INSTITUTE for RESEARCH on POVERTY
UNIVERSITY OF WISCONSIN–MADISON

IRP Discussion Paper

No. 1399-12

Food Insecurity and Food Access

Alessandro Bonanno
Department of Agricultural Economics and Rural Sociology
The Pennsylvania State University
Email: abonanno@psu.edu

Jing Li
Department of Agricultural Economics and Rural Sociology
The Pennsylvania State University
Email: jing.li@psu.edu

Final Report prepared for the Institute for Research on Poverty at the University of Wisconsin–Madison – Research Innovation and Development Grants in Economics (RIDGE) program of the USDA. The authors thank Judith Bartfeld, Timothy Smeeding, Sean Cash, Sarah Lessem, and the other participants in the IRP RIDGE Center for National Food and Nutrition Assistance Research Small Grant Workshop (Madison, WI); Mark Nord; the participants in the 2011 AAEA & NAREA Joint annual meeting (Pittsburgh, PA); and those of the Research Innovation and Development Grants in Economics (RIDGE) Conference (Washington, DC) for precious suggestions and comments on earlier versions of this work. The Institute for Research on Poverty at the University of Wisconsin–Madison, and the Research Innovation and Development Grants in Economics Program of the USDA are thankfully acknowledged for financial support.

Abstract

This paper measures the relationship between food access and households' food adult insecurity using two years of Current Population Survey – Food Security Supplement data, matched with MSA-level data on different food outlets (small grocery stores and convenience stores, medium and large grocery stores, convenience stores associated with gas stations and Wal-Mart Supercenters). Endogeneity of food stores' location is tested and accounted for to eliminate spurious correlation between household food insecurity status and food access. The results indicate that while medium-large grocery stores and small food stores have a mitigating effect on adult food insecurity, especially among low-income households and households with children, convenience stores attached to gas stations seem to contribute to higher food insecurity levels among low-income households. The presence of Wal-Mart Supercenters seems to have no overall impact on adult food insecurity. Some results point to the company having a modest *direct* adult food insecurity-easing effect, and a detrimental *indirect* effect (via a negative impact on the number of other food stores helping reduce food insecurity), suggesting a null *net* effect of Wal-Mart Supercenters on adult food insecurity.

JEL Codes: Q18; L81; I14

Key Words: Food Security, Food Access, Endogeneity, Two-Stage Residual Inclusion.

Food Insecurity and Food Access

1. FOOD INSECURITY AND FOOD ACCESS INTRODUCTION

Food insecurity (henceforth FI) is the outcome of a household being unable to acquire (or being uncertain of having) enough food to meet the needs of all its members (Nord et al., 2010).[1] During the most recent economic downturn, estimates of households' FI levels in the United States have soared. Nord et al. (2011) report that, at some point during the year 2010 there were 17.2 million (14.7 percent) households affected by FI; of these, 10.9 million (9.1 percent) were characterized as Low Food Secure (LFS) households and 6.4 million (5.4 percent) as Very Low Food Secure (VLFS) households.[2] Although these figures are slightly smaller than those of the previous two years, they are considerably larger than past values: in the year 1999, only 10.1 percent of U.S. households were food insecure, of which 7.1 percent were LFS and 3.0 percent VLFS (Nord et al., 2010).

The phenomenon of household FI, resulting in social, psychological, and physical negative outcomes for both children and adults affected by it (Haering and Syed, 2009) has been intensively studied. For instance, the number of studies investigating the characteristics of food insecure households is plentiful (see, for example, Rose, Gundersen and Oliveira, 1998; Nord, Andrews, and Carlson, 2004; Nord et al., 2010). Even more numerous are the analyses assessing the effectiveness of Supplemental Nutrition Assistance Program (SNAP) participation in reducing households' FI, with mixed findings.[3]

---

[1]For other definitions of food security, food insecurity, and related terminology, see Haering and Syed (2009) and the literature cited therein.

[2]Low Food Secure households have obtained enough food to avoid substantial disruption in their eating patterns or reduced food intake by using a variety of coping strategies, such as eating less varied diets, participating in Federal food assistance programs, or getting emergency food from community food pantries. These households were previously described as "food insecure without hunger." Very Low Food Secure households are affected by disruption of normal eating patterns of one or more household members. Food intake was reduced at times during the year because they had insufficient money or other resources for food. These households were previously described as "food insecure with hunger."

[3]See, for example, Gundersen and Oliveira (2001); Jensen (2002); Borjas (2004); Kabbani and Kmeid (2005); Bartfeld and Dunifon (2006); Gibson-David and Foster (2006); Yen et al. (2008); Nord and Golla (2009); and Ratcliffe, Mc Klerman and Zhang (2011). As Jensen (2002) and Gibson-David and Foster (2006) have discussed, one of the main challenges in assessing the impact of SNAP participation on FI is that unobserved factors

Surprisingly, in spite of the copious evidence suggesting that areas inhabited by a prevalence of less-privileged individuals are characterized by limited access to large (or "high quality") food stores[4] and that limited access (or the existence of *food deserts*—areas with limited access to supermarkets and supercenters) can constitute a barrier to obtaining an adequate amount of nutritious food (Hearing and Syed, 2009; Ver Ploeg et al., 2009), there has been no empirical analysis assessing the consequences of lack of food access on outcomes such as food insecurity and hunger (Cummins and Macintyre, 2002).[5] This lack of rigorous analysis is also surprising given the existence of programs both at the national and at the local level aiming to improve food security through food access (Haering and Syed, 2009),[6] and the fact that improving access to nutritious food (both physical access and affordability) has been postulated as one of the possible methods to help reducing FI (Nelson, 2000).

The food environment can affect FI on different fronts: in the first place, limited access, or access to isolated food stores, could be characterized by higher food prices, either because of monopoly position (pricing power) or because of cost inefficiencies (King, Leibtag, and Behl, 2004). Also, different types of food outlets can affect food insecurity status through different mechanisms. On the one hand, the presence of easy-to-reach stores (proximity/convenience stores) can result in improved access and obviate the lack of means of transportation that may prevent low-income households from adopting cost-saving strategies

---

affecting both outcomes could bias the results. Recent studies addressing issues of simultaneity and food insecure households' self-selection into joining the program (Borjas, 2004; Bartfeld and Dunifon, 2006; Yen et al., 2008; Nord and Golla, 2009; Ratcliffe, McKernan, and Zhang, 2011) have provided evidence supporting SNAP participation having a mitigating effect on FI.

[4]See, for example, Alwitt and Donley (1997); Ball, Timperio, and Crawford (2008); Cotterill and Franklin (1995); Morland, Wing, and Diez Roux (2002); King, Leibtag, and Behl (2004); Moore and Diez Roux (2006); Powell et al. (2007); and Zenk et al. (2005). Furthermore, a positive relationship exists between the quality of the food choices that low-income (food stamps recipient) households make and the access to food outlets (Rose and Richards, 2004).

[5]A very limited number of studies have attempted to understand the relationship between household's FI status and the surrounding environment. One exception is the Bartfeld and Dunifon (2006) analysis of how state-level "food security infrastructure" contextual (economic and social) attributes affect the likelihood of food security among households with children.

[6]For example, following the passage of the Community Food Security Act in 1996, the USDA launched the Community Food Security Initiative in 1999 to help establish partnerships between USDA and local communities (Scott Kantor, 2001).

(Leibtag and Kaufman, 2003). On the other hand, large stores, which could be arguably harder to reach, can provide more variety and lower prices, alleviating the probability of experiencing FI as households may have more flexibility and assortment for their food choices.

Furthermore, more complex mechanisms may be in place. For example, the presence of nontraditional food retailers, such as Wal-Mart Supercenters,[7] could provide the beneficial effect of lower prices and larger assortments and also ignite a pro-competitive effect (Hausman and Leibtag, 2007; Volpe and Lavoie, 2008; Basker and Noel, 2009; Cleary and Lopez, 2011), which could create a spillover effect resulting in lower prices.[8] However, as the company has been found to impact negatively the performance of other businesses (Singh, Hansen, and Blattberg, 2006; Ailawadi et al., 2010) and their survival (Haltiwanger, Jarmin, and Krizan, 2010), there could be indirect spillovers of the company's impact on FI, which could be hard to measure.

The goal of this analysis is to understand whether access to food retailers of different types can have an effect on households' FI. Using household-level data from the Current Population Survey – Food Security Supplement (CPS-FSS) for the years 2004 and 2005, matched with MSA-level food stores data, we assess the impact of Wal-Mart Supercenters, medium-large grocery stores (proxy for traditional, full-service food stores), small food stores (convenience stores and small groceries), and convenience stores attached to gas stations on adult FI, focusing on different subsamples of households, segmented depending upon presence (or absence) of children and income (185 percent below the poverty level). The food access measures are obtained dividing the MSA-level number of stores by population. Due to the potential presence of confounding factors that could affect both the food access measures and the

---

[7]The company has gradually moved away from its discount stores format (carrying a limited number of food products, mostly shelf-stable) to the Supercenter format, which offers fresh produce, meat, bakery, deli, and fresh seafood departments, becoming the largest food retailer in the United States (Food Marketing Institute, 2007). As of January 31 2011, Wal-Mart operated (in the U.S. alone) 2,747 Supercenters and 803 Discount Stores (Wal-Mart Stores Inc, 2010).

[8]Furthermore, as Wal-Mart locates its stores preferentially in areas where competition is scant (Jia, 2008; Bonanno, 2010), and there appears to be a positive relationship between the company's location and the rate of food stamps recipients (Bonanno, 2010), its expansion could improve food access for low-income households who may have limited access otherwise.

likelihood of a household experiencing FI, we control for the endogeneity of food stores' location decision using a two-step IV probit estimator (Newey, 1987; Rivers and Vuong, 1988) and a different identification strategy for each food-store type. In the specific case of Wal-Mart Supercenters, we use an identification strategy exploiting the unique expansion strategy of the company combining lagged number of per-capita Discount Stores and the distance from the company's headquarters (Bentonville, Arkansas), while for other food outlets we capture variation in number of stores per capita due to geographic differences in the supply-side determinants of location decision and aggregate market potential. The results show that food access does have an impact on household FI, although the effects differ across samples and store types. We find that proximity stores and medium-large grocery stores have a considerable mitigating effect on adult FI, in particular for low-income households, and that convenience stores attached to a gas station have instead a positive effect. We also find little support of Wal-Mart Supercenters affecting adult food insecurity: this is the result of a direct mitigating effect being offset by an indirect aggravating effect due to Wal-Mart Supercenters' negative impact on the density of those outlets that have an FI-reducing effect.

2. AN EMPIRICAL MODEL OF FOOD INSECURITY AND FOOD ACCESS

The following model is a stylized representation of FI as the outcome of a household's optimization problem. Household *i* located in area *l* maximizes its utility, function of income (spent on goods) and leisure (or hours worked), subject to time and budget constraints. Although the formal derivation of the model is not illustrated here, the interested reader can refer to Jensen (2002) for a thorough discussion. In Jensen's model (which does not account for the role of the built environment, but which considers participation in the Food Stamp Program), FI enters the utility function since it causes disutility due to concerns about having an adequate food supply, and under-consumption of food for some household members. In the context of this analysis, the household FI status will depend upon both the characteristics of the household and the features of the surrounding environment. Thus, the FI status of household *i* in area *l*, or FI_{il} will be represented by the following function:

$$FI_{il} = f(\mathbf{X_{il}}, \mathbf{FA_l}, \mathbf{d_l} \,|\, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\gamma}) + e_{il} = f(\mathbf{Z} \,|\, \boldsymbol{\theta}) + e_{il}, \tag{1}$$

where $\mathbf{X_{il}}$ is a vector of household characteristics, $\mathbf{FA_l}$ is a proxy capturing the level of access to food for all households in area l (measured by the number of outlets of a given store type N_j divided by the total population in area l or N_j / pop_l), $\mathbf{d_l}$ is a vector of fixed effects to control for unobservable factors that could impact FI, $\boldsymbol{\beta}$, $\boldsymbol{\delta}$, and $\boldsymbol{\gamma}$ are vectors of parameters conformable to $\mathbf{X_{il}}$, $\mathbf{FA_l}$, and $\mathbf{d_l}$, respectively, and e_{il} is an error term. The first part of the central term in equation (1) can be summarized as $f(\mathbf{Z} \,|\, \boldsymbol{\theta})$ where $\mathbf{Z}$ is the vector of all the variables that can influence FI and $\boldsymbol{\theta}$ **is** a conformable vector of parameters characterizing the relationship between the covariates in $\mathbf{Z}$ and FI_{il}.

Let h be a realization of FI_{ij}, that is, a FI state. Consider the simplest case, that is, that of only two food security states for the adult members of a household, Food Secure (FI=0) and Food Insecure (FI=1), so that $h = \{0, 1\}$. In this case, the probability of observing a given realization of h is:

$$\Pr(FI_{il} = 1 | \boldsymbol{Z}) = \Xi\,(\,\mathbf{Z}'\boldsymbol{\theta}\,); \tag{2}$$

where $\Xi(.)$ can represent either the standard normal or the logistic cumulative density function (CDF) , or in other words, $\Xi(.)=\Phi\,(.)$ or $\Xi(.)= \Lambda(.)$, which, would allow the vector of coefficients $\boldsymbol{\theta}$ to be estimated via either a probit or a logit maximum likelihood estimator. In the remainder of the illustration we are going to assume $\Xi(.)=\Phi(.)$.[9]

A probit estimator in this case would lead to unbiased and consistent estimates of the vector $\boldsymbol{\theta}$ if none of the independent variables were spuriously correlated with FI. However, endogeneity bias may be present since food retailers' location decision is non-random and could be affected by (unobserved) factors also contributing to a household's FI state. Also, the source of endogeneity bias may be different

---

[9]Alternatively one could have $h = \{0,1,2\}$ where the three states are Food Secure (FI=0) Low Food Insecure (FI=1) and Very Low Food Insecure (FI=2). In this case, equation (2) would be rewritten as

$$\Pr(FI_{il} = h | \boldsymbol{Z}) = \Xi\,(\delta_{h-1} - \mathbf{Z}'\boldsymbol{\theta}\,) - \Xi\,(\delta_h - \mathbf{Z}'\boldsymbol{\theta}); \tag{2a}$$

where $\delta_0 = -\infty$, $\delta_3 = +\infty$ and the vector of coefficients $\boldsymbol{\theta}$ to be estimated via either an ordered probit or an ordered logit estimator. In section 4.3, we propose an alternative estimation of equation (2) via different types of ordered logits.

for different types of food outlets. For example, empirical evidence (Ellickson, 2006; 2007) supports that for low-quality, smaller stores, the equilibrium number of entrant firms increases with market size, while higher-quality stores, which invest in fixed costs, tend to create a natural oligopoly where the equilibrium number of entrants is not destined to grow indefinitely with market size (Shaked and Sutton, 1987; Sutton, 1991).[10] Thus, according to this notion, the equilibrium number of food retailing firms providing a "higher-quality" product is likely to be impacted by other factors besides market size, such as consumers' characteristics, demand for services (Bonanno and Lopez, 2009), and similar factors.

Food retailing firms (both incumbents and potential entrants) played a dynamic entry game (as they maximize expected profits), with infinite possible equilibria whose detailed analysis is a daunting task (see Jia [2008] for an example). Using instead the simplifying assumptions that (1) food retail companies have limited ability to choose their store formats[11] (i.e., that the product-type offered by each retailer is given); and (2) that, in a given area and for a given establishment type, food retail firms (facing symmetric demand and cost) can be ordered by decreasing profitability (i.e., the most profitable firms enter the market first, as in Berry, 1992),[12] there exists a Nash equilibrium (although not unique), which allows the researcher to treat the observed number of market participants as one of the possible equilibria of a game played by all potential entrants. As such, the observed number of food retail outlets of type *j* in area *l*, defined as N^*_{jl}, will be one possible equilibrium outcome of the location game discussed above, and it will be determined by market-level variables.

---

[10]Such considerations apply to most industries whose firms commit to a specific location. Asplund and Sandin (1999) point out in their analyses of Swedish regional markets for driving schools, as profits per-capita decrease in market size, capacity will tend to impose a limit to the possibility of observing a higher number of equilibrium firms.

[11]Strictly speaking, retail firms present different formats (making them differentiated products). Including the format-type decision in the game, will complicate the analysis further, as illustrated by Mazzeo (2002) and Seim (2006).

[12]This assumption is consistent with those of seminal models of firms' entry (e.g., Bresnahan and Reiss, 1991; Berry, 1992)

Define with S_l the total (aggregate, potential) market size in area l, and let's assume that S_l is a proportion of the population pop_l or that $V_l * pop_l = S_l$.[13] Let's consider the following reduced form equation representing the equilibrium number of stores of the j–the type in area l divided by the area's population:

$$\frac{N_{jl}^{*}}{Pop_l} = \alpha_j^V V_l + g^D(\mathbf{X}_{jl}, \mathbf{X}_{-jl}; \boldsymbol{\alpha}_\mathbf{j}^\mathbf{D}) + g^C(\mathbf{C}_\mathbf{jl}, \mathbf{K}_\mathbf{jl}; \boldsymbol{\alpha}_\mathbf{j}^\mathbf{C}) + \varepsilon_{jl} \tag{3}$$

where the $g^D(.)$ and $g^C(.)$ are functions representing, respectively, the role of demand factors (other than market potential) and cost factors on the equilibrium store density levels; $\boldsymbol{X}_{jl}$ and $\boldsymbol{X}_{-jl}$ are vectors of demand characteristics in area l (both for the j-th store type and for that of other store-types)[14]; $\mathbf{C}_{jl}$ and $\mathbf{K}_{jl}$ are vectors of format- and market-specific cost variables (variable and fixed cost, respectively), the $\boldsymbol{\alpha}_\mathbf{j}^\mathbf{D}$ and $\boldsymbol{\alpha}_\mathbf{j}^\mathbf{C}$ are conformable vectors of parameters, and ε_{jl} is an idiosyncratic error term. It should be noted that the measure of food access, in equation (2) is the observed number of stores per capita, or that

$$FA_{jl} \equiv \frac{N_{jl}^{*}}{Pop_l}.$$

Under the assumption of $E(\varepsilon_{jl} e_{il}) = 0, \forall i, j, l$, if market size and the variables in $g^D(.)$ and $g^C(.)$ are also uncorrelated with e_{il}, then FA_{jl} would not be spuriously correlated with FI_{il} and estimates of the parameters of equation (2) will be unbiased. If instead some of the variables in (3) are potentially correlated with an unobserved driver of FI, one needs an empirical remedy for the resulting estimated parameters' bias.

---

[13] As market size, representing the potential demand for the goods offered by the firm j, is a function of market characteristics, such proportion is not a constant but depends on other factors such as income and consumers' heterogeneity, which, in equation (3) are represented by the vectors $\boldsymbol{X}_j$ and $\boldsymbol{X}_{-j}$. See Asplund and Sandin (1999) for more details.

[14] Demand characteristics across store types enter equation (3) because in the second stage of the game firms are likely to compete with those of other formats.

Assuming linearity of $g^D(.)$ and $g^C(.)$ in both variables and parameters, one can separate the variables in (3) in the two vectors $\mathbf{W_{jl}}$ and $\mathbf{R_{jl}}$, the former including factors uncorrelated with unobserved drivers of households' FI, the latter including instead those potentially correlated with it. Indicating the respective vectors of parameters for these variables as $\boldsymbol{\alpha}_{\mathbf{j}}^{W}$ and $\boldsymbol{\alpha}_{\mathbf{j}}^{R}$ one can rewrite equation (3) as

$$\frac{N_{jl}^{*}}{Pop_l} = \mathbf{W_{jl}}'\boldsymbol{\alpha}_{\mathbf{j}}^{W} + \mathbf{R_{jl}}'\boldsymbol{\alpha}_{\mathbf{j}}^{R} + \varepsilon_{jl} = \mathbf{W_{jl}}'\boldsymbol{\alpha}_{\mathbf{j}}^{W} + r_{jl} \tag{4}$$

where the term $r_{jl} = \mathbf{R_{jl}}'\boldsymbol{\alpha}_{\mathbf{j}}^{R} + \varepsilon_{jl}$ is by construction, correlated with the errors in equation (1). It is easy to show that, if one introduced r_{jl} (or an unbiased estimate) in equation (1) the source of correlation between FA_{jl} and e_{il} will be accounted for in the model and the new resulting error term uncorrelated with $FA_{jl.}$ Let $\hat{r}_{jl}$ be the errors obtained from a first-stage linear regression of the *j*-th FA indicator on the (weakly) exogenous vector $\mathbf{W_{jl}}$:

$$\frac{N_{jl}^{*}}{Pop_l} = \mathbf{W_{jl}}'\hat{\boldsymbol{\alpha}}_{\mathbf{j}}^{W} + \hat{r}_{jl} \tag{5}$$

The model in equation (2) can then be rewritten as follows:

$$\Pr(FI_{il} = 1 \mid \mathbf{Z}, \hat{r}) = \Phi([\mathbf{Z}, \hat{r}]'\boldsymbol{\lambda}^{\mathbf{2SRI}}) \tag{6}$$

whose parameters, under the assumptions in (2), can be estimated using a probit estimator. It should be noted that the vector of coefficients in (6) is $\boldsymbol{\lambda}^{\mathbf{2SRI}} = [\boldsymbol{\theta}^{2SRI}, \lambda]$ where λ is the coefficient associated with $\hat{r}_{jl}$ and $\boldsymbol{\theta}^{\mathbf{2SRI}}$ differs from $\boldsymbol{\theta}$ in equation (2). Following the notation in Wooldridge (2002, pp. 473 and 474) the relationship linking the parameters of equation (6) and (2) can be expressed (using a generic coefficient δ as example) as $\delta^{2SRI} = \delta / (1-\rho^2)^{1/2}$, where $\rho = \mathrm{Corr}(r, e)$. It follows that, in absence of endogeneity, ρ =0 and $\delta^{2SRI} = \delta$.

The approach illustrated above, which resembled Rivers and Vuong (1988) two-step estimator, is also referred to as the 2-Stage Residual Inclusion (2SRI) method, and it is superior to classical two-stage instrumental variable methods in non-linear models (Wooldridge, 2002). Terza, Basu, and Rathouz (2008) show that, while classical "2-stage" approaches can produce inconsistent estimates, the 2SRI method produces unbiased and consistent estimates for a broad family of non-linear estimators. Applications of similar methods can be found in several areas, such as policy analysis (Alvarez and Glasgow, 1999), health economics (Terza, Basu, and Rathouz, 2008) and marketing (Petrin and Train, 2010).

3. DATA AND ESTIMATION

3.1 Data Sources and Variables Definition

The data used come from different sources. Data on households' FI status and their characteristics come from two years of individual-level observations of the Current Population Survey Food Security Supplement (CPS-FSS) of the U.S. Census Bureau and Bureau of Labor Statistics, December 2004 and 2005.[15]

The CPS-FSS reports different measures of household FI. Survey respondents are asked a series of eighteen questions related to the availability of food in their households, including limitations in food consumption and the number of meals skipped, distinguishing for disruptions in eating habits for adults (ten questions) and children (eight questions). Raw food insecurity scores are constructed using the responses to these questions, and then coded to obtain discrete FI indicators for adults, children, and for the household as a whole. In this analysis we focus on adult FI only[16] and use the "12-month adult food security summary status" to create an indicator for the level of household food insecurity among adult

[15]The choice of the years 2004 and 2005 was made on two grounds. First, the CPS-FSS data had to be matched with MSA-level data from another source; although MSA definitions change across years *and* databases, those for the years 2004 and 2005 allowed for relatively easy matching of the different databases. Second, as data on Wal-Mart Supercenters' location is only available until January 2006, through T. J. Homes Store location database (Holmes, 2010), no subsequent years were used.

[16]We thank Mark Nord at the ERS/USDA for suggesting this approach.

household members.[17] A binary adult food insecurity indicator (*FI*) is calculated combining the High and Marginal Food Security (*FI*=0) and Low Food Security and Very Low Food Security (*FI*=1) status.

While the public access files of the CPS-FSS do not include the exact location of the individuals' surveyed, most households have attached state and MSA-code identifiers, allowing the individual CPS-FSS data to be matched with other, MSA-level and state-level databases.

Data on traditional food retailers' location were obtained from the County Business Pattern (CBP) database of the U.S. Census Bureau/Bureau of Labor Statistics (BLS). The industries considered are NAICS 445110: Grocery Stores; NAICS 445120, Convenience Stores, and NAICS 447110, Convenience Stores with Gas Station. Data on Wal-Mart Supercenters' store number and location are obtained from T. J. Holmes database (Holmes, 2010). County-level CBP data and the Wal-Mart data were aggregated to the MSA level to match the geographic indicators of the CPS-FSS.[18]

We re-categorized some of the food store-types based on their size, using the number of establishments for each employee-class contained in the CBP: grocery stores' establishments with less than 50 employees are combined with the number of establishments belonging to NAICS 445120 to obtain a proxy for the number of small (proximity / low assortment) food stores. The variable SMALL is then obtained dividing this number by total population (in tens of thousands), obtained from the U.S. Bureau of Census Population Estimates Program (PEP) database. A proxy for access to supermarkets and other traditional food outlets, GROC, is obtained dividing the MSA-level number of NAICS 445120 establishment with 50 or more employees by population in hundreds of thousands; the variable GSCNV, a proxy for access to outlets characterized by limited accessibility (as for cars are usually necessary) and

---

[17]Food security summary status indicators for both adults and children are only in survey years from 2005 onward; for the year 2004, adult FI score was calculated subtracting the child raw score form the household raw scores (for households with children). The adult FI summary indicator was then coded accordingly to the number of positive responses to the FI questions for adults (0=High Food Security, 1–2=Marginal Food Security, 3–5=Low Food Security, 6–10=Very Low Food Security). See Nord (2002) for more details.

[18]CBP data at the MSA–level could not be directly used due to discrepancies in some of the classifications across the two databases.

assortment, is obtained dividing the number of NAICS 447110 establishments by population in hundreds of thousands. Lastly the variable WMSC is obtained dividing the aggregated, MSA-level number of Wal-Mart Supercenters by population in millions.

Household-level variables from the CPS-FSS survey are used to control for households' characteristics: age of the household head (AGE); number of children in the household (CHILD); highest education level in household (three binary variables indicating, respectively, high-school, HIGHSC, some college, SOMCOL, and bachelor degree or more, COLMOR); household head being male (MALE); and a series of binary variables accounting for race of the household head (Black, Asian, and Hispanic, respectively), home ownership (HMOWN), single-head household (SINGLEH), unitary household (SINGLUN), and for the presence of any non-citizens (NOCITIZ), unemployed (UNEMPL), and disabled (DISABL) individuals in the household.

A proxy for household income was obtained following Jensen (2002), that is, assigning to each household an income level equal to the mid-point of the household income bracket the household belongs to. Per-capita household income (INC_PR) was then obtained dividing this measure by household size. Lastly, the general Consumer Price Index (from the BLS), a dummy for the year 2004 and state-level fixed effects obtained using the state identifiers in the CPS-FSS are included in the model to control, respectively, for different price levels across areas and time, and for unobserved heterogeneity across geographic location of residence.

Only households presenting valid entries of adult FI status, geographic indicators, as well as valid entries of the household characteristics and of the 16-level household income brackets illustrated above are retained in the database. The total number of data points used in the estimation consists of 36,887 observations. From this database, referred to as the All HH (all households) sample, a subsample including only households whose income is below the 185 percent of the current poverty threshold, referred to as the low-income (All HH Low-Income) subsample is obtained (n = 7,487). As the existing evidence points towards households with children being more likely to be affected by food insecurity (see, e.g., Nord et al., 2010), the database was further divided into two Households with children samples

(Full sample, $N = 14{,}691$ and Low-Income Sample, $n = 4{,}081$), and households without children (Full Sample, $N = 22{,}196$ and Low-Income Sample, $n = 3{,}406$).

Summary statistics for the FI indicator and FA variables across the different subsamples are reported in Table 1. The values show that, as expected, the percentages of FI households are much larger in the low-income sample than in the full sample. In particular, 23.7 percent of the households in the All HH Low-Income sample experienced adult FI in the 12-month period prior to the survey versus 8.4 percent in the full sample. Incidence of adult FI is also larger among HH with children, than in those without children, this difference being more marked in the full samples (10.2 percent in HH with children vs. 7.3 percent in HH without children) than in the low-income samples (24.9 percent and 22.2 percent, respectively).

The average values of the FA variables shows that Wal-Mart tends to locate preferentially in areas with a higher concentration of low-income individuals: the average number of WMSC in the low-income subsamples is 15 percent higher in the low-income samples than in the full samples (4.61 percent vs. 5.32 percent in the All HH sample). The sample averages for the other food stores' density are instead relatively similar across samples: the average number of medium-large grocery stores per 100,000 people is approximately 6.3, that of small food stores 2.4 (2.3 in the low-income samples), while that of convenience stores attached to gas stations is approximately 2.7 (2.8 in the low-income samples). Lastly, a list of all the household-level variables and summary statistics for the two All HH samples are reported in the top half of Table 2.

3.2 Identification Strategy

This section illustrates the different identification strategy used to correct for the potential endogeneity of each of the FA variables. Although the variables will be referred to as "instruments," the reader should be aware that the estimation method adopted here differs from standard, two-stage instrumental variable methods.

Table 1
FI Indicators and Food Access Variables: Descriptive Statistics across Samples

	All Households				Households with Children				Households w/o Children			
	Full		Low-Income		Full		Low-Income		Full		Low-Income	
	% FS	%FI	% FS	%FI	% FS	%FI	% FS	%FI	% FS	%FI	% FS	%FI
FI	91.6	8.4	76.4	23.7	92.8	7.2	77.8	22.3	89.8	10.2	75.2	24.8
	Mean	St.Err	Mean	St.Err	Mean	St.Err	Mean	St.Err	Mean	St.Err	Mean	St.Err
WMSC	4.61	4.90	5.32	5.09	4.63	4.90	5.49	5.11	4.58	4.90	5.17	5.07
GROC	6.34	1.31	6.31	1.38	6.36	1.31	6.32	1.40	6.32	1.30	6.31	1.37
SMALL	2.41	0.99	2.31	0.96	2.42	0.98	2.31	0.93	2.41	1.00	2.31	0.98
GSCNV	2.72	1.19	2.86	1.22	2.73	1.19	2.94	1.21	2.70	1.19	2.79	1.22
N	36,887		7,487		14,691		4,081		22,196		3,406	

Legend and data sources:

WMSC: Number of WM Supercenters/ 1,000,000 people. Source: Holmes(2010) Database / PEP
GROC: Number of NAICS 445110 stores >=50 employees /100,000 people. Source: CBP / PEP
SMALL: Number of NAICS 445120+NAICS 445110 stores <50 employees /10,000 people. Source: CBP / PEP
GSCNV: Number of NAICS 447110 Stores /100,000 people. Source: CBP / PEP
FI: Adult Food Insecurity Indicator
(FS = Food Secure, Marginally Food Secure; FI = 0)
(FI = Low Food Security, Very Low Food Security; FI =1)

Table 2
Sample Statistics

Variable	Description	All HH Full Sample Mean	All HH Full Sample St.dev.	All HH Low-Income Sample Mean	All HH Low-Income Sample St.dev.
Demographic Variables – Continuous					
AGE	Age of household head	43.07	12.38	38.95	13.16
INC_PR	Estimated average income of each household member	29.64	23.21	17.52	4.05
CHILD	Number of children in household<18 year	0.73	1.06	1.17	1.33
		Frequency of 1			
Control Variables - Discrete					
MALE	Head is male		57.13		47.87
HIGHSC	Max educational attainment: high school		24.82		36.52
SOMCOL	Max educational attainment: some college		29.89		29.40
COLMOR	Max educational attainment: college degree or higher		37.82		12.94
ASIAN	Race is Asian household head		3.56		3.35
BLACK	Race is Black household head		10.64		18.33
HISP	Hispanic ethnicity household head		10.73		25.99
HMOWN	Own living quarters (for household)		70.14		41.85
SINGLEH	Single head household		16.38		32.32
SINGLUN	Single unit household		29.28		28.46
NOCITIZ	Non-citizen in household		12.29		24.44
UNEMPL	Unemployed in household		5.87		10.20
DISABL	Disabled in the household		0.49		0.80
CPI	Consumer price index	187.41	10.06	187.87	8.76
FA Instruments					
WMSC					
DIST_BC	Distance from Benton County	870.60	397.42	844.14	402.24
NDS_LAG	Density of WM DSs / 1,000,000 people (3-yr lag)	4.05	2.66	3.96	2.79
Common to GROC, SMALL, GSCNV					
PC_INC	Per capita income ($ thousand)	36.25	6.29	35.01	6.03
AVHPI	Average housing price index	208.97	34.98	206.64	35.33
P_ELECT	Electricity price for commercial use ($/Kwh) (cents/KWh)	8.51	2.20	8.40	2.16
GROC					
SHARE_LAND	Land share of housing value	0.35	0.23	0.34	0.24
P_DIESEL	Area-level diesel price (On-Highway) All Types ($/gal)	2.15	0.31	2.14	0.31
SMALL					
POPDEN	Population density (.000/square mile)	0.80	0.79	0.72	0.79
GSCNV					
CARS_PC	State-level per-capita privately owned vehicles	5.26	10.64	6.24	12.23
P_GAS	State-level refiner gasoline price ($/gal)	1.50	0.20	1.49	0.21

Note: All the demographic variables come from the CPS-FSS.

Our strategy to account for the endogeneity of WMSCs[19] uses two facts that are based on the company's unique store location strategy. First, as the company's expansion into food retailing capitalizes on converting its mass merchandize Discount Stores (DSs) into Supercenters (see Bonanno, 2010), a 3-year lagged number of DSs per 1,000,000 people is used as instrument for WMSCs as it represents a good predictor of SCs density (for a more thorough discussion of the rationale behind this approach, see the appendix of Basker and Noel, 2009). Furthermore, as illustrated in other analyses (Neumark, Zhang, and Ciccarella, 2008; Courtemanche and Carden, 2011), the distance from the company's headquarters in Bentonville, Arkansas (or a function of it) can be used as a predictor of the company's location decision. Until the mid-1990s Wal-Mart tended to open its DSs at distances progressively increasing from Benton County, Arkansas, as it capitalized on locating stores at driving distance from distribution centers, consistently with its "hub-and-spoke" logistic system (Walton and Huey, 1992). This phenomenon, which appeared to be less relevant for DSs in more recent years, is instead still valid for the company's SCs (see Courtemanche and Carden, 2011, for a more detailed discussion). Thus, distance from Benton County, calculated from longitude and latitude coordinates retrieved from the U.S. Census Gazetteer of Counties (2001) is also used to capture exogenous variation in per-capita number of Supercenters across MSAs.[20]

The identification strategy used for food access measures based on "traditional" food stores density is based on isolating aggregated, market-level determinants of store location, which are unlikely to be correlated with unobserved, household-level determinants of FI.

In the first place, some common factors affecting location decision for all types of food stores (i.e., affecting their expected profitability of locating in a given area) are the potential aggregate market demand, fixed investment costs, and operating costs (Ellickson, 2006; 2007). One commonly used

---

[19]Specifically, Wal-Mart Supercenter locations may be correlated with particular socio-demographic profile, which may in turn be correlated with poorer diets (e.g., high poverty rates, as in Goetz and Swaminathan, 2006; or share of population food stamps recipients, as in Bonanno, 2010).

[20]Other identification strategies were attempted, with mixed results. For a thorough discussion of the sensitivity of the results to the use of different instruments, see the discussion in section 4.3.

measure for the potential total size of the market is aggregate (disposable) income; as the food access variables are ratios of the number of establishments over population, we use a proxy of the annual aggregated MSA-level total income from county-level observations from the Bureau of Economic Analysis divided by population, to obtain a proxy for per capita MSA-level income. The proxy for fixed capital investment cost used is the Monthly House Price Indexes for Census Divisions from the U.S. Federal Housing Finance Agency. Monthly Retail prices of electricity per commercial use ($/Kwh) from the U.S. Department of Energy is instead used as common source of variable costs.21

Store-type specific variables are used for each FA measure: state-level Land Share (ration between the value of land over the total value of a home), from the "Land Prices by State" database of the Lincoln Institute of Land Policy as described in Morris and Heathcote (2007) is used for GROC, to capture the additional component of fixed costs these establishments need to face to deliver the "higher" levels of quality they provide to consumers (Ellickson, 2006; 2007). Also, as larger stores necessitate more frequent delivery of goods and they may operate their own truck fleet, the "On-highway" price of diesel (all types) in $/gal (from the U.S. Department of Energy) is used as additional instrument for GROC. Population density, in thousands of individuals divided by squared miles of land (from the U.S. Bureau of Census Gazetteer of counties [2001]), is used for SMALL, as smaller stores, which incur lower fixed-cost (mostly due to the smaller size and square footage), tend to locate in more densely populated areas. Lastly, two specific instruments used for GSCNV are the state-specific wholesale (refiner) gasoline price ($/gal), from the U.S. Department of Energy, and the state-level per-capita number of privately owned automobiles, from the U.S. Department of Transportation, Federal Highway Administration, to account for the specific existing demand for this type of outlet. A list of all the instruments used and some summary statistics are illustrated in the bottom half of Table 2.

[21] Although labor is another major source of cost in retailing, we opted for not including proxies for retail wages for two reasons. First, wages are decided by the retail firms themselves and may be affected by the composition of the local retail industry, which makes them less likely to be exogenous. Second, a practical matter is that for some of the sub-industries considered in our definition of food access variables, we encountered missing observations due to the non-disclosure policy adopted by the BLS.

3.3 Estimation

Equation (2) was estimated using a simple maximum likelihood probit estimator. The estimation of equation (6) was performed via the two-step IV-probit procedure in STATA, which uses the approach proposed by Newey (1987) and Rivers and Voung (1988). We tested for the validity of the identification assumptions in the two-step IV-Probit by means of Amemiya-Lee-Newey (ALN) minimum χ^2 statistic (Amemiya, 1978; Newey, 1987; Lee, 1992).[22] A Wald test of endogeneity is performed on the significance of the correlation coefficient between the errors of the first stage equation and those of equation (3) under the null of exogeneity (i.e., rejection of the null hypothesis indicates endogeneity).[23]

One complication in using a two-stage IV-probit estimator is that, since its parameters are identified "up to scale," the explanatory variables' marginal effects cannot be directly calculated and, most importantly, obtaining the standard errors associated with such marginal effects would be a rather complex task (see Wooldridge, 2002, pp. 475–476). We circumvent these issue following the approach suggested by Wooldridge (2002, p. 476), consisting in estimating directly the parameters of equation (6) (i.e., obtaining estimates of the vector $\boldsymbol{\lambda}^{\mathbf{2SRI}}$) via a two-step procedure, then calculating the marginal effects using the estimated $\hat{\boldsymbol{\lambda}}^{\mathbf{2SRI}}$, thus averaging across the first stage errors in the sample, and using bootstrapped standard errors to account for the presence of the first-stage residual in the model. To differentiate these estimates from those of the two-step IV-probit, we will refer to them as 2SRI-probit estimates. Estimating directly the parameters of the 2SRI-probit has another advantage, that is it provides an additional test of whether the FA variables are endogenous: using the same logic behind Hausman's (1978) endogeneity test, a non-statistically significant parameter associated with the residual of the first stage regression will provide evidence against the parameters of equation (2) being affected by endogeneity bias.

---

[22]This test is performed using the STATA routine **overid**. See Baum et al. (2006) for more details regarding the computation of the test.

[23]The test is not performed on $\hat{\rho}$ but on $\operatorname{atanh}\hat{\rho} = 0.5\ln[(1+\hat{\rho})/(1-\hat{\rho})]$.

Lastly, we evaluate the power of the instruments via *F*-tests on the joint-significance of their coefficients in the first stage regression, using Staiger and Stock (1997) rule of thumb of *F*-statistic larger than 10 being as enough evidence to dismiss issues of weak instruments. All data manipulation and estimation were performed in STATA v. 11.

4. EMPIRICAL RESULTS

As the number of MSA-level observations for every year is limited (approximately 270 observations), and most of the instruments data are at the state-level, the limited variation resulted in causing most attempts to include the FA variables simultaneously in the FI equation to produce insignificant results. The estimated coefficients and marginal effects reported below in sections 4.1 and 4.2, respectively, as well as the additional estimates and robustness checks discussed in section 4.3, come from models the FA variables were used separately one at the time. The results of an indirect estimation procedure attempted at isolating the effects of the different food store types on FI as well as some of the indirect effect associated with the presence of Wal-Mart Supercenters is illustrated in section 4.4.

4.1 FA and FI: Estimated Coefficients and Model Performance

The estimated parameters of equation (2) obtained using probit, and those of equation (6) obtained via two-step IV-probit and 2SRI probit, for the All HH samples are reported in Table 3. In general, the use of different FA variables does not affect the overall performance of the model (the pseudo R-squared remain substantially unchanged showing approximately the same magnitude of 0.18, and the value of the likelihood ratio tests for the joint significance of the coefficients shows similar values across models). The sign, magnitude, and behavior of the estimated FA parameters illustrated in Table 3 exemplify patterns common across subsamples, highlighting that not all the FA variables appear to be

Table 3a
Estimated Coefficients equation 2: WMSC and GROC Full sample (N=36,887)

	WMSC			GROC		
Variables	Probit	IV-Probit	2SRI-Probit	Probit	IV-Probit	2SRI-Probit
FA	0.0024	-0.0150	-0.0163	-0.0022	-0.1331**	-0.1305**
	0.0035)	(0.0173)	(0.0201)	(0.0124)	(0.0628)	(0.0623)
AGE	-0.0017*	-0.0017*	-0.0017*	-0.0017*	-0.0017*	-0.0017**
	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0008)
MALE	-0.1599***	-0.1594***	-0.1594***	-0.1600***	-0.1614***	-0.1614***
	(0.0227)	(0.0227)	(0.0188)	(0.0227)	(0.0227)	(0.0184)
HIGHSC	-0.0840**	-0.0847**	-0.0848**	-0.0840**	-0.0856**	-0.0856***
	(0.0370)	(0.0370)	(0.0353)	(0.0370)	(0.0371)	(0.0332)
SOMCOL	-0.0291	-0.0316	-0.0319	-0.0294	-0.0330	-0.0331
	(0.0374)	(0.0375)	(0.0383)	(0.0374)	(0.0375)	(0.0387)
COLMOR	-0.3467***	-0.3530***	-0.3537***	-0.3475***	-0.3463***	-0.3466***
	0.0417)	(0.0422)	(0.0367)	(0.0417)	(0.0418)	(0.0382)
HISP	(0.0413	0.0282	0.0267	0.0391	0.0192	0.0189
	(0.0362)	(0.0384)	(0.0411)	(0.0362)	(0.0375)	(0.0339)
ASIAN	-0.1414**	-0.1475**	-0.1481*	-0.1425**	-0.1451**	-0.1453*
	(0.0690)	(0.0693)	(0.0769)	(0.0690)	(0.0691)	(0.0751)
BLACK	0.1863***	0.1774***	0.1766***	0.1848***	0.1808***	0.1807***
	(0.0314)	(0.0326)	(0.0367)	(0.0313)	(0.0315)	(0.0333)
HMOWN	-0.3352***	-0.3330***	-0.3327***	-0.3348***	-0.3339***	-0.3338***
	0.0249)	(0.0250)	(0.0290)	(0.0249)	(0.0249)	(0.0268)
INC_PR	-0.0249***	-0.0251***	-0.0251***	-0.0250***	-0.0249***	-0.0249***
	(0.0010)	(0.0010)	(0.0015)	(0.0010)	(0.0010)	(0.0016)
SINGLEH	0.3484***	0.3467***	0.3465***	0.3481***	0.3506***	0.3505***
	(0.0290)	(0.0291)	(0.0266)	(0.0290)	(0.0291)	(0.0224)
SINGLUN	0.4234***	0.4235***	0.4234***	0.4234***	0.4281***	0.4280***
	(0.0314)	(0.0314)	(0.0342)	(0.0314)	(0.0315)	(0.0303)
CHILD	0.0190*	0.0181	0.0180	0.0188*	0.0183	0.0183*
	(0.0114)	(0.0115)	(0.0127)	(0.0114)	(0.0115)	(0.0104)
NONCITIZ	0.0043	0.0028	0.0038	0.0033	0.0082	0.0075
	(0.0348)	(0.0355)	(0.0377)	(0.0348)	(0.0349)	(0.0315)
UNEMPL	0.2599***	0.2595***	0.2595 ***	0.2599***	0.2588 ***	0.2590***
	(0.0370)	(0.0371)	(0.0399)	(0.0370)	(0.0371)	(0.0491)
DISABL	0.2148*	0.2181*	0.2186 *	0.2154 *	0.2082*	0.2088*
	(0.1238)	(0.1239)	(0.1242)	(0.1238)	(0.1242)	(0.1208)
CPI	-0.0003	-0.0004	0.0000	-0.0003	-0.0014	-0.0006
	(0.0017)	(0.0017)	(0.0016)	(0.0017)	(0.0018)	(0.0020)
FA_RES			0.0194			0.1354**
			(0.0207)			(0.0633)
CONSTANT	-1.0144**	-0.7152	-0.7886*	-0.9566**	0.0034	-0.2408
	0.4355)	(0.5234)	(0.4740)	(0.4411)	(0.6314)	(0.4629)
Wald Joint	3,858.11	2,802.39	3,859.20	3,857.68	2,801.31	3,862.22
Pseudo R2	0.1813		0.1813	0.1812		0.1814
P-value Exog		0.3024			0.0331	
ALN test (p-val)		0.5879			0.3721	
F-stat (inst)		731.55			364.43	

*Note: *, **, and *** represent 10, 5 and 1% significance levels*; standard errors in parenthesis (St. Err for 2-SRI bootstrapped). State–level fixed effects coefficients omitted for brevity.
Wald Joint: Wald-test for joint significance of the model's coefficients.
P-value Exog: *p*-value of the Wald test of exogeneity for the suspected endogenous variable.
ALN test (p-val): *p*-value of the Amemyia-Lee-Newey minimum distance chi-square statistic.
F-stat (inst): *F*-statistic for test for joint significance of IVs coefficients in first stage equation.

Table 3b
Estimated coefficients equation 2: SMALL and GSCNV Full sample (N=36,887)

	SMALL			GSCNV		
Variables	Probit	IV-Probit	2SRI-Probit	Probit	IV-Probit	2SRI-Probit
FA	-0.0031	-0.1319**	-0.1302***	0.0123	0.0377	0.0372
	(0.0238)	(0.0523)	(0.0486)	(0.0169)	(0.0270)	(0.0305)
AGE	-0.0017*	-0.0015	-0.0015*	-0.0016*	-0.0016*	-0.0016**
	(0.0009)	(0.0009)	(0.0008)	(0.0009)	(0.0009)	(0.0008)
MALE	-0.1600***	-0.1616***	-0.1616***	-0.1600***	-0.1602***	-0.1602***
	(0.0227)	(0.0227)	(0.0197)	(0.0227)	(0.0227)	(0.0186)
HIGHSC	-0.0839**	-0.0836**	-0.0835**	-0.0840**	-0.0833**	-0.0833**
	(0.0370)	(0.0370)	(0.0335)	(0.0370)	(0.0370)	(0.0336)
SOMCOL	-0.0292	-0.0280	-0.0277	-0.0290	-0.0282	-0.0280
	(0.0374)	(0.0374)	(0.0333)	(0.0374)	(0.0374)	(0.0355)
COLMOR	-0.3474***	-0.3442***	-0.3438***	-0.3466***	-0.3442***	-0.3441***
	(0.0417)	(0.0418)	(0.0336)	(0.0417)	(0.0418)	(0.0391)
HISP	0.0397	0.0474	0.0486	0.0418	0.0471	0.0476
	(0.0362)	(0.0363)	(0.0345)	(0.0363)	(0.0366)	(0.0350)
ASIAN	-0.1423**	-0.1323*	-0.1320*	-0.1407**	-0.1376**	-0.1374*
	(0.0690)	(0.0691)	(0.0790)	(0.0691)	(0.0691)	(0.0764)
BLACK	0.1851***	0.1968***	0.1970***	0.1867***	0.1904***	0.1906***
	(0.0314)	(0.0317)	(0.0328)	(0.0314)	(0.0316)	(0.0305)
HMOWN	-0.3349***	-0.3395***	-0.3397***	-0.3354***	-0.3364***	-0.3366***
	(0.0249)	(0.0250)	(0.0288)	(0.0249)	(0.0249)	(0.0283)
INC_PR	-0.0250***	-0.0247***	-0.0247***	-0.0249***	-0.0248***	-0.0248***
	(0.0010)	(0.0010)	(0.0014)	(0.0010)	(0.0010)	(0.0013)
SINGLEH	0.3480***	0.3478***	0.3480***	0.3481***	0.3483***	0.3485***
	(0.0290)	(0.0290)	(0.0258)	(0.0290)	(0.0290)	(0.0267)
SINGLUN	0.4232***	0.4210***	0.4211***	0.4229***	0.4223***	0.4225***
	(0.0314)	(0.0314)	(0.0277)	(0.0314)	(0.0314)	(0.0353)
CHILD	0.0188*	0.0187	0.0188**	0.0190*	0.0193*	0.0194
	(0.0114)	(0.0115)	(0.0092)	(0.0114)	(0.0115)	(0.0124)
NONCITIZ	0.0034	0.0138	0.0146	0.0056	0.0102	0.0106
	(0.0348)	(0.0350)	(0.0360)	(0.0349)	(0.0352)	(0.0388)
UNEMPL	0.2599***	0.2604***	0.2602***	0.2597***	0.2599***	0.2599***
	(0.0370)	(0.0371)	(0.0437)	(0.0370)	(0.0370)	(0.0414)
DISABL	0.2155*	0.2175*	0.2166*	0.2154*	0.2158*	0.2155*
	(0.1238)	(0.1239)	(0.1123)	(0.1238)	(0.1238)	(0.1196)
CPI	-0.0003	0.0008	-0.0004	-0.0002	0.0000	-0.0005
	(0.0017)	(0.0018)	(0.0019)	(0.0017)	(0.0017)	(0.0016)
FA_RES			0.1569***			-0.0403
			(0.0590)			(0.0391)
CONSTANT	-0.9749**	-1.0533**	-0.3745	-0.8035*	-0.9662**	-1.0942**
	(0.4315)	(0.4327)	(0.5061)	(0.4119)	(0.4332)	(0.4430)
Wald Joint	3857.67	2808.49	3865.20	3858.18	2803.83	3859.54
Pseudo R2	0.1812		0.1816	0.1813		0.1813
P-value Exog		0.0056			0.2286	
ALN test (p-val)		0.2836			0.1182	
F-stat (inst)		2415.94			4460.13	

*Note: *, **, and *** represent 10, 5 and 1% significance levels*; standard errors in parenthesis (St. Err for 2SRI bootstrapped). State–level fixed effects coefficients omitted for brevity
Wald Joint: Wald-test for joint significance of the model's coefficients
P-value Exog: *p*-value of the Wald test of exogeneity for the suspected endogenous variable;
ALN test (p-val): *p*-value of the Amemyia-Lee-Newey minimum distance chi-square statistic
F-stat (inst): *F*-statistic for test for joint significance of IVs coefficients in first stage equation

spuriously correlated with adult food insecurity and that each food outlet impacts the likelihood of a household to experience adult FI differently.[24]

The estimated probit WMSC coefficient is positive (0.0036) but not statistically significant; after correcting for endogeneity, the coefficient becomes negative but remains not statistically significant (-0.0150 and -0.0162 for IV-probit and 2-SRI probit, respectively). The Wald test for exogeneity, as well as the lack of statistical significance of the first-stage residual coefficient in the 2SRI probit, suggests that WMSC may need not to be treated as endogenous. However, the value of the ALN minimum chi-square statistic being 0.294 (*p*-value 0.5897) shows that the instruments used are orthogonal to the error terms of the second stage equation and the value of the *F*-stat for the joint significance of the instruments' estimated coefficients in the first stage regression is 731.55, dismissing the risk of having weak instruments. As similar results persist across subsamples we find little to no evidence that Wal-Mart's presence contributes to ease adult food insecurity. This result is counterintuitive, as it is not consistent with the expectations that, thanks to the availability of larger assortment (Hausman and Leibtag, 2007), or its documented pro-competitive effect (Basker, 2005; Volpe and Lavoie, 2008; Basker and Noel, 2009; Cleary and Lopez, 2011) the company should indeed promote lower levels of FI. A more thorough investigation of this result and of its likely causes is provided in section 4.4.

The presence of medium-large sized grocery stores per 100,000 people appears instead to be associated with lower levels of FI. The estimated probit coefficient is much smaller than the IV-probit and 2SRI-probit ones (respectively, -0.0022, -0.1331, and -0.1305), suggesting that reverse causality is considerably biasing the probit coefficient. For this variable, we find strong evidence supporting the necessity of correcting for endogeneity: the *p*-value of the Wald test of exogeneity is 0.0331 (i.e. we reject the null of exogeneity), and the coefficient of the first stage residuals in the 2SRI-probit results is

---

[24]Full sets of results for the other samples are omitted for brevity and available upon request to the authors. The values of the Pseudo R^2 appear stable: for the All HH, Low-Income sample its value is of approximately 0.045; HH without children, Full circa 0.174; HH without children, Low-income, 0.044; HH with Children, Full, 0.20; HH with Children, Low-Income 0.056.

statistically significant at the 1 percent level. Also, the identification strategy seems valid: the *p*-value of the ALN statistic is 0.3712, and we observe a large *F*-stat for the joint significance of the instruments in the first stage-regression (634.43).

The estimated coefficient for SMALL in the full sample of all HH is also affected by endogeneity bias; both the small *p*-value of the exogeneity test in the IV-probit (0.0056) and the 1 percent statistical significance of the first stage residuals in the 2SRI-probit, indicate that this variable is likely to be endogenous. The probit estimated coefficient for this variable is not statistically significant, while both the IV-probit and the 2SRI probit coefficients are statistically different than zero at the 1 percent level and show similar magnitude (-0.13 circa), indicating that a larger presence of per-capita convenience and proximity stores have, among the All HH sample, a mitigating effect on adult FI. Lastly, using the full sample we find neither evidence of convenience stores attached to gas station (GSCNV) to be endogenous to the likelihood of being food insecure, nor do we find a statistically significant impact. Although this result seems to suggest that, on average, this food outlet is unlikely to be associated with adult FI, the results differ across subsamples, as will be discussed below.

The estimated coefficients assessing the effect of household characteristics on the probability of observing adult FI, are mostly consistent with the characteristics of FI households highlighted in other studies (see, e.g., Daponte and Stephens, 2004; Bartfeld and Dunifon, 2006; Nord et al., 2004; Nord et al., 2010). Furthermore, sign, magnitude, and overall significance of the estimates is largely unaffected by the use of different food access measures. The factors affecting adult FI likelihood in a negative and statistically significant way are: age of the household head, household head being male, higher levels of education in the household (in particular the COLMOR and SOMCOL dummies), the proxy for per-capita HH income, and homeownership. Factors showing instead a positive relationship with the likelihood of a household having experienced adult FI during the previous year are single head households, household head being Black, living in a single-unit household, number of children in the household, as well as the presence of unemployed persons in the household.

Table 4 reports the estimated parameters of the two-step IV-probit for the SMALL food access variable across subsamples. With the exception of the HH with Children subsample, the results of the Wald test for exogeneity support the need for correcting for the variables' endogeneity. The identification strategy used appears valid as *p*-values of the ALN chi-squared tests are above 0.1, although in the HH with Children-Low Inc sample the p-value is slightly below the rejection threshold (0.085). In all cases the instruments appear to have enough explanatory power (the values of the *F*-statistic are well above Staiger and Stock [1997] rule of thumb of 10), ruling out weak instruments' problems. The magnitude of the estimated SMALL coefficients across the four samples indicates a larger effect on low-income households, especially among HH with children: the coefficients are -0.1319 (All HH); -0.1317 (HH with Children); -0.3103 (All HH / Low-Income); and -0.3868 (HH with Children / Low-income). This result, which seems to hold also for GROC and, in the opposite direction, for GSCNV (see the discussion below) indicates that food access may affect more adult FI among low-income households, and those with children. This result suggests that the food environment can play a large role in mitigating FI for households facing hardships due to lack or resources (low income levels) or who have to provide nourishment for their children. In either case, improved access to some specific food outlets may help mitigate the cost of food sourcing.

The fact that FA has little to no impact on adult FI among HH without children (result which appears robust across store-types) suggests the existence of strong heterogeneity in how households are affected by the food environment. Households without children may have little incentive to explore the resources available to them to diminish their risk of being food insecure. However, the food environment does seem to be of particular importance for more numerous households as an improved food access could facilitate the adoption of economizing shopping habits (Leibtag and Kaufman, 2003).

Some variation in the magnitude and significance of the estimated coefficients for household characteristics indicates that the profile of FI households differs slightly across subsamples. Two striking features are that among the low-income sample, no level of education other than "College or higher" is associated with lower adult FI levels, and that the role of other demographic variables is weakened among

Table 4
IV-Probit Estimated Coefficients across Household Samples SMALL

	All HH		HH w/o Child		HH with Child	
Sample	Full	Low Inc	Full	Low Inc	Full	Low Inc
N obs	36,887	7,487	22,196	3,406	14,691	4,081
SMALL	-0.1319**	-0.3103***	-0.0780	-0.2124	-0.1317*	-0.3863***
	(0.0523)	(0.0949)	(0.0711)	(0.1319)	(0.0791)	(0.1388)
AGE	-0.0015	0.0012	0.0011	0.0013	-0.0029	0.0012
	(0.0009)	0.0014)	(0.0011)	(0.0018)	(0.0020)	(0.0028)
MALE	-0.1616***	-0.1318***	0.1306***	-0.0580	-0.1970***	-0.2164***
	(0.0227)	(0.0375)	(0.0294)	(0.0519)	(0.0374)	(0.0566)
HIGHSC	-0.0836**	-0.0406	-0.1169**	-0.0531	-0.0350	-0.0294
	(0.0370)	(0.0473)	(0.0512)	(0.0733)	(0.0544)	(0.0627)
SOMCOL	-0.0280	0.0362	-0.0574	0.0362	0.0614	0.0500
	(0.0374)	(0.0502)	(0.0515)	(0.0776)	(0.0555)	(0.0671)
COLMOR	-0.3442***	-0.2852***	-0.3611***	-0.3190***	-0.2002***	-0.2322**
	(0.0418)	(0.0656)	(0.0566)	(0.0950)	(0.0647)	(0.0945)
HISP	0.0474	-0.0689	0.1343***	0.0307	-0.0992*	-0.1460**
	(0.0363)	(0.0525)	(0.0517)	(0.0847)	(0.0525)	(0.0683)
ASIAN	-0.1323*	-0.1099	-0.1255	-0.2281	-0.1819*	-0.0135
	(0.0691)	(0.1018)	(0.0945)	(0.1600)	(0.1055)	(0.1363)
BLACK	0.1968***	0.0685	0.2287***	0.1563**	0.1125**	-0.0169
	(0.0317)	(0.0476)	(0.0424)	(0.0709)	(0.0488)	(0.0654)
HMOWN	-0.3395***	-0.2990***	-0.3470***	-0.3021***	-0.2619***	-0.2877***
	(0.0250)	(0.0387)	(0.0337)	(0.0611)	(0.0387)	(0.0523)
INC_PR	-0.0247***	-0.0239***	-0.0204***	-0.0165***	-0.0509***	-0.0450***
	(0.0010)	(0.0048)	(0.0010)	(0.0057)	(0.0026)	(0.0091)
SINGLEH	0.3478***	0.2144***	0.3137***	0.1505*	0.3067***	0.1893***
	(0.0290)	(0.0458)	(0.0472)	(0.0850)	(0.0404)	(0.0592)
SINGLUN	0.4210***	0.1349**	0.4001***	0.1292*		
	(0.0314)	(0.0548)	(0.0369)	(0.0733)		
CHILD	0.0187	0.0139			-0.0440**	-0.0070
	(0.0115)	(0.0157)			(0.0179)	(0.0220)
NONCITIZ	0.0138	0.0129	-0.0703	-0.0467	0.0586	0.0375
	(0.0350)	(0.0499)	(0.0504)	(0.0799)	(0.0505)	(0.0652)
UNEMPL	0.2604***	0.2454***	0.3177***	0.3318***	0.1781***	0.1733**
	(0.0371)	(0.0523)	(0.0509)	(0.0802)	(0.0551)	(0.0702)
DISABL	0.2175*	0.3207*	0.1525	0.2062	0.3116	0.4467*
	(0.1239)	(0.1721)	(0.1596)	(0.2474)	(0.2065)	(0.2485)
CPI	0.0008	0.0012	-0.0008	0.0010	0.0023	0.0015
	(0.0018)	(0.0032)	(0.0024)	(0.0045)	(0.0027)	(0.0046)
CONSTANT	-1.0533**	-1.3380*	-0.7123	0.9679	-0.8321	0.1318
	(0.4327)	(0.7644)	(0.5605)	(1.0297)	(0.7537)	(1.0794)
Wald joint	2808.49	355.76	1443.78	151.19	1372.01	245.20
P-value Exog	0.0056	0.0027	0.0735	0.0447	0.1728	0.0351
ALN test (p-val)	0.2836	0.3694	0.4078	0.7175	0.4370	0.0845
F-test (inst)	2415.94	324.33		178.75	934.50	146.40

*Note: *, **, and *** represent 10, 5 and 1% significance levels* – Standard errors in parenthesis. State–level fixed effects coefficients omitted for brevity

Wald Joