Five-category description (N=46,686 households)				
Within one mile of high-volume HIP participating supermarket (N=2,965)	13.52 (0.52)	12.70 (0.21)	0.82 (0.14)	
Within one mile of high-volume non-participating supermarket but not a high-volume HIP participating supermarket (N=9,768)	10.59 (0.32)	9.99 (0.13)	0.60 (0.07)*	
Within one mile of both high-volume HIP participating supermarket and non-participating supermarket (N=15,925)	13.04 (0.88)	12.13 (0.36)	0.91 (0.34)	
Not within one mile of a high-volume supermarket but within one mile of non-high-volume supermarket (N=1,651)	12.43 (0.30)	11.40 (0.11)	1.03 (<0.01)***	
Not within one mile of any supermarket (N=16,377)	13.19 (0.29)	11.70 (0.11)	1.47 (<0.01)***	
P-value for difference in levels across subgroups(treatment and control combined) [a]				(<0.01)***
P-value for difference in treatment control impacts across subgroups[a]				(0.43)

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.
Due to rounding, reported impacts (T-C differences) and reported differences in impacts across subgroups may differ from differences between reported regression-adjusted means for the treatment and control groups and subgroups.

Standard errors adjusted for non-independence by clustering.
[a] P-value represents significance level for joint test across all categories.
Source: EBT Transaction Data, pooled across March 2012, July 2012, and October 2012 (total of 46,686 unique households across all three months; average of 42,219 households per month).

Exhibit 5.5 Targeted Fruit and Vegetable Purchases in HIP Participating Supermarkets, by Distance to Food Retailers



Source: EBT Transaction Data, pooled across March 2012, July 2012, and October 2012 (total of 46,686 unique households across all three months, average of 42,219 households per month).

In contrast to the findings regarding EBT-based measures of TFV purchases, we do not observe a consistent relationship between distance to food retailers and either TFV consumption or fruit and vegetable purchases using survey based measures (see Exhibits 5.6 and 5.7 below).

Exhibit 5.6: Impact of HIP on Consumption of Targeted Fruits and Vegetables, by Subgroup

	Regression-adjusted treatment mean (SE)	Regression-adjusted control mean (SE)	Treatment-control impact (P-value)	Difference in impacts (P-value)
Food retailer access: Two-category description (N=3,890 recalls from 1,996 respondents)				
Closest supermarket is HIP (N=665) [a]	1.191 (0.074)	0.937 (0.064)	0.254 (0.009)***	
Closest supermarket is not HIP (N=1,531) [a]	1.128 (0.052)	0.904 (0.040)	0.224 (0.001)***	
Difference in levels: Closest HIP - Not Closest HIP (treatment and control combined)				0.033 (0.653)
Difference in treatment control impact: Closest HIP, Not Closest HIP				0.031 (0.795)
Food retailer access: Three-category description (N=3,890 recalls from 1,996 respondents)				
Within one mile of HIP participating supermarket (N=541) [a]	1.178 (0.084)	0.962 (0.072)	0.216 (0.051)*	
Within one mile of non -HIP participating supermarket but not a HIP participating supermarket (N=938) [a]	1.121 (0.066)	0.926 (0.058)	0.195 (0.026)**	
Not within one mile of any supermarket (N=719) [a]	1.152 (0.075)	0.871 (0.054)	0.281 (0.003)***	
P-value for difference in levels across subgroups(treatment and control combined)				(0.604)
P-value for difference in treatment control impacts across subgroups				(0.802)
Food retailer access: Five-category description (N=3,890 recalls from 1,996 respondents)				
Within one mile of high-volume HIP participating supermarket (N=135) [a]	0.984 (0.137)	1.007 (0.140)	-0.023 (0.908)	
Within one mile of high-volume non-participating supermarket but not a high-volume HIP participating supermarket (N=448) [a]	1.137 (0.092)	0.809 (0.066)	0.328 (0.004)***	
Within one mile of both high-volume HIP participating supermarket and non-participating supermarket (N=713) [a]	1.434 (0.211)	0.702 (0.145)	0.732 (0.004)***	
Not within one mile of a high-volume supermarket but within one mile of non-high-volume supermarket (N=87)	1.131 (0.074)	1.030 (0.074)	0.101 (0.319)	
Not within one mile of any supermarket (N=713) [a]	1.155 (0.075)	0.874 (0.055)	0.281 (0.003)***	
P-value for difference in levels across subgroups(treatment and control combined)				(0.118)
P-value for difference in treatment control impacts across subgroups				(0.090)*

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.
Due to rounding, reported impacts (T-C differences) and reported differences in impacts across subgroups may differ from differences between reported regression-adjusted means for the treatment and control groups and subgroups.

Standard errors adjusted for non-independence by clustering.
[a] The sum of the number of respondents in each subgroup exceeds the total number of respondents for given analysis. This is due to the fact that respondents may be interviewed up to four times across the observation period. A respondent's food retailer environment sometimes changed from one observation to the next due to changes in respondent location, or changes in retailer location or retailer HIP-participation. As a result, an individual respondent may be included in more than one subgroup.
Source: Participant Survey (AM/PM dietary recall module), pooled Round 2 and Round 3 responses, including 10 percent second-day subsamples for each round (unweighted N=3,890 recalls from 1,996 respondents).

Exhibit 5.7: Impact of HIP on Self-Reported Monthly Total Fruit and Vegetable Expenditures ($), by Subgroup

	Regression-adjusted treatment mean (SE)	Regression-adjusted control mean (SE)	Treatment-control impact (P-value)	Difference in impacts (P-value)
Food retailer access: Two-category description(N=3,284 recalls from 1,792 respondents)				
Closest supermarket is HIP (N=583) [a]	79.15 (4.32)	75.99 (4.75)	3.16 (0.62)	
Closest supermarket is not HIP (N=1,358) [a]	79.12 (2.67)	68.72 (2.43)	10.41 (<0.01)***	
Difference in levels: Closest HIP - Not Closest HIP (treatment and control combined)				7.28 (0.17)
Difference in treatment control impact: Closest HIP, Not Closest HIP				-7.25 (0.32)
Food retailer access: Three-category description (N=3,284 recalls from 1,792 respondents)				
Within one mile of HIP participating supermarket (N=479) [a]	73.31 (3.85)	74.32 (5.90)	-1.01 (0.89)	
Within one mile of non -HIP participating supermarket but not a HIP participating supermarket (N=821) [a]	77.47 (3.41)	70.82 (3.37)	6.65 (0.16)	
Not within one mile of any supermarket (N=651) [a]	85.00 (4.28)	68.72 (3.50)	16.27 (<0.01)***	
P-value for difference in levels across subgroups(treatment and control combined)				(0.75)
P-value for difference in treatment control impacts across subgroups				(0.18)
Food retailer access: Five-category description (N=3,284 recalls from 1,792 respondents)				
Within one mile of high-volume HIP participating supermarket (N=120) [a]	68.37 (8.54)	54.45 (5.73)	13.92 (0.17)	
Within one mile of high-volume non-participating supermarket but not a high-volume HIP participating supermarket (N=385) [a]	76.28 (5.10)	75.79 (5.39)	0.49 (0.95)	
Within one mile of both high-volume HIP participating supermarket and non-participating supermarket (N=73)[a]	100.83 (12.05)	78.44 (12.86)	22.38 (0.20)	
Not within one mile of a high-volume supermarket but within one mile of non-high-volume supermarket (N=647)	74.21 (3.23)	71.85 (4.53)	2.36 (0.65)	
Not within one mile of any supermarket (N=651) [a]	84.86 (4.29)	68.56 (3.51)	16.30 (<0.01)***	
P-value for difference in levels across subgroups(treatment and control combined)				(0.03)**
P-value for difference in treatment control impacts across subgroups				(0.25)

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.

Due to rounding, reported impacts (T-C differences) and reported differences in impacts across subgroups may differ from differences between reported regression-adjusted means for the treatment and control groups and subgroups.
Standard errors adjusted for non-independence by clustering.
[a] The sum of the number of respondents in each subgroup exceeds the total number of respondents for given analysis. This is due to the fact that respondents may be interviewed up to four times across the observation period. A respondent's food retailer environment sometimes changed from one observation to the next due to changes in respondent location, or changes in retailer location or retailer HIP-participation. As a result, an individual respondent may be included in more than one subgroup
Source: Participant Survey (AM/PM dietary recall module), pooled Round 2 and Round 3 responses, including 10 percent second-day subsamples for each round (unweighted N=3,284 recalls from 1,792 respondents).

HIP Impact on TFV Purchases and Consumption

As with the continuous measures of distance reported in Section 5.1, we did not find consistent evidence that the distance to food retailers, using categorical measures, suppressed the HIP impact on key outcomes.

First, there was no difference across categories in the magnitude of the HIP impact on TFV purchases in participating supermarkets. For each of the three categories describing food retailer access, families who were eligible to earn the HIP incentive spent a larger amount of their benefits on TFVs in participating supermarkets than did those families who were not eligible (Exhibit 5.4, middle tier, and Exhibit 5.5). However, the size of this difference is relatively consistent across the three categories. In Exhibit 5.5, the HIP/non-HIP difference—that is, the difference between red bars and corresponding grey bars—is similar for SNAP households living in all three distance categories. In each case, the HIP/non-HIP difference in monthly TFV spending in participating supermarkets is approximately $1 to $2, and the difference across categories is not statistically significant.

These results from the survey-based measures were somewhat surprising. Because food retailer access was related to TFV spending in participating supermarkets, one might reasonably expect that food retailer access would influence the HIP impact on self-reported spending and consumption. Yet, on the contrary, the HIP/non-HIP difference in these outcomes appeared not to respond strongly to any of the categorical descriptions we analyzed.

5.3 Discussion

This chapter provides some evidence that a SNAP household's distance to food retailers is associated with its level of TFV spending in participating supermarkets. Households that lived farther away from a HIP participating supermarket, on average, spent fewer benefit dollars on targeted fruits and vegetables than did households that lived in closer proximity to these retailers. However, we find no consistent evidence that the effects of being eligible to earn the HIP incentive differed systematically based on a household's distance-defined access to different types of food retailers. To ensure that this result is not an artifact of how we defined our food retailer access variables, we investigated multiple continuous and categorical methods for describing food retailer access. This finding proved robust: the HIP impact appeared to be similar across diverse food retailer access conditions.

6. Neighborhood Effects

This chapter studies the associations between "neighborhood effects" and SNAP participants' total purchases and their purchases of targeted fruits and vegetables. According to a principle of geospatial analysis, "Everything is related to everything else, but near things are more related than distant things" (Tobler, 1970). Accordingly, SNAP households may be more similar to other SNAP households nearby than they are to SNAP households that live farther away. Geographic clustering in outcome variables could arise from many possible explanatory variables, including variables describing the food retail environment. Some explanatory variables are observed in the EBT data and are included in our regression analyses of the impact of the HIP incentive. Yet, some sources of clustering may be unobserved in the EBT data but remain related to the outcomes.

The ordinary-least-squares (OLS) regression-adjusted findings presented in the HIP Final Report provide unbiased estimates of the impact of HIP on SNAP participants' purchase and consumption of fruits and vegetables. The error terms in those regression models incorporate all unaccounted-for factors (i.e., those not explicitly included in the regression model) that may influence food spending and food intake outcomes. For each SNAP household, these "other factors" may be comparatively similar to those of nearby neighbors and comparatively different than those of SNAP households that live farther away. Hence, the error terms in such regression models may be spatially correlated, meaning that there is more correlation in the errors for people who live closer to each other. Such a correlation implies that unobserved neighborhood characteristics may have inappropriately inflated standard errors, potentially leading to incorrect inferences regarding the impact of the HIP incentive.

The analyses presented in this chapter address the possibility of unobserved "neighborhood effects" in two ways. First, we calculate indices of spatial autocorrelation, which measure whether SNAP participants' purchasing patterns are spatially correlated and, if so, to what extent this is likely to be significant in our models. Indices of spatial autocorrelation measure whether and to what extent like values (i.e., particularly high or low values) of a given variable cluster together spatially. Specifically, these measures estimate the degree to which households' HIP outcomes are more like those of their neighbors than would be expected given a random distribution of those outcome variables across the county.

Second, we re-estimate the regression-adjusted differences between the HIP and non-HIP groups of SNAP participants using a spatial error model, such that spatial correlation is incorporated into the analysis. We use spatial regressions to account for these unobserved neighborhood characteristics and provide a more robust estimate of HIP impacts. Our spatial error model takes the following form:

$$y_i = \beta_0 + \beta_1 HIP_i + \beta_2 ControlVars_i + \delta W\theta + \varepsilon_i$$

where y denotes the outcome of interest, *HIP* is a binary variable that identifies the HIP group, *ControlVars* is a vector of household characteristic and demographic variables that may influence Y, δ denotes the spatial autoregressive parameter, W represents a spatial weights matrix based on the Euclidian distance between households, θ denotes the vector of spatial errors, and ϵ denotes a vector of normally distributed, homoskedastic, and uncorrelated errors. Essentially, the values of δ are produced by the spatial autocorrelation analyses and indicate whether outcomes are more similar for geographically proximate households and more disparate for geographically distant households.

Spatial autocorrelation and spatial regression represent computationally intensive analyses. Given the large number of SNAP households in the monthly EBT files, it was not possible to conduct these analyses pooling data across multiple months as we did with the OLS regression analyses in Chapter 5. Instead, we conducted the spatial regression analyses using a single month of data (July 2012). Second, in order to properly estimate "neighborhood effects," it was necessary to remove from the analyses households that did not have at least eight "neighbors" within a 0.25 mile radius.[19] This reduced our July 2012 sample from the 42,180 SNAP households used in the main analysis to 40,015 SNAP households. The 2,165 SNAP households that do not have at least eight neighbors and are thus excluded from the analyses reside in both the densely populated central and sparsely populated western and eastern regions of Hampden County. We examine the degree to which removing these families from the analyses changes our estimates in Exhibits 6.3 and 6.4 below.

6.1 Indices of Spatial Autocorrelation

To determine whether there are neighborhood effects in Hampden County for SNAP spending in HIP participating stores or TFV spending, we first calculated a *Getis-Ord Gi** statistic for each SNAP household. *Getis-Ord Gi** provides a localized analysis of spatial autocorrelation. It shows for whom HIP outcome levels are spatially correlated, where that spatial autocorrelation occurs, and which types of values (i.e., particularly high or low values) are clustered at a given location.

As Exhibit 6.1 shows, the *Getis-Ord Gi** statistic indicates that the level of TFV spending in HIP participating supermarkets is, to some extent, spatially clustered.[20] In these figures red dots represent "hot spots", that is, households that are both similar to their neighbors and have higher values than would be expected. Blue dots represent "cold spots" or households that are similar to their neighbors and have lower than expected values.

The exhibit also provides some evidence that this clustering may be associated with the proximity of individual SNAP households to different types of retailers. This is particularly evident in Exhibit 6.1

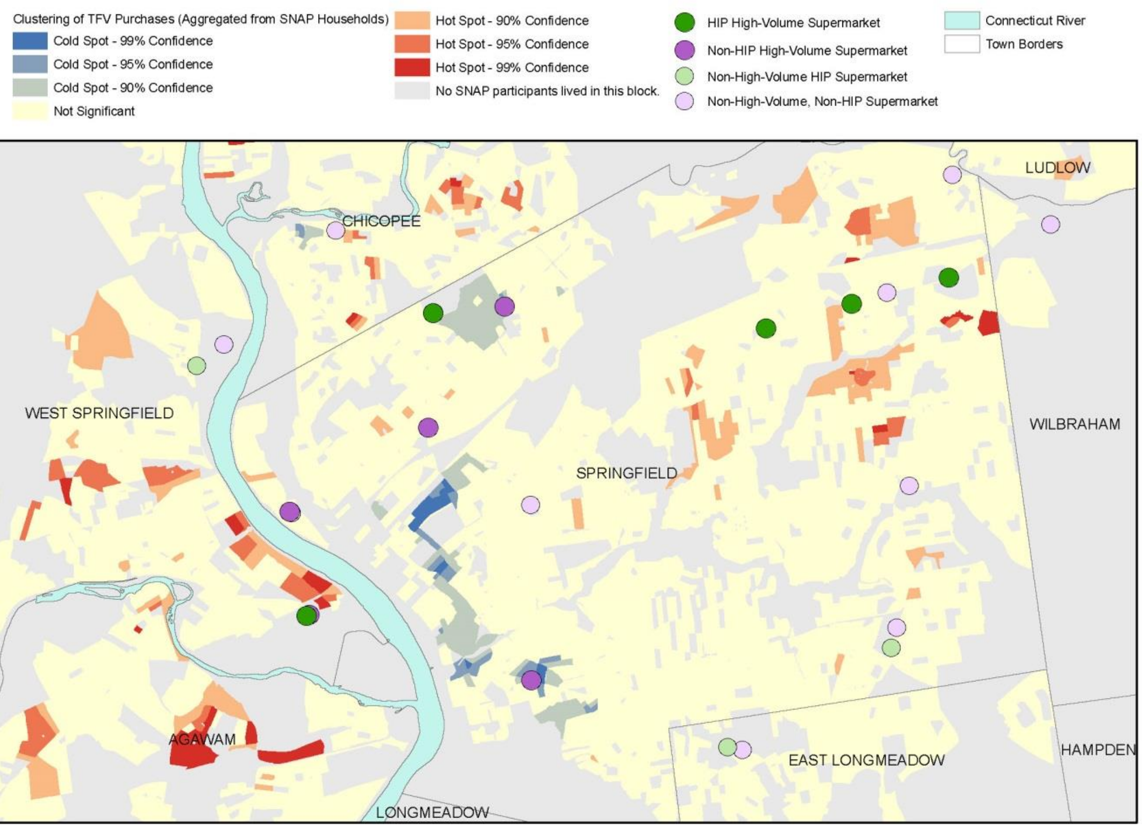
---

[19] Spatial autocorrelation is assessed by comparing the value (e.g., the amount spent on TFVs) of each input (e.g., a SNAP household) to its neighbors. If the result of these comparisons is statistically significantly different than what one would expect given a random geographic distribution of a given value (e.g., the amount a SNAP household spent on TFVs), then spatial error models incorporate this finding into the regression. In effect, this process measures whether a "neighborhood effect" is present, then adjusts the estimates of the regression coefficients and their standard errors accordingly to control for neighborhood effects. However, this process is only meaningful if we can be reasonably certain that the comparisons made by ArcGIS constitute a meaningful estimate of the existence of a neighborhood effect. This requires two things: (1) a restrained geographic area (two people in an urban county that are five miles apart are not neighbors) and (2) a sufficient number of neighbors living in that geographic area such that the sum of neighbor-to-neighbor comparisons is unlikely to yield spurious spatial autocorrelation values (if two people who live near each other behave similarly, this is much more likely to be random than if ten people do). Given that the vast majority of SNAP participants in Hampden County live within a densely populated urban area, a neighborhood-defining distance of greater than 0.25 miles would have increased the number of neighbors involved in some calculations to incalculable extremes given computational memory constraints, in addition to rendering z scores unreliable. A minimum of eight neighbors is a generally accepted standard for minimizing the risk of generating inappropriate autocorrelation values.

[20] We also calculated the statistic for SNAP redemptions at HIP participating supermarkets and found similar results.

in southeastern West Springfield and southwestern Springfield, where high TFV spending values tend to cluster around high-volume, HIP participating supermarkets and low values tend to cluster around non-participating, high-volume supermarkets. However, the results also indicate that food retailer distance measures may not be capturing all of the relevant spatial phenomena affecting SNAP participants' food purchasing behavior, as pockets of high and low values sometimes cluster in ways that appear to be independent of their proximity to a given type of retailer.

Exhibit 6.1: Local Spatial Autocorrelation of TFV Spending in HIP Participating Supermarkets (July 2012)



Note: Getis-Ord Gi* statistics were calculated using SNAP households and then aggregated to the block level. This map should *not* be interpreted as a display of the results of Getis-Ord Gi* calculations that used blocks or block groups as the unit of analysis.

The map of local spatial autocorrelation values suggests that SNAP household TFV spending in HIP participating supermarkets may be systematically higher and lower in certain areas. Whether this is related to the food retail environment or to some other unobserved neighborhood characteristic(s), the possibility remains that this neighborhood variation may not be properly accounted for in standard OLS regression analyses of the impact of the HIP incentive.

To determine whether this is the case, we next calculated a *global* index of spatial autocorrelation (*Moran's I*). This statistic measures aggregate household-level spatial autocorrelation values to indicate whether a given outcome is spatially correlated across the entire sample. Values for *Moran's*

I can range from −1 to +1 (indicating that space and value levels are perfectly correlated, i.e., that high and low values always cluster spatially). A value of zero indicates that there is no spatial relationship in the data.

We calculated this index of spatial autocorrelation for two HIP-related outcomes: (1) total SNAP purchases in HIP participating supermarkets, and (2) eligible fruit and vegetable purchases in HIP participating supermarkets. First, we calculated the degree to which the outcomes are spatially correlated. To measure the degree to which regression covariates and distance measures of food retailer access accounted for any observed spatial autocorrelation, we also calculated the degree to which regression residuals are spatially correlated. We did this three times: (1) using our standard set of regression covariates but excluding our measures of food retailer access, (2) including our standard set of regression covariates and a continuous measure of the distance between a SNAP participant's residence and the nearest HIP participating supermarket, and (3) including our standard set of regression covariates along with our five-category characterization of food retailer access.

We find evidence of statistically significant spatial autocorrelation for both measures of SNAP spending regardless of which covariates are included (Exhibit 6.2).

Exhibit 6.2: Global Indices of Spatial Autocorrelation

Outcome/input variable	Moran's I	p-value
SNAP purchases ($) in HIP participating supermarkets		
Spending at HIP participating supermarkets	0.0280	<0.01***
Residuals for HIP spending regressions, excluding distance measures	0.0200	<0.01***
Residuals for HIP spending regressions, including continuous distance measures	0.0185	<0.01***
Residuals for HIP spending regressions, including five-category distance measures	0.0146	<0.01***
Targeted fruit and vegetables purchases ($) in HIP participating supermarkets		
TFV spending at HIP participating supermarkets	0.0154	<0.01***
Residuals for TFV spending regressions, excluding distance measures	0.0135	<0.01***
Residuals for TFV spending regressions, including continuous distance measures	0.0127	<0.01***
Residuals for TFV spending regressions, including five-category distance measures	0.0108	<0.01***

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.
Source: EBT Transaction Data, July 2012 (40,015 households).

Moran's I values ranged from 0.0146 to 0.0280 for SNAP spending in HIP participating supermarkets and 0.0108 to 0.0154 for fruit and vegetable spending. These data indicate that among SNAP participants, households with similar levels of spending in HIP participating supermarkets and similar levels of targeted fruit and vegetable spending live systematically nearer to one another than would be expected if spending in HIP participating supermarkets was unrelated to geography. More simply, neighborhood effects are apparent. This pattern is statistically significant, and is more evident for total SNAP spending in HIP participating supermarkets than it is for TFV spending. For both outcomes, however, on average across the entire sample, it is not substantively strong. We return to how to think about the magnitude of spatial correlation below in our discussion of the regressions controlling for spatial autocorrelation.

As expected, *Moran's I* values decrease monotonically as we incorporate our explanatory covariates. However, the inclusion of measures of food retailer access does not eliminate statistically significant spatial autocorrelation. For TFV spending, for example, the inclusion of non-spatial covariates accounts for a difference in *Moran's I* values of 0.0019, or 12.3 percent of the observed spatial correlation, while categorical distance measures of food retailer access account for an additional difference in *Moran's I* values of 0.0027, or 17.5 percent of the overall observed spatial autocorrelation. The remaining 70.2 percent of the spatial autocorrelation is unaccounted for, indicating additional, unobserved characteristics about the neighborhoods in which Hampden County SNAP households live that are related to their food purchasing behavior. Similarly, for spending in HIP participating supermarkets, our spatial and non-spatial covariates account for 47.9 percent of spatial autocorrelation, leaving 52.1 percent unaccounted for.

The presence of spatial autocorrelation among HIP outcomes of interest implies that more detailed analyses, such as spatial regression models, may provide a yet more precise estimate of the impact of HIP.

6.2 Spatial Regression

*Getis-Ord Gi** and *Moran's I* statistics presented in the previous section indicate the presence of spatial autocorrelation in both the outcomes of interest and the residuals from regression models, including our set of covariates plus measures of food retailer access. They suggest that some of the unaccounted for variation in our impact estimates is the result of shared variation among neighboring SNAP households.

As noted above, while not correcting for such spatial correlation is standard, failure to correct for it potentially yields underestimates of the true standard errors (and thus inferring statistically significant impacts when they are not truly present). In addition, controlling for spatial auto-correlation has the potential to yield different and better estimates. In this section, we report such regressions corrected for spatial autocorrelation and consider how doing so shifts the standard errors and the estimated impacts.

Exhibits 6.3 and 6.4 below, display the results of spatial regression analyses of two outcome measures—SNAP spending and TFV spending in HIP participating supermarkets. For each exhibit, column (1) displays the coefficients and standard errors for the treatment variable and intercepts from OLS models using the full sample. Column (2) provides the same information from OLS models that use the reduced sample, and column (3) provides the results of spatial error models, also using the reduced sample. Thus, comparing column (2) to column (1) indicates how switching to the narrower sample used for the spatial regressions shifts the estimated impacts and their standard errors, maintaining the conventional OLS—not corrected for spatial autocorrelation—methods. Then, comparing column (3) to column (2) indicates how correcting for spatial autocorrelation shifts the estimated impacts and their standard errors.

In July 2012, for the full sample used in the HIP Final Report, households that received the HIP incentive spent an additional $4.08 of their benefits in HIP participating supermarkets compared to households that did not receive the HIP incentive, controlling for relevant household and demographic characteristics. When we conduct the same OLS analyses on the reduced sample (those households that had at least eight "neighbors" within 0.25 miles of their homes), the estimated treatment effect is slightly smaller ($3.82). This suggests that removing the 2,165 households that did not have at least eight neighbors reduced the estimated impact of HIP. When we analyze these same households using spatial error models, the corresponding treatment effect is again slightly smaller ($3.62).

Similarly, for spending on TFVs in HIP participating supermarkets, the HIP group spent an additional $1.00 on targeted fruits and vegetables. When we conduct the same analyses on the reduced sample, the estimated treatment effect is slightly smaller, $0.92. The estimated treatment effect produced by the spatial error model is trivially smaller, $0.89.

There is no evidence that failure to control for spatial autocorrelation biases our uncorrected estimates of the standard errors. Standard errors are slightly smaller in the spatial analysis sample and smaller still when correcting for spatial autocorrelation. The change in the estimated standard error of the treatment effect with correction for autocorrelation is trivially small; in fact, the standard error shrinks. We conclude that using the uncorrected models is not problematic. Furthermore, the lack of change in the standard errors implies that the level of spatial autocorrelation—according to the crucial dimension, effect on estimated impacts and their standard errors—is small.

Exhibit 6.3: SNAP Purchases ($) in HIP Participating Supermarkets

	Regression coefficient (SE)		
	(1) OLS regression on full sample	(2) OLS regression on spatial analysis sample	(3) Spatial error regression on spatial analysis sample
Treatment	4.08***	3.82*	3.62*
	(1.55)	(1.59)	(1.57)
Constant	259.85***	278.11***	277.65***
	(4.58)	(4.80)	(4.82)
Number of households included in the analyses	42,180	40,015	40,015

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.
Regression models estimated including household demographic covariates.
Source: EBT Transaction Data, July 2012.

Exhibit 6.4 Eligible Fruit and Vegetables Purchases ($) in HIP Participating Supermarkets

	Regression coefficient (SE)		
	(1) OLS regression on full sample	(2) OLS regression on spatial analysis sample	(3) Spatial error regression on spatial analysis sample
Treatment	1.00***	0.92***	0.89***
	(0.29)	(0.23)	(0.23)
Constant	31.68***	32.78***	32.38***
	(0.68)	(0.70)	(0.71)
Number of households included in the analyses	42,180	40,015	40,015

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.
Regression models estimated including household demographic covariates.
Source: EBT Transaction Data, July 2012.

6.3 Discussion

In summary, we do find evidence of “neighborhood effects” on households’ spending on fruits and vegetables. Households that lived nearer to one another were more similar in their purchases of fruits and vegetables than they were to families that lived farther away. The ordinary least squares regression analyses used in the main HIP report accounted for some but not all of these neighborhood effects. Correcting for this shared variation using spatial regression analyses did not substantively change the estimated impact, the estimated precision of impact, or the inferences about statistical significance.

7. Conclusions

This report has presented the results of the HIP Spatial Analysis. The Spatial Analysis is a follow-on to the earlier HIP evaluation (Bartlett et al., 2014), which estimated the impact of HIP without accounting for any spatial phenomena that may have affected food purchases and the impact of HIP. The HIP Spatial Analysis further investigates how the food retail environment influenced Hampden County SNAP households' fruit and vegetable purchases in general and the HIP impact estimates in particular.

Specifically, the HIP Spatial Analysis addresses the following two research questions posed by USDA's Food and Nutrition Service (FNS):

1. How did relative physical access to HIP participating stores affect household shopping patterns, HIP incentive earnings, and targeted fruit and vegetable (TFV) consumption?
2. Are there "neighborhood effects" with respect to consumer responses to the incentive? Did HIP participants in close proximity to one another (and thereby with shared food retail environment characteristics) exhibit similarities in response to the pilot?

The balance of this chapter summarizes the findings for each research question.

1. How did relative physical access to HIP participating stores affect household shopping patterns, HIP incentive earnings, and targeted fruit and vegetable (TFV) consumption?

The HIP Spatial Analysis investigated whether the distance to food retailers affected fruit and vegetable purchases and dietary intake outcomes in Hampden County. In particular, we sought to uncover whether the HIP impact on fruit and vegetable outcomes was suppressed in some areas of the county because of limited access to participating retailers where participants could purchase a variety of targeted fruits and vegetables.

Using multiple measurement approaches and definitions of food retailer access, we found no evidence that the distance to food retailers affected the impact of HIP on fruit and vegetable outcomes.

To some extent, this finding was a surprise. We found that supermarkets—and HIP participating supermarkets in particular—were concentrated in the densely populated central region of Hampden County. In the more rural western regions of the county, distances to the nearest supermarket could be quite large. Yet, precisely because most SNAP participants lived in the densely populated central cities of Springfield, Holyoke, and Chicopee, 65 percent of SNAP participants had a supermarket no more than one mile away from home. The median distance to the nearest supermarket was less than one mile (0.81 miles) and 75 percent of SNAP households lived within 1.15 miles of a supermarket.

Consistent evidence showed that SNAP participants in Hampden County did not prioritize shopping at a SNAP retailer within one mile from home, even when such a retailer was available. The median distance to the retailer where SNAP participants actually spent most of their benefits—almost always a supermarket—was 1.51 miles. Most SNAP participants bypassed a nearby retailer to reach a preferred retailer at a greater distance. We found that most SNAP benefits were spent in just 13 individual high-volume retail store locations, which were patronized by SNAP participants residing both nearby and farther away.

This does not mean distance to retailers had no influence. For example, all else being equal, for each additional mile from home to the nearest supermarket, the mean monthly purchases of targeted fruits and vegetables at participating supermarkets fell by $0.69. But we found no evidence of an interaction between distance to retailers and HIP's impact on patterns of SNAP spending and food intake. Whether we investigated distance to the nearest supermarket (as a continuous variable), or the presence of a supermarket within one mile, or the presence of a HIP participating supermarket within one mile, or even the presence of one of the select high-volume retailers in the nearby neighborhood, the difference in outcomes between HIP and non-HIP SNAP participants was unchanged.

2. ***Are there "neighborhood effects" with respect to consumer responses to the incentive? Did HIP participants in close proximity to one another (and thereby with shared food retail environment characteristics) exhibit similarities in response to the pilot?***

It is clear that the neighborhood you live in affects the food you buy. Spending outcomes for one household were more similar to spending patterns for another household that lived nearby and more dissimilar to spending patterns for another household that lived farther away. Yet, these spatial correlations were not as strong as we might have expected.

Moreover, we sought but did not find strong evidence that these spatial correlations could be fully explained by measurable characteristics of food retailer access.

Finally, correcting the impact regressions for spatial autocorrelation did not substantively shift the estimated impact, the estimated precision of impact (the standard errors), or the inferences about statistical significance. We conclude that the degree of spatial autocorrelation was small.

Discussion

This report used the EBT and survey data collected for the HIP evaluation to consider the spatial dimension of food purchasing and dietary intake behavior and to consider whether previous, conventional estimates were biased because they ignored spatial autocorrelation. Clearly, there are interesting spatial aspects of food shopping behavior. Conventional approaches to characterizing the food retailer access do not seem to capture how SNAP households actually shop. We show that households passed over nearby food retailers—even supermarkets and superstores—in favor of a small set of more distant retailers, which may have a combination of prices and products that low-income consumers seek. Our analyses suggest new approaches to characterizing food retailers. Applying these analytic approaches to other data sources seems a fruitful direction for future work on environmental factors affecting dietary quality.

Despite the clear evidence of a spatial component to shopping patterns, failure to control for spatial autocorrelation does not appear to have caused us to underestimate the impacts of HIP presented in the HIP Final Report.

References

Andreyeva, T., Luedicke, J., Middleton, A. E., Long, M. W., and Schwartz, M. B. (2012). Positive influence of the Revised Special Supplemental Nutrition Program for Women, Infants, and Children food packages on access to healthy foods. *Journal of the Academy of Nutrition and Dietetics* 112, 850–858.

Bartlett, S., Klerman, J., Olsho, L., Logan, C., Blocklin, M., Beauregard, M., Enver, A., Owens, C., and Melthem, M. (2014). *Evaluation of the Healthy Incentives Pilot (HIP): Final Report.* Alexandria, VA: U.S. Department of Agriculture, Food and Nutrition Service.

Block, J. P., Christakis, N. A., O'Malley, A. J., and Subramanian, S. V. (2011). Proximity to food establishments and body mass index in the Framingham Heart Study offspring cohort over 30 years. *American Journal of Epidemiology* 174, 1108–1114.

Currie, J., Vigna, S. D., Moretti, E., and Pathania, V. (2010). The effect of fast food restaurants on obesity and weight gain. *American Economic Journal: Economic Policy* 2(3), 32–63.

IOM (Institute of Medicine) and NRC (National Research Council). (2013). Supplemental Nutrition Assistance Program: Examining the evidence to define benefit adequacy. Washington, DC: The National Academies Press.

Larson, N., Story, M., and Nelson, M. (2009). Neighborhood environments: Disparities in access to healthy foods in the U.S. *American Journal of Preventive Medicine* 36, 74–81.

Powell, L.M., Slater, S., Mirtcheva, D., Bao, Y., and Chaloupka, F.J. (2007). Food store availability and neighborhood characteristics in the United States. *Preventative Medicine* 44, 189-195. doi:10.1016/j.ypmed.2006.08.008

The Reinvestment Fund (TRF). (2011). Limited supermarket access analysis: Summary of TRF's methodology. Available at: http://www.trfund.com/resource/downloads/policypubs/LSAMethodology2011.pdf

Rose, D., and Richards, R. (2004). Food store access and household fruit and vegetable use among participants in the U.S. Food Stamp Program. *Public Health Nutrition* 7(8), 1081–1088.

Tobler, W. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46(2), 234-240.

U.S. Department of Agriculture, Economic Research Service. (2009). Access to affordable and nutritious food: Measuring and understanding food deserts and their consequences, report to Congress. Washington, DC. Available at: www.ers.usda.gov/Publications/AP/AP036/.

U.S. Department of Agriculture, Economic Research Service. (2013). Food Access Research Atlas (FARA). Available at: http://www.ers.usda.gov/data-products/food-access-research-atlas.aspx

Ver Ploeg, M., Breneman, V., Dutko, P., Williams, R., Snyder, S., Dicken, C., and Kaufman, P. (2012). *Access to affordable and nutritious food: Updated estimates of distance to economic supermarkets using 2010 data*. Washington, DC: U.S. Department of Agriculture, Economic Research Service.

Wilde, P. E., Llobrera, J., and Valpiani, N. (2013.). Household food expenditures and obesity risk. *Current Obesity Reports* 1(3), 123–133. DOI:10.1007/s13679-012-0022-y.

Wilde, P. E., Llobrera, J., and Ver Ploeg, M. (2014.). Population density, poverty, and food retail access in the United States: An empirical approach. *International Food and Agribusiness Management Review* 17(Special Issue A), 171–186. Available at: http://www.ifama.org/i4a/pages/index.cfm?pageid=3512

Appendix A: Technical Methodology

This appendix provides further detail about how we conducted the spatial analyses presented in this report. It proceeds in three sections, ordered chronologically to mirror both the order of our work and the order of this report: (1) geocoding, (2) distance calculations, and (3) spatial analysis.

A.1 Geocoding

This section provides further detail about the three geocoded files described in Chapter 2: the SNAP case file, the retailer location file, and the survey location file.

Construction of SNAP Case Location Files

Our first step in this process was to geocode the home addresses of the 54,861 households in Hampden County that were active SNAP households at the end of July 2011 (and thus eligible to receive benefits in August 2011).[21] The Massachusetts Department of Transitional Assistance provided Abt Associates with monthly files on these households from July 2011 to December 2012. We merged these 18 separate files into a single data file, in which each row represented a unique combination of individual case file ID and household address. The resulting dataset contained a total of 74,910 unique address observations, including multiple addresses for some households during the 18-month period of this study (Exhibit A.1).

Exhibit A.1: Number of Residences: SNAP Households From Case Files

# of residences	Households	%	Cum. households	Cum. %
1	39,443	71.90	39,443	71.90
2	11,877	21.65	51,320	93.55
3	2,697	4.92	54,017	98.46
4	640	1.17	54,657	99.63
5	168	0.31	54,825	99.93
6	30	0.05	54,855	99.99
7	6	0.01	54,861	100.00

For the period July 2011–December 2012.

We then imported address information for each unique combination of case file ID and household address into ArcMap spatial analysis software. We generated longitude and latitude coordinates from information on household street number, street name, city, state, and ZIP code, using ESRI's World Geocode Service (ArcGIS Online) TA US 10 geolocator. We plotted each address's latitude and longitude onto a map of the United States, projected using Lambert's conformal conic projection centered on Massachusetts. We used this projection throughout our geospatial analysis, as it preserves distance calculations for geographic bodies, such as Massachusetts, that are elongated along an east-west axis.

This preliminary round of geocoding provided accurate information on the latitude and longitude for 94.5 percent (70,793) of case file unique ID addresses. We took two additional steps to geocode the

[21] This number is 234 households less that the population of households that were randomly assigned to either receive or not receive the HIP intervention due to exits from SNAP between mid-July, when random assignment occurred, and the end of July.

remaining 5.5 percent (4,117) of the case files. Our first step was to process unmatched case file addresses through a second geolocator, ESRI's StreetMap Premium software. We then reformatted the remaining unmatched case file addresses to better align with the specifications of the geocoding software. For example, we identified cases in which information such as "3RD FLR" preceded the street name or street number of an address, or a place name such as the name of an apartment building was concatenated with an address' street number. We corrected these incorrectly formatted addresses and then reprocessed the information using the StreetMap Premium geolocator. This provided accurate latitude and longitude information on an additional 400 unique case file ID-addresses, 41 of which were located within Hampden County.[22] Exhibit A.2 displays the final results of our geocoding, which shows that almost 99 percent of Hampden County households were geocoded.

Exhibit A.2: SNAP Case File Geocoding

Geocoding results	Frequency	% of unique ID-addresses	% of unique ID-addresses within Hampden County [a]
Matched	71,193	95.0	-
Within Hampden County	69,510	-	98.7
Unmatched	3,717	5.0	-
Within Hampden County	933	-	1.3
Total	74,910	100.0	
*Within Hampden County**	70,443	-	100.0

[a] Unmatched case file ID-addresses were considered "within Hampden County" if the ZIP code or town/city name associated with a given record matched a ZIP code or town/city name found within the county. Note that while ZIP code boundaries are nearly identical to county boundaries in Hampden County, they do not align exactly, especially in the county's northwest corner. Very few records were identified as being "within Hampden County" using ZIP code alone, however, so the impact of this discrepancy is small.

Construction of the Retailer Location File

Throughout the pilot, Abt received monthly Retailer Electronic Data Exchange (REDE) files that included all Massachusetts SNAP-authorized retailers from FNS.

Our process for identifying latitude and longitude information for SNAP retailers was similar to the approach we used for geocoding case files, detailed above. We drew address information on all Massachusetts SNAP-authorized retailers from monthly REDE files for the months (August 2011–December 2012) provided by FNS. Connecticut retailers' addresses were drawn from a single file containing a single observation for every SNAP-authorized retailer in Tolland and Hartford Counties, also provided by FNS. As with the household files, we merged these 17 separate files into a single data file, in which each row represented a unique combination of retailer ID and retailer address.

The resulting dataset contained a total of 6,981 unique retailer ID-addresses representing 6,666 distinct retailers across the Commonwealth of Massachusetts and Tolland and Hartford Counties in Connecticut. We imported address information for each unique combination of retailer ID and

[22] All households lived in Hampden County at the time of random assignment to HIP and non-HIP groups. Approximately 8.5 percent of the 54,861 households in the evaluation sample lived outside Hampden County for one or more months during the evaluation period. They remained eligible to earn HIP incentives. These households were, however, excluded from the spatial analysis for the months they resided outside Hampden County.

address into ArcMap, then generated longitude and latitude coordinates from information on retailer street number, street name, city, state, and ZIP code, using ESRI's World Geocode Service (ArcGIS Online) TA US 10 geolocator. We then plotted each address's latitude and longitude onto a map of the United States, again projected using Lambert's conformal conic projection centered on Massachusetts.

This preliminary round of geocoding provided accurate information on the latitude and longitude for 97.8 percent of retailer unique ID-addresses. We focused our efforts to recode the remaining 2.2 percent (169) on the 17 retailers that were located within 10 miles of Hampden County. We began by manually choosing correct retailer addresses out of lists of candidate locations identified by the geolocator, when possible and appropriate. We then ran unmatched addresses through a second U.S. Address geolocator(ESRI's StreetMap Premium software). Finally, we cross-referenced the remaining, unmatched retailer addresses with publicly available SNAP location data from http://snap-retailers.findthedata.org/. This website provided retailer locations through Google Maps (maps.google.com). This manual approach provided the most accurate method to locate the appropriate latitude and longitude values of farmers' markets, which are often located in places without a formal address. We then input the latitude and longitude of unmatched retailers into the ArcMap software. These steps allowed us to geocode 100 percent of SNAP retailers within 10 miles of Hampden County (see Exhibit A.3).

Exhibit A.3: SNAP Retailer Geocoding

Geocoding results	Frequency	% of unique ID-addresses	% of unique ID-addresses within 10 mi of Hampden County*
Matched	6,829	97.8	-
Within 10 mi. of Hampden County	941	-	100.0
Unmatched	152	2.2	-
Total	6,981	100.0%	

The results of SNAP retailer geocoding, using ESRI's World Geocoding Service (ArcGIS Online) US 10 geolocator; individual internet searches for unmatched retailer addresses, and a second geolocator enhanced results.

An inclusion distance of 10 miles for retailers satisfied findings from food access literature as to the distance the average person in a rural area travels to shop for food, as well as limited the food retailer environment so as not to include the Worcester, Pittsfield, or Greenfield metropolitan areas. Most SNAP participants in Hampden County lived in urban areas, but we took the cautious approach of choosing a buffer distance large enough to be adequate for most rural residents in Hampden County. Rose and Richards (2004) reported that 27 percent of a nationally representative sample of SNAP participants purchased most of their food at a retailer farther than five miles from their homes, which set a lower bound for inclusion distance. A 2012 report by the USDA Economic Research Service clarified further, showing that 93 percent of rural residents in the United States live within 10 miles of a supermarket. A 10-mile inclusion distance appears adequate even for a rural resident right on the county boundary, and more than adequate for residents in the interior of Hampden County.

Construction of the Survey Location Files

As part of the primary HIP evaluation, Abt Associates administered a detailed survey to a subsample of Hampden County SNAP participants. Although the SNAP participants included in the survey sample were also included in the case files, we geocoded their household locations separately using the information provided in the survey files. This was done because there may have been differences between the household address in the case file records and the household address in the survey records. Address information on the participant survey files was updated as each of the three survey waves was administered. From this, we created a household location file based on survey addresses, which contained 6,302 unique addresses across all three waves of data collection. Latitude and longitude data attached to this file were used for analyses that were focused exclusively on survey data.

Our procedure for geocoding the survey data followed the same process as for case files and retailers. Geocoding the survey data was simple and straightforward, as address information recorded through the survey was well aligned with the specifications of the geocoding software. Specifically, supplemental address information (such as apartment or floor number) for participant addresses was entered into a separate field, which allowed for simple and efficient pre-processing. As town and city names were recorded by interviewers over the phone, we also checked for and corrected misspellings of cities or towns in the Hampden County area.

These changes allowed us to accurately geocode the household location of all but three non-missing records across all three rounds of survey data collection. We manually geocoded the three remaining unmatched addresses using Google Maps. Overall, participant survey geocoding returned latitude and longitude data for 98.9 percent of records. See Exhibit A.4 below for details.

Exhibit A.4: Participant Surveys

Data collection round	Geocoding results	Frequency	% of unique ID-addresses
1	Matched	2,734	98.20
	Within Hampden County	2,717	97.59
	Unmatched	49	1.76
	Total	2,784	100.00
2	Matched	1,982	99.20
	Within Hampden County	1,965	98.35
	Unmatched	14	0.70
	Total	1,998	100.00
3	Matched	1,515	99.67
	Within Hampden County	1,504	98.95
	Unmatched	5	0.33
	Total	1,520	100.00

The results of geocoding participant survey addresses using ESRI's World Geocoding Service (ArcGIS Online) US 10 geolocator. Results are presented for each of the three rounds of data collection.

A.2 Distance Calculations

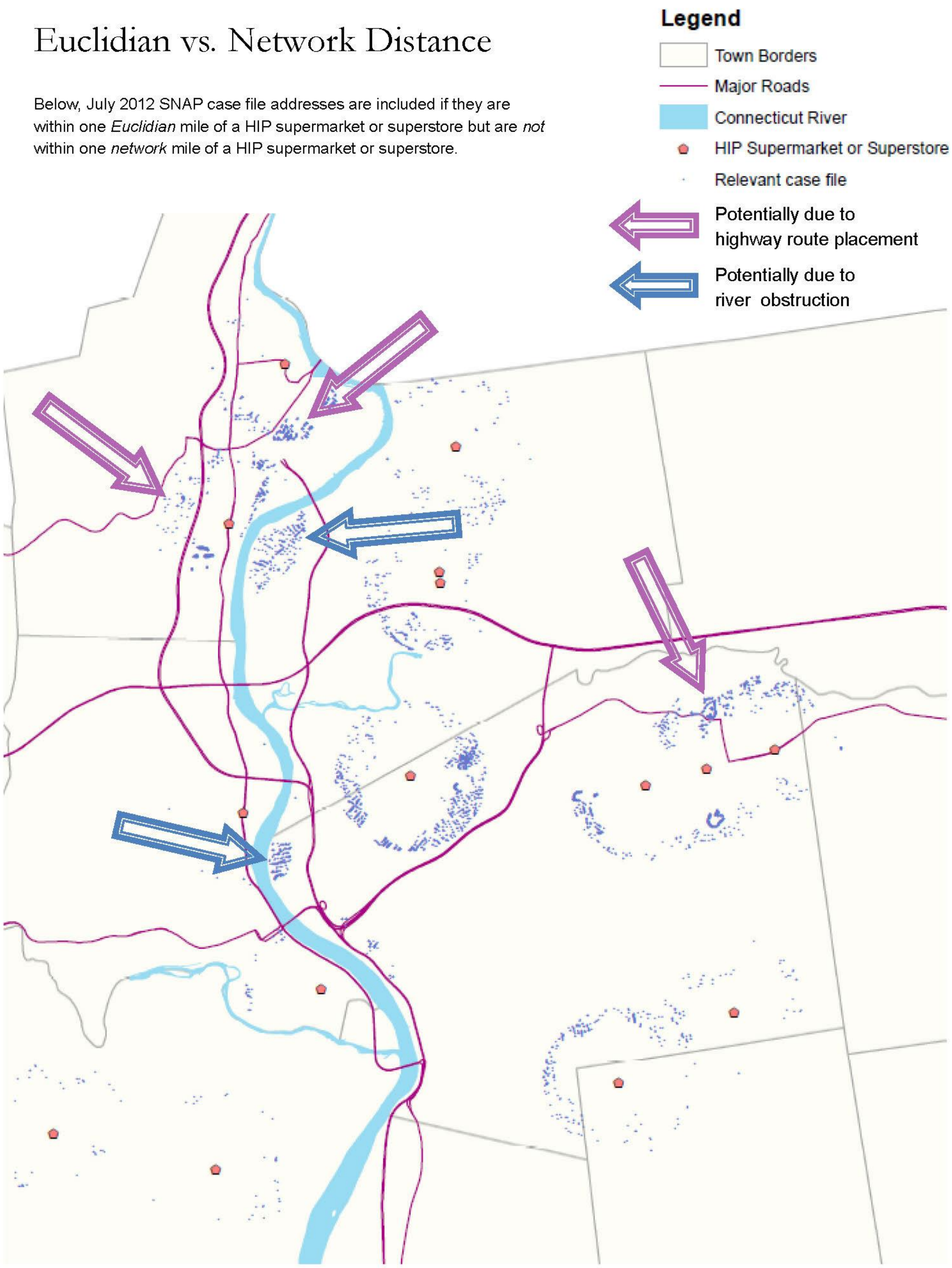
This section provides further detail about the process within ArcGIS software that was used to calculate distance measures, discussed primarily in Chapter 4 and used in analyses found in Chapters 5 and 6.

Overview of Distance Calculation Methods

We computed the distances between SNAP households and SNAP retailers using ArcGIS. Specifically, we computed both the distance to the nearest retailer of a given type and the distance to retailers at which SNAP participants redeemed their benefits. We performed both of these computations for six types of retailers: any retailer; any HIP participating retailer; any supermarket/superstore; any HIP participating supermarket/superstore; any high-volume HIP participating retailer; and any high-volume, non-HIP participating retailer. We computed three versions of each of these measurements: one measuring distance as the crow flies ("Euclidian" distance) between a household and a retailer, one measuring the distance someone could *drive* using a roadway network, and a third measuring the distance someone could *walk* using a roadway network. Walking and driving distances are calculated differently based on road-specific travel restrictions. For instance, pedestrians are not allowed on highways, while service roads/walking paths are not open to privately owned vehicle traffic. We compared walking and driving distances and found that they were nearly identical to one another for Hampden County SNAP households.

Using multiple methodologies to calculate the distance between SNAP participants and retailers for the month of July allowed us to compare all of these measures as a case study that would inform the rest of our analyses. We found that Euclidean and driving distances differed substantially from one another, with driving distances being longer than Euclidean distances. The difference between Euclidean and driving distances varied across Hampden County. This difference was particularly striking in the densely populated central portion of the county located to the immediate east and west of the Connecticut River. The closest supermarket, as measured by Euclidean distance, was for some residents of this region located on the opposite side of the river. Since the river can only be traversed using one of the county's three bridges, Euclidean-derived distances overestimate access to supermarkets for many Hampden County SNAP participants. Exhibit A.5 presents a visual depiction of this phenomenon. In this report, we therefore use only network-derived measures of distance to different types of food retailers.

Exhibit A.5: Euclidian vs. Network Distance



Calculating Distances in ArcGIS

Calculating Euclidian distances using ArcGIS software was relatively straightforward. For the nearest retailer, July retailer shapefiles containing all retailers, only HIP retailers, only supermarkets and superstores, only HIP supermarkets and superstores, only HIP high-volume retailers, and only non-HIP high-volume retailers were joined to a July SNAP participant shapefile using the "Spatial Join (Analysis)" tool.[23] This calculated distances in meters, which were then converted to miles in STATA. For distances to the retailers at which SNAP participants actually spent their benefits, we compiled a data file using EBT transaction data and retailer and case file latitudes and longitudes: each SNAP participant was represented by as many rows as retailers at which they spent some portion of their SNAP benefits in July 2012. Each of those rows contained information about the retailer at which the participant spent their benefits, as well as the amount they spent there, the geographic coordinates of their household, and the geographic coordinates of the retailer. In this format, ArcGIS software was able to read in each pair of latitudes and longitudes (one set from the retailer, one set from the participant) and calculate the Euclidian distance between the two, taking into account the curvature of the earth. This was accomplished using the "XY to Line (Data Management)" tool, which returns both the calculated distance and an ID variable. As only one ID variable is returned, we concatenated each row's Client ID with the given retailer's FNS number, separated by a dash, and then separated them again during post-processing. These distances could then be merged onto case file data by each participant's Client ID.

Calculating distances along a roadway network ("network" or "driving" distances) was somewhat more complex. First, we purchased roadway network information from ESRI for Massachusetts and Connecticut as part of the StreetMap Premium software. This network dataset was comprehensive, including data on roadway topography, travel restrictions, and data from an algorithm developed by ESRI known as "hierarchy," which models driver preference. To calculate distances as described below.

1. We uploaded the roadway network onto ArcMap, whereupon a new origin-destination (OD) cost matrix layer was created using the Network Analyst extension for each type of retailer (any retailer, any HIP participating retailer, any supermarket or superstore, any HIP participating supermarket or superstore, any HIP high-volume retailer, and any non-HIP high-volume retailer). Each of these was then replicated to calculate distances for different modes of travel (walking or driving).

2. We changed each layer's analysis settings depending on whether we were calculating walking distance or driving distance. In both cases, we restricted the number of destinations to find to one (i.e., the closest), set the measurement units to length in miles, and changed the ID ("Name") variables-to-find for the Origins and Destinations layers to participant Client ID and FNS number, respectively.

---

23 Parameters for running the spatial analysis included searching for the "closest" retailer while simultaneously specifying that we wanted ArcGIS to also return the distance to that retailer, accomplished by inputting an optional name for the distance variable.

a. For walking, travel restrictions limited routes to streets on which 1) walking was allowed 2) while avoiding private roads. [24]

b. For driving, travel restrictions limited routes to streets on which 1) driving was allowed 2) with a private automobile 3) while avoiding private roads.

3. After each matrix was solved, the concatenated ID variable output by ArcMap was split to enable merging by client ID onto case file data.

To calculate distances between survey respondent locations and the nearest retailer of various types, we repeated the process described above using survey data files and a unique identifier for each household. We calculated these distances on a month-by-month basis for each month between March 2012 and November 2012. The number of retailers participating in HIP varied across the intervention period, and many SNAP households changed addresses during this time. To ensure that the distances reflected the appropriate food retailer environment, we calculated each household's distance to various types of retailers based on the food retailer environment and respondent address at the time of their survey response.

The ArcGIS Network Analyst tool is not designed to readily calculate distances between specific pairs of geographic coordinates. To address this limitation we wrote a program using the *Python* programming language to calculate the distance between each SNAP participant in the EBT files and the retailers at which they spent their SNAP benefits. The script does the following:

1. Loads the latitude and longitude of each participant's home address (the Origin) and the latitude and longitude of a given retailer at which they shopped (the Destination) into a new, temporary OD cost matrix.

2. Calculates the distance between the two and exports the distance and a concatenated ID variable into a .dbf file.

3. Repeats this process for every row, skipping transactions that occurred at retailers farther than 10 miles outside of Hampden County, as these were unlikely to reflect regular purchasing patterns of Hampden County residents.

For simplicity, individuals whose home address' latitude and longitude were not near a public road (this amounted to one or two SNAP participants, depending on the month) were excluded from the distance calculations.

---

[24] Walking distance to the nearest retailer via roadway network was impossible to calculate for 21 SNAP participants. The addresses of these participants are technically along a highway; some of their front doors literally open onto an off-ramp. The roadway network is programmed to recognize major highways as walking-prohibited; as such, the computer does not recognize that walking out of these participants' houses is possible. As these participants make up less than a tenth of a percent of our sample, however, this did not make a significant impact.

A.3 Spatial Analysis

This section provides further details regarding our approach to calculating the spatial autocorrelation statistics and running the spatial error regressions presented in Chapter 6. This section also provides justification for our decision to use a spatial error model.

Conceptualizing Spatial Relationships

As described in Chapter 6, spatial autocorrelation is assessed by comparing the value (e.g., the amount spent on TFVs) of each SNAP household to its neighbors. If the result of these comparisons is statistically significantly different than what one would expect given a random geographic distribution of a given value (e.g., the amount a SNAP household spent on TFVs), then spatial regression models incorporate this information into the regression. In effect, this process measures whether a "neighborhood effect" is present, then adjusts the estimates of the regression coefficients and their standard errors accordingly to control for neighborhood effects. There are several ways to do this. Determining which analyses to run requires theorizing the way neighborhood effects are likely to work in the analysis, with (broadly) two possibilities. The first is that people living nearer to one another *affect one another's behavior* and do so proportional to their proximity to one another. This kind of conceptualization is appropriate when modeling phenomena such as the spread of knowledge about a given topic through a given community, or the spread of a particular plant species in cases where the speed with which that species spreads is dependent on the density of that species' growth. The second possibility is that people living nearer to one another are affected in similar ways by their *shared environment*. In such a case, the purpose of the spatial analysis is to capture unexplained spatial variation based on factors such as racial segregation, shared transportation infrastructure, or food retail environment. The latter is the focus of our analysis here.

Given the assumption that spatial autocorrelation would result from SNAP participants' shared environment, spatial error models (which model the second possibility) are preferable to spatial lag models (which model the first). Further, the determination of who ought to be considered a "neighbor" of a given SNAP household should be based on whether one SNAP household lives within the same, clearly defined geographic area as another (i.e., they share a "neighborhood"). Moreover, any definition of who should be a neighbor of a given SNAP household should be based on (1) a *restrained* geographic area (two people in an urban county that are five miles apart are not neighbors) and (2) a sufficient number of neighbors living in that geographic area such that the sum of neighbor-to-neighbor comparisons are unlikely to yield spurious spatial autocorrelation values (if two people who live near each other behave similarly, this is much more likely to be random than if 10 people do).

Therefore, to estimate spatial autocorrelation, we used a "fixed distance" of 0.25 miles to define neighbors for each SNAP household; that is, any other SNAP household within 0.25 miles of a given SNAP household was considered that SNAP household's "neighbor."[25] Each relationship between a SNAP household and his/her neighbors was weighted equally with a value of one (i.e., we did not

---

[25] Given that the vast majority of SNAP participants in Hampden County live within a densely populated urban area, a neighborhood-defining distance of greater than 0.25 miles would have increased the number of neighbors involved in some spatial autocorrelation calculations to incalculable extremes given computational memory constraints, in addition to rendering z scores unreliable.

standardize by row). A minimum of eight neighbors is also a generally accepted standard for minimizing the risk of generating inappropriate autocorrelation values. Thus, in order to properly estimate "neighborhood effects," it was necessary to remove from our analyses households that did not have at least eight "neighbors" within a 0.25 mile radius using R.[26] (This reduced our July 2012 sample from the 42,180 SNAP households used in the main analysis to 40,015 SNAP households.) Households were excluded from these analyses by flagging those households with less than eight rows of data in the spatial weights matrix, then removing the relationships that went into their spatial autocorrelation calculations from the matrix. The revised spatial weights matrix was then used to run analyses on the reduced sample of 40,015, effectively excluding SNAP households with an insufficientnumber of neighbors from the autocorrelation calculations while still enabling all participants to act as each other's neighbors.

Spatial Analyses

As described in Chapter 6, we first calculated a *Getis-Ord Gi** statistic for each SNAP household using the "Hot Spot Analysis (Getis-Ord Gi*) (Spatial Statistics)" tool in ArcGIS. We calculated this statistic twice, once for each of two HIP outcomes drawn from the EBT data files. These two outcomes were

(1) SNAP redemptions at HIP participating supermarkets, and

(2) SNAP spending on TFVs at HIP participating supermarkets.

Next, we calculated *Moran's I* for eight different dependent variables using the "Spatial Autocorrelation (Global *Moran's I*)" tool, also in ArcGIS. We repeated this calculation for the two outcome variables listed above as well as a total of six additional times on regression residuals. These six calculations were based on two sets of three regression residuals, one for each of the outcomes above: (1) using our standard set of regression covariates but excluding our measures of household food retailer access; (2) including our standard set of regression covariates and a continuous measure of the distance between a SNAP participant's residence and the nearest HIP participating supermarket; and (3) including our standard set of regression covariates along with our five-category characterization of food retailer access. In all cases, the reduced (i.e., limited to participants who had at least eight neighbors within 0.25 miles of their residences) spatial weights matrix was used.

We then used R to run (1) ordinary OLS regressions on the full sample, (2) ordinary OLS regressions on the reduced sample (i.e., using the eight-neighbor-limited spatial weights matrix), and (3) spatial error regressions on the reduced sample using the "LU" method.[27]

---

26 R is a software programming language also known as GNU S. For more information on R, please refer to http://www.r-project.org/.

27 The LU method, or the "Pace & Barry" approach, is an alternate sparse matrix decomposition approach without updating. As our files were very large, the standard optimization method (a.k.a. the Ord approach) was unfeasible given computational memory constraints. For more information, see Pace, R.K. & Barry, R. (1997). Sparse spatial autoregressions. *Statistics & probability letters* 33 (3), 291-297; or LeSage, J.P. & Pace, R.K. (2009) *Introduction to Spatial Econometrics*, Boca Raton, FL: Chapman & Hall/CRC

Appendix B: Results of Additional Analyses from Chapters 4 and 5

This appendix contains analyses referenced in Chapter 4 regarding factors that are associated with spending in high-volume supermarkets and with the probability of shopping primarily in a retailer other than a supermarket or superstore (Section B.1). This appendix also contains exhibits discussed in Chapter 5 showing regression analyses that estimate: (1) spending levels and (2) the impact of HIP when incorporating distance measures of food retailer access (i) as a continuous variable (Section B.2) or (ii) as categorical variables (Section B.3).

B.1 Models Predicting SNAP Spending in High-volume Supermarkets and SNAP Household by Primary Retailer Type (Chapter 4)

Exhibit B.1: Ordinary Least Squares Regression Predicting the Percentage of SNAP Purchases in a High-volume Supermarket

	Regression coefficient (SE)
Treatment indicator	0.43 (0.39)
Distance to the nearest high-volume supermarket	-3.50*** (0.07)
Household size=1 (Household size=6 EXCLUDED CATEGORY)	-3.35*** (0.94)
Household size=2 (Household size=6 EXCLUDED CATEGORY)	0.50 (0.78)
Household size=3 (Household size=6 EXCLUDED CATEGORY)	2.90*** (0.76)
Household size=4 (Household size=6 EXCLUDED CATEGORY)	3.66*** (0.78)
Household size=5 (Household size=6 EXCLUDED CATEGORY)	-2.60*** (0.85)
Respondent race/ethnicity indicator, Hispanic (White EXCLUDED CATEGORY)	10.09*** (0.38)
Respondent race/ethnicity indicator, Black (White EXCLUDED CATEGORY)	6.23*** (0.49)
Respondent race/ethnicity indicator, Other (White EXCLUDED CATEGORY)	-1.50** (0.62)
Head of household 16 to 30 years (Head of household 55 + years EXCLUDED CATEGORY)	-2.11*** (0.55)
Head of household 31 to 40 years (Head of household 55 + years EXCLUDED CATEGORY)	-1.20** (0.54)
Head of household 41 to 54 years (Head of household 55 + years EXCLUDED CATEGORY)	-0.80* (0.48)
Head of household is female (Head of household 55 + years EXCLUDED CATEGORY)	4.04*** (0.34)
Location of residence is Springfield	12.36*** (0.48)
Location of residence is Chicopee	-0.74 (0.46)
Children in household	0.78 (0.54)
Elderly in household	0.16 (0.57)

	Regression coefficient (SE)
Wave 2 household	0.06 (0.32)
Wave 3 household	0.35 (0.32)
Household head is disabled	-0.47 (0.32)
Spanish spoken in household	0.64* (0.34
Income is $0 (Income $1,084+ EXCLUDED CATEGORY)	-2.14*** (0.43)
Income between $1 and $787 (Income $1,084+ EXCLUDED CATEGORY)	-0.96** (0.41)
Income between $788 and $1083 (Income $1,084+ EXCLUDED CATEGORY)	0.45 (0.43)
Benefit size between $1 and $160 (Benefit size $201+ EXCLUDED CATEGORY)	-1.03*** (0.52)
Benefit size between $161 and $200 (Benefit size $201+ EXCLUDED CATEGORY)	<0.01 (0.53)
Constant	45.23*** (1.12)
R-squared	0.21

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.
Standard errors adjusted for non-independence by clustering.
Source: EBT Transaction Data, pooled across March 2012, July 2012, and October 2012 (average of 42,219 households per month).

Certain household types were more likely to shop primarily at retailers other than a supermarket or superstore. We estimated three sets of models to investigate this issue (see Exhibits B.2, B.3, and B.4). First, we estimated logistic regression models of the probability that most (more than half of) a household's SNAP expenditures in a given month were at a supermarket/superstore. Such logistic regression models are the conventional approach to modelling binary outcomes. Second, we estimated linear probability models (i.e., the linear probability model on the binary outcome) of the same outcome. Unlike the estimates from the logistic regression model, these estimates can be directly interpreted as a change in the probability of most SNAP expenditures in a supermarket/superstore. The estimates reported in the text were produced by linear probability models. Third, we estimated linear regression models of the fraction of expenditures in supermarkets/superstores. This is a continuous dependent variable. The source for these analyses was EBT Transaction Data, pooled across March, July, and October 2012. Standard errors in these analyses were adjusted for non-independence by clustering.

These analyses suggest that the following groups are less likely to shop primarily at a supermarket/superstore: households in which Spanish was spoken (5 percentage points more than non-Spanish-speaking households); non-white households (Hispanic households were 1 percentage point, black households 5 percentage points, and households that identified as some other race 11 percentage points more likely than white households); households with no income (2 percentage points more than those with incomes greater than $1,084 per month); and households headed by a person with a disability (4 percentage points more than the households headed by a non-disabled person). Conversely, female-headed households (6 percentage points less than male-headed household) were *less* likely to shop primarily at a non-supermarket/superstore. Models of the percent of expenditures in a supermarket/superstore (rather than the probability of spending more than half of SNAP at a supermarket/superstore) show similar results.

Exhibit B.2: Linear Probability for Having a Primary Retailer That Is Not a Supermarket

	Regression coefficient (SE)
Treatment indicator	>-0.01 (0.004)
Person lives within one mile (Driving) of a supermarket or superstore	0.02*** (0.003)
Interaction of Disability status and Person lives within one mile (Driving) of a supermarket or superstore	-0.01*** (0.005)
Household size=1 (Household size=6 EXCLUDED CATEGORY)	0.02** (0.009)
Household size=2 (Household size=6 EXCLUDED CATEGORY)	-0.01 (0.008)
Household size=3 (Household size=6 EXCLUDED CATEGORY)	-0.03*** (0.008)
Household size=4 (Household size=6 EXCLUDED CATEGORY)	-0.03*** (0.008)
Household size=5 (Household size=6 EXCLUDED CATEGORY)	-0.02* (0.009)
Respondent race/ethnicity indicator, Hispanic (White EXCLUDED CATEGORY)	0.01*** (0.003)
Respondent race/ethnicity indicator, Black (White EXCLUDED CATEGORY)	0.05*** (0.005)
Respondent race/ethnicity indicator, Other (White EXCLUDED CATEGORY)	0.11*** (0.007)
Head of household 16 to 30 years (Head of household 55 + years EXCLUDED CATEGORY)	0.03*** (0.005)
Head of household 31 to 40 years (Head of household 55 + years EXCLUDED CATEGORY)	0.03*** (0.005)
Head of household 41 to 54 years (Head of household 55 + years EXCLUDED CATEGORY)	0.02*** (0.005)
Head of household is female (Head of household 55 + years EXCLUDED CATEGORY)	-0.06*** (0.003)
Location of residence is Springfield	0.01** (0.004)
Location of residence is Chicopee	-0.01* (0.004)
Children in household	-0.02*** (0.005)
Elderly in household	-0.02*** (0.005)
Wave 2 household	>-0.01 (0.003)
Wave 3 household	>-0.01 (0.003)
Household head is disabled	0.05*** (0.004)
Spanish spoken in household	0.05*** (0.004)
Income is $0 (Income $1,084+ EXCLUDED CATEGORY)	0.03*** (0.004)

	Regression coefficient (SE)
Income between $1 and $787 (Income $1,084+ EXCLUDED CATEGORY)	0.02*** (0.004)
Income between $788 and $1083 (Income $1,084+ EXCLUDED CATEGORY)	-0.01* (0.004)
Benefit size between $1 and $160 (Benefit size $201+ EXCLUDED CATEGORY)	0.01*** (0.005)
Benefit size between $161 and $200 (Benefit size $201+ EXCLUDED CATEGORY)	>0.01 (0.005)
Constant	0.10*** (0.011)
R-squared	0.037

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.
Standard errors adjusted for non-independence by clustering.
Source: EBT Transaction Data, pooled across March 2012, July 2012, and October 2012 (average of 42,219 households per month).

Exhibit B.3: Odds that a SNAP Household's Primary Retailer Is Not a Supermarket

	Regression coefficient (SE)
Treatment indicator	0.98 (0.030)
Person lives within one mile (Driving) of a supermarket or superstore	1.21*** (0.043)
Interaction of Disability status and Person lives within one mile (Driving) of a supermarket or superstore	0.83*** (0.037)
Household size=1 (Household size=6 EXCLUDED CATEGORY)	1.08 (0.093)
Household size=2 (Household size=6 EXCLUDED CATEGORY)	0.88* (0.066)
Household size=3 (Household size=6 EXCLUDED CATEGORY)	0.69*** (0.052)
Household size=4 (Household size=6 EXCLUDED CATEGORY)	0.74*** (0.058)
Household size=5 (Household size=6 EXCLUDED CATEGORY)	0.84** (0.074)
Respondent race/ethnicity indicator, Hispanic (White EXCLUDED CATEGORY)	1.10*** (0.034)
Respondent race/ethnicity indicator, Black (White EXCLUDED CATEGORY)	1.54*** (0.055)
Respondent race/ethnicity indicator, Other (White EXCLUDED CATEGORY)	2.13*** (0.088)
Head of household 16 to 30 years (Head of household 55 + years EXCLUDED CATEGORY)	1.27*** (0.052)
Head of household 31 to 40 years (Head of household 55 + years EXCLUDED CATEGORY)	1.19*** (0.048)
Head of household 41 to 54 years (Head of household 55 + years EXCLUDED CATEGORY)	1.15*** (0.039)
Head of household is female (Head of household 55 + years EXCLUDED CATEGORY)	0.67*** (0.016)
Location of residence is Springfield	1.06* (0.034)
Location of residence is Chicopee	0.94** (0.032)
Children in household	0.86*** (0.039)
Elderly in household	0.85*** (0.035)
Wave 2 household	1.00 (0.025)
Wave 3 household	0.98 (0.025)
Household head is disabled	1.55*** (0.061)
Spanish spoken in household	1.52*** (0.042)
Income is $0 (Income $1,084+ EXCLUDED CATEGORY)	1.33*** (0.048)
Income between $1 and $787 (Income $1,084+ EXCLUDED CATEGORY)	1.16*** (0.041)
Income between $788 and $1083 (Income $1,084+ EXCLUDED CATEGORY)	0.97 (0.035)

	Regression coefficient (SE)
Benefit size between $1 and $160 (Benefit size $201+ EXCLUDED CATEGORY)	1.15*** (0.050)
Benefit size between $161 and $200 (Benefit size $201+ EXCLUDED CATEGORY)	1.05 (0.047)
Constant	0.10*** (0.010)

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.
Standard errors adjusted for non-independence by clustering.
Source: EBT Transaction Data, pooled across March 2012, July 2012, and October 2012 (average of 42,219 households per month).

Exhibit B.4: Ordinary Least Squares Regression Predicting the Percentage of SNAP Purchases ($) in a Supermarket

	Regression coefficient (SE)
Treatment indicator	0.17 (0.32)
Person lives within one mile (Driving) of a supermarket or superstore	-1.89*** (0.28)
Interaction of Disability status and Person lives within one mile (Driving) of a supermarket or superstore	1.76*** (0.44)
Household size=1 (Household size=6 EXCLUDED CATEGORY)	1.25* (0.73)
Household size=2 (Household size=6 EXCLUDED CATEGORY)	3.47*** (0.60)
Household size=3 (Household size=6 EXCLUDED CATEGORY)	4.73*** (0.57)
Household size=4 (Household size=6 EXCLUDED CATEGORY)	3.96*** (0.60)
Household size=5 (Household size=6 EXCLUDED CATEGORY)	2.26*** (0.65)
Respondent race/ethnicity indicator, Hispanic (White EXCLUDED CATEGORY)	-3.17*** (0.30)
Respondent race/ethnicity indicator, Black (White EXCLUDED CATEGORY)	-6.12*** (0.38)
Respondent race/ethnicity indicator, Other (White =EXCLUDED CATEGORY)	-9.61*** (0.55)
Head of household 16 to 30 years (Head of household 55 + years EXCLUDED CATEGORY)	-6.27*** (0.47)
Head of household 31 to 40 years (Head of household 55 + years EXCLUDED CATEGORY)	-4.14*** (0.47)
Head of household 41 to 54 years (Head of household 55 + years EXCLUDED CATEGORY)	-2.79*** (0.41)
Head of household is female (Head of household 55 + years EXCLUDED CATEGORY)	5.29*** (0.30)
Location of residence is Springfield	-1.91*** (0.32)
Location of residence is Chicopee	0.02 (0.33)
Children in household	1.73*** (0.42)
Elderly in household	2.93*** (0.46)
Wave 2 household	-0.25 (0.27)
Wave 3 household	-0.08 (0.27)
Household head is disabled	-5.46*** (0.38)
Spanish spoken in household	-5.65*** (0.29)
Income is $0 (Income $1,084+ EXCLUDED CATEGORY)	-5.05*** (0.38)
Income between $1 and $787 (Income $1,084+ EXCLUDED CATEGORY)	-3.08*** (0.33)
Income between $788 and $1083 (Income $1,084+ EXCLUDED CATEGORY)	-1.20*** (0.33)

	Regression coefficient (SE)
Benefit size between $1 and $160 (Benefit size $201+ EXCLUDED CATEGORY)	-0.26 (0.39)
Benefit size between $161 and $200 (Benefit size $201+ EXCLUDED CATEGORY)	0.54 (0.42)
Constant	84.70*** (0.85)
R-squared	0.058

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.
Standard errors adjusted for non-independence by clustering.
Source: EBT Transaction Data, pooled across March 2012, July 2012, and October 2012 (average of 42,219 households per month).

B.2 Additional Analyses Using a Continuous Distance Measure (Chapter 5)

Exhibit B.5: SNAP Purchases in All Retailers ($) by Distance to HIP Participating Supermarket

	Regression coefficient[a] (SE)		
	Model (1)	**Model (2)**	**Model (3)**
Treatment	4.42*** (1.30)	4.41*** (1.30)	5.92*** (1.94)
Driving distance to nearest HIP supermarket		-0.24 (0.26)	-0.15 (0.28)
Distance to nearest HIP participating supermarket by treatment interaction			-0.86 (0.81)
Constant	666.33*** (7.836)	667.06*** (7.86)	666.89*** (7.86)
R-squared	0.63	0.63	0.623

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.
Standard errors adjusted for non-independence by clustering.
[a] All regression models included additional household demographic covariates: household residential location, household size, gender and age of household head, gender and age of respondent, race/ethnicity of respondent. Model 2 also contains a variable representing the driving distance between the SNAP household and the nearest HIP supermarket. Model 3 contains the interaction of treatment status and the driving distance between the SNAP household and the nearest HIP supermarket.
Source: EBT Transaction Data, pooled across March 2012, July 2012, and October 2012 (average of 42,219 households per month).

Exhibit B.6: SNAP Purchases in HIP Participating Supermarkets ($) by Distance to HIP Participating Supermarket

	Regression coefficient[a] (SE)		
	Model (1)	**Model (2)**	**Model (3)**
Treatment	4.02*** (1.27)	3.78*** (1.26)	4.46** (1.86)
Driving distance to nearest HIP supermarket		-6.15*** (0.24)	-6.11*** (0.25)
Distance to nearest HIP participating supermarket by treatment interaction			-0.39 (0.772)
Constant	272.20*** (6.05)	290.61*** (6.09)	290.53*** (6.09)
R-squared	0.32	0.33	0.33

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.
Standard errors adjusted for non-independence by clustering.
[a] All regression models included additional household demographic covariates: household residential location, household size, gender and age of household head, gender and age of respondent, and race/ethnicity of respondent. Model 2 also contains a variable representing the driving distance between the SNAP household and the nearest HIP supermarket. Model 3 contains the interaction of treatment status and the driving distance between the SNAP household and the nearest HIP supermarket.
Source: EBT Transaction Data, pooled across March 2012, July 2012, and October 2012 (average of 42,219 households per month).

Exhibit B.7: Targeted Vegetable Intake, Cup-Equivalents, by Distance to HIP Participating Supermarket

	Regression coefficient[a] (SE)		
	Model (1)	**Model (2)**	**Model (3)**
Treatment	0.131*** (0.032)	0.130*** (0.032)	0.105* (0.054)
Driving distance to nearest HIP supermarket		-0.011 (0.010)	-0.013 (0.011)
Distance to nearest HIP participating supermarket by treatment interaction			0.014 (0.023)
Constant	-0.214 (0.196)	-0.178 (0.197)	-0.175 (0.197)
R-squared	0.096	0.097	0.096

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.
Standard errors adjusted for non-independence by clustering.
[a] All regression models included additional household demographic covariates: household residential location, household size, gender and age of household head, gender and age of respondent, and race/ethnicity of respondent. Model 2 also contains a variable representing the driving distance between the SNAP household and the nearest HIP supermarket. Model 3 contains the interaction of treatment status and the driving distance between the SNAP household and the nearest HIP supermarket.
Source: Participant Survey (AM/PM dietary recall module), pooled Round 2 and Round 3 responses, including 10 percent second-day subsamples for each round (unweighted N=3,890 recalls from 1,996 respondents).

Exhibit B.8: Targeted Fruit Intake, Cup-Equivalents, by Distance to HIP Participating Supermarket

	Regression coefficient[a] (SE)		
	Model (1)	**Model (2)**	**Model (3)**
Treatment	0.100*** (0.038)	0.100*** (0.038)	0.062 (0.065)
Driving distance to nearest HIP supermarket		-0.004 (0.017)	-0.007 (0.019)
Distance to nearest HIP participating supermarket by treatment interaction			0.021 (0.034)
Constant	-0.149 (0.236)	-0.136 (0.235)	-0.130 (0.235)
R-squared	0.068	0.068	0.069

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.
Standard errors adjusted for non-independence by clustering.
[a]All regression models included additional household demographic covariates: household residential location, household size, gender and age of household head, gender and age of respondent, race/ethnicity of respondent, baseline fruit and vegetable consumption, and baseline composite scales describing the home food environment, barriers to grocery shopping, and attitudes about consumption of fruits and vegetables. Model 2 also contains a variable representing the driving distance between the SNAP household and the nearest HIP supermarket. Model 3 contains the interaction of treatment status and the driving distance between the SNAP household and the nearest HIP supermarket.
Source: Participant Survey (AM/PM dietary recall module), pooled Round 2 and Round 3 responses, including 10 percent second-day subsamples for each round (unweighted N=3,890 recalls from 1,996 respondents).

Exhibit B.9: All Fruit and Vegetable Intake, Cup-Equivalents, by Distance to HIP Participating Supermarket

	Regression coefficient[a] (SE)		
	Model (1)	Model (2)	Model (3)
Treatment	0.319***	0.316***	0.282**
	(0.080)	(0.081)	(0.127)
Driving distance to nearest HIP supermarket		-0.026	-0.029
		(0.027)	(0.030)
Distance to nearest HIP participating supermarket by treatment interaction			0.019
			(0.055)
Constant	1.117**	1.203**	1.208**
	(0.503)	(0.506)	(0.506)
R-squared	0.103	0.103	0.103

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.
Standard errors adjusted for non-independence by clustering.
[a] All regression models included additional household demographic covariates: household residential location, household size, gender and age of household head, gender and age of respondent, race/ethnicity of respondent, baseline fruit and vegetable consumption, and baseline composite scales describing the home food environment, barriers to grocery shopping, and attitudes about consumption of fruits and vegetables. Model 2 also contains a variable representing the driving distance between the SNAP household and the nearest HIP supermarket. Model 3 contains the interaction of treatment status and the driving distance between the SNAP household and the nearest HIP supermarket.
Source: Participant Survey (AM/PM dietary recall module), pooled Round 2 and Round 3 responses, including 10 percent second-day subsamples for each round (unweighted N=3,890 recalls from 1,996 respondents).

B.3 Additional Analyses Using a Categorical Distance Measure (Chapter 5)

Exhibit B.10: Total SNAP Purchases ($) in All Retailers, by Subgroup—Pooled Months

	Regression-adjusted treatment mean (SE)	Regression-adjusted control mean (SE)	Treatment-control impact (P-value)	Difference in impacts (P-value)
Food retailer access: Two-category description (N=46,686 households)				
Closest supermarket is HIP (N=15,451)	268.98 (1.98)	264.92 (0.76)	4.06 (0.06)*	
Closest supermarket is not HIP (N=31,235)	270.28 (1.41)	265.67 (0.56)	4.62 (<0.01)***	
Difference in levels: Closest HIP and not closest HIP (treatment and control combined)				-0.75 (0.41)
Difference in treatment control impact: Closest HIP and not closest HIP				-0.56 (0.82)
Food retailer access: Three-category description (N=46,686 households)				
Within one mile of HIP participating supermarket (N=12,122)	268.71 (2.26)	265.53 (0.84)	3.19 (0.184)	
Within one mile of non-HIP participating supermarket but not a HIP participating supermarket (N=18,187)	271.00 (1.85)	265.52 (0.73)	5.48 (0.01)***	
Not within one mile of any supermarket (N=16,377)	269.45 (1.96)	265.23 (0.81)	4.22 (0.04)**	
P-value for difference in levels across subgroups (treatment and control combined)[a]				(0.96)
P-value for difference across subgroups[a]				(0.72)
Food retailer access: Five-category description (N=46,686 households)				
Within one mile of high-volume HIP participating supermarket (N=2,965)	264.05 (4.03)	262.23 (1.66)	1.83 (0.67)	
Within one mile of high-volume non-participating supermarket but not a high-volume HIP participating supermarket (N=9,768)	267.41 (2.53)	264.22 (1.10)	3.19 (0.24)	
Within one mile of both high-volume HIP participating supermarket and non-participating supermarket (N=15,925)	276.79 (7.52)	265.78 (2.70)	11.01 (0.17)	
Not within one mile of a high-volume supermarket but within one mile of non-high-volume supermarket (N=1,651)	272.13 (2.19)	266.84 (0.85)	5.29 (0.02)**	
Not within one mile of any supermarket (N=16,377)	269.54 (1.96)	265.30 (0.81)	4.24 (0.04)**	
P-value for difference in levels across subgroups (treatment and control combined)[a]				(0.15)
P-value for difference across subgroups[a]				(0.85)

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.

Due to rounding, reported impacts (T-C differences) and reported differences in impacts across subgroups may differ from differences between reported regression-adjusted means for the treatment and control groups and subgroups.
Standard errors adjusted for non-independence by clustering.
[a] P-value represents significance level for joint test across all categories.
Source: EBT Transaction Data, pooled across March 2012, July 2012, and October 2012 (total of 46,686 unique households across all three months; average of 42,219 households per month).

Exhibit B.11: Total SNAP Purchases ($) in HIP Participating Supermarkets, by Subgroup—Pooled Months

	Regression-adjusted treatment mean (SE)	Regression-adjusted control mean (SE)	Treatment-control impact (P-value)	Difference in impacts (P-value)
Food retailer access: Two-category description (N=46,686 households)				
Closest supermarket is HIP (N=15,451)	151.18 (1.90)	144.71 (0.75)	6.47 <0.01***	
Closest supermarket is not HIP (N=31,235)	122.64 (1.38)	120.09 (0.53)	2.56 (0.08)*	
Difference in levels: Closest HIP and not closest HIP (treatment and control combined)				24.63 (<0.01)***
Difference in treatment control impact: Closest HIP and not closest HIP				3.91 (0.10)
Food retailer access: Three-category description (N=46,686 households)				
Within one mile of HIP participating supermarket (N=12,122)	148.36 (2.12)	142.05 (0.83)	6.31 (<0.01)***	
Within one mile of non -HIP participating supermarket but not a HIP participating supermarket (N=18,187)	116.59 (1.84)	116.65 (0.69)	-0.06 (0.98)	
Not within one mile of any supermarket (N=16,377)	137.10 (1.92)	130.94 (0.79)	6.16 (<0.01)***	
P-value for difference in levels across subgroups (treatment and control combined)[a]				(<0.01)***
P-value for difference across subgroups[a]				(0.03)**
Food retailer access: Five-category description (N=46,686 households)				
Within one mile of high-volume HIP participating supermarket (N=2,965)	150.58 (3.94)	147.36 (1.61)	3.23 (0.45)	
Within one mile of high-volume non-participating supermarket but not a high-volume HIP participating supermarket (N=9,768)	111.20 (2.43)	112.84 (1.01)	-1.64 (0.52)	
Within one mile of both high-volume HIP participating supermarket and non-participating supermarket (N=15,925)	128.19 (5.90)	120.89 (2.42)	7.30 (0.25)	
Not within one mile of a high-volume supermarket but within one mile of non-high-volume supermarket (N=1,651)	137.09 (2.14)	132.45 (0.82)	4.64 (0.04)**	
Not within one mile of any supermarket (N=16,377)	137.10 (1.92)	130.94 (0.79)	6.16 <0.01***	
P-value for difference in levels across subgroups (treatment and control combined)[a]				(<0.01)***
P-value for difference across subgroups[a]				(0.18)

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.

Due to rounding, reported impacts (T-C differences) and reported differences in impacts across subgroups may differ from differences between reported regression-adjusted means for the treatment and control groups and subgroups.
Standard errors adjusted for non-independence by clustering.
[a] P-value represents significance level for joint test across all categories.
Source: EBT Transaction Data, pooled across March 2012, July 2012, and October 2012 (total of 46,686 unique households across all three months, average of 42,219 households per month).

Exhibit B.12: Differences in Impacts of HIP on Consumption of Targeted Vegetables, Cup-Equivalents, by Subgroup

	Regression-adjusted treatment mean (SE)	Regression-adjusted control mean (SE)	Treatment-control impact (P-value)	Difference in impacts (P-value)
Food retailer access: Two-category description (N=3,890 recalls from 1,996 respondents)				
Closest supermarket is HIP (N=665) [a]	0.642 (0.051)	0.404 (0.032)	0.237 (<0.001)***	
Closest supermarket is not HIP (N=1,531) [a]	0.554 (0.029)	0.464 (0.027)	0.091 (0.021)**	
Difference in levels: Closest HIP - Not Closest HIP (treatment and control combined)				-0.059 (0.161)
Difference in treatment control impact: Closest HIP, Not Closest HIP				0.147 (0.043)**
Food retailer access: Three-category description (N=3,890 recalls from 1,996 respondents)				
Within one mile of HIP participating supermarket (N=541) [a]	0.606 (0.051)	0.441 (0.037)	0.165 (0.009)***	
Within one mile of non -HIP participating supermarket but not a HIP participating supermarket (N=938) [a]	0.553 (0.040)	0.483 (0.043)	0.071 (0.217)	
Not within one mile of any supermarket (N=719) [a]	0.590 (0.042)	0.410 (0.031)	0.179 (0.001)***	
P-value for difference in levels across subgroups(treatment and control combined)[b]				(0.432)
P-value for difference in treatment control impacts across subgroups[b]				(0.358)
Food retailer access: Five-category description (N=3,890 recalls from 1,996 respondents)				
Within one mile of high-volume HIP participating supermarket (N=135) [a]	0.499 (0.084)	0.442 (0.067)	0.058 (0.590)	
Within one mile of high-volume non-participating supermarket but not a high-volume HIP participating supermarket (N=448) [a]	0.512 (0.053)	0.428 (0.043)	0.084 (0.206)	
Within one mile of both high-volume HIP participating supermarket and non-participating supermarket (N=87) [a]	0.796 (0.142)	0.247 (0.076)	0.549 (0.001)***	
Not within one mile of a high-volume supermarket but within one mile of non-high-volume supermarket (N=713) [a]	0.591 (0.045)	0.521 (0.053)	0.070 (0.288)	
Not within one mile of any supermarket (N=719) [a]	0.590 (0.042)	0.412 (0.031)	0.179 (0.001)***	
P-value for difference in levels across subgroups(treatment and control combined)[b]				(0.120)
P-value for difference in treatment control impacts across subgroups[b]				(0.062)*

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.

Due to rounding, reported impacts (T-C differences) and reported differences in impacts across subgroups may differ from differences between reported regression-adjusted means for the treatment and control groups and subgroups.
Standard errors adjusted for non-independence by clustering.
[a] The sum of the number of respondents in each subgroup exceeds the total number of respondents for given analysis. This is due to the fact that respondents may be interviewed up to four times across the observation period. A respondent's food retailer environment sometimes changed from one observation to the next due to changes in respondent location, or changes in retailer location or retailer HIP-participation. As a result, an individual respondent may be included in more than one subgroup.
[b] P-value represents significance level for joint test across all categories.
Source: Participant Survey (AM/PM dietary recall module), pooled Round 2 and Round 3 responses, including 10 percent second-day subsamples for each round (unweighted N=3,890 recalls from 1,996 respondents).

Exhibit B.13: Differences in Impacts of HIP on Consumption of Targeted Fruits, Cup-Equivalents, by Subgroup

	Regression-adjusted treatment mean (SE)	Regression-adjusted control mean (SE)	Treatment-control impact (P-value)	Difference in impacts (P-value)
Food retailer access: Two-category description(N=3,890 recalls from 1,996 respondents)				
Closest supermarket is HIP (N=665) [a]	0.550 (0.045)	0.533 (0.049)	0.017 (0.801)	
Closest supermarket is not HIP (N=1,531) [a]	0.573 (0.036)	0.440 (0.025)	0.133 (0.002)***	
Difference in levels: Closest HIP - Not Closest HIP (treatment and control combined)				0.093 (0.082)*
Difference in treatment control impact: Closest HIP, Not Closest HIP				-0.116 (0.140)
Food retailer access: Three-category description (N=3,890 recalls from 1,996 respondents)				
Within one mile of HIP participating supermarket (N=541) [a]	0.572 (0.055)	0.522 (0.055)	0.050 (0.518)	
Within one mile of non -HIP participating supermarket but not a HIP participating supermarket (N=938) [a]	0.567 (0.043)	0.443 (0.033)	0.124 (0.021)**	
Not within one mile of any supermarket (N=719) [a]	0.562 (0.056)	0.461 (0.040)	0.101 (0.144)	
P-value for difference in levels across subgroups(treatment and control combined)[b]				(0.411)
P-value for difference in treatment control impacts across subgroups[b]				(0.722)
Food retailer access: Five-category description (N=3,890 recalls from 1,996 respondents)				
Within one mile of high-volume HIP participating supermarket (N=135) [a]	0.485 (0.093)	0.566 (0.120)	-0.081 (0.601)	
Within one mile of high-volume non-participating supermarket but not a high-volume HIP participating supermarket (N=448) [a]	0.625 (0.066)	0.381 (0.047)	0.243 (0.003)***	
Within one mile of both high-volume HIP participating supermarket and non-participating supermarket (N=87) [a]	0.638 (0.119)	0.455 (0.106)	0.183 (0.247)	
Not within one mile of a high-volume supermarket but within one mile of non-high-volume supermarket (N=713) [a]	0.540 (0.045)	0.510 (0.041)	0.030 (0.612)	
Not within one mile of any supermarket (N=719) [a]	0.564 (0.056)	0.462 (0.040)	0.102 (0.143)	
P-value for difference in levels across subgroups(treatment and control combined)[b]				(0.171)
P-value for difference in treatment control impacts across subgroups[b]				(0.152)

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.

Due to rounding, reported impacts (T-C differences) and reported differences in impacts across subgroups may differ from differences between reported regression-adjusted means for the treatment and control groups and subgroups.
Standard errors adjusted for non-independence by clustering.
[a] The sum of the number of respondents in each subgroup exceeds the total number of respondents for given analysis. This is due to the fact that respondents may be interviewed d up to four times across the observation period. A respondent's food retailer environment sometimes changed from one observation to the next due to changes in respondent location, or changes in retailer location or retailer HIP-participation. As a result, an individual respondent may be included in more than one subgroup.
[b] P-value represents significance level for joint test across all categories.
Source: Participant Survey (AM/PM dietary recall module), pooled Round 2 and Round 3 responses, including 10 percent second-day subsamples for each round (unweighted N=3,890 recalls from 1,996 respondents).

Exhibit B.14: Differences in Impacts of HIP on Consumption of All Fruits and Vegetables, by Subgroup

	Regression-adjusted treatment mean (SE)	Regression-adjusted control mean (SE)	Treatment-control impact (P-value)	Difference in impacts (P-value)
Food retailer access: Two-category description (N=3,666 recalls from 1,909 respondents)				
Closest supermarket is HIP (N=638) [a]	2.669 (0.110)	2.336 (0.103)	0.333 (0.026)**	
Closest supermarket is not HIP (N=1,458) [a]	2.593 (0.072)	2.278 (0.063)	0.315 (0.001)***	
Difference in levels: Closest HIP - Not Closest HIP (treatment and control combined)				0.057 (0.628)
Difference in treatment control impact: Closest HIP, Not Closest HIP				0.019 (0.916)
Food retailer access: Three-category description (N=3,666 recalls from 1,909 respondents)				
Within one mile of HIP participating supermarket (N=524) [a]	2.647 (0.117)	2.432 (0.112)	0.215 (0.189)	
Within one mile of non -HIP participating supermarket but not a HIP participating supermarket (N=885) [a]	2.640 (0.098)	2.280 (0.085)	0.36 (0.005)***	
Not within one mile of any supermarket (N=668) [a]	2.557 (0.102)	2.234 (0.093)	0.324 (0.018)**	
P-value for difference in levels across subgroups(treatment and control combined)[b]				(0.371)
P-value for difference in treatment control impacts across subgroups[b]				(0.763)
Food retailer access: Five-category description (N=3,666 recalls from 1,909 respondents)				
Within one mile of high-volume HIP participating supermarket (N=135) [a]	2.382 (0.205)	2.495 (0.237)	-0.114 (0.718)	
Within one mile of high-volume non-participating supermarket but not a high-volume HIP participating supermarket (N=422) [a]	2.553 (0.134)	2.285 (0.114)	0.268 (0.123)	
Within one mile of both high-volume HIP participating supermarket and non-participating supermarket (N=686) [a]	2.908 (0.283)	2.569 (0.306)	0.339 (0.416)	
Not within one mile of a high-volume supermarket but within one mile of non-high-volume supermarket (N=81) [a]	2.704 (0.108)	2.308 (0.097)	0.396 (0.005)***	
Not within one mile of any supermarket (N=668) [a]	2.558 (0.102)	2.234 (0.093)	0.324 (0.018)**	
P-value for difference in levels across subgroups(treatment and control combined)[b]				(0.718)
P-value for difference in treatment control impacts across subgroups[b]				(0.695)

Two-sided test: *p<0.1, **p<0.05, ***p<0.01.

Due to rounding, reported impacts (T-C differences) and reported differences in impacts across subgroups may differ from differences between reported regression-adjusted means for the treatment and control groups and subgroups.
Standard errors adjusted for non-independence by clustering.
[a] The sum of the number of respondents in each subgroup exceeds the total number of respondents for given analysis. This is due to the fact that respondents may be interviewed up to four times across the observation period. A respondent's food retailer environment sometimes changed from one observation to the next due to changes in respondent location, or changes in retailer location or retailer HIP-participation. As a result, an individual respondent may be included in more than one subgroup.
[b] P-value represents significance level for joint test across all categories.
Source: Participant Survey (AM/PM dietary recall module), pooled Round 2 and Round 3 responses, including 10 percent second-day subsamples for each round (unweighted N=3,890 recalls from 1,996 respondents).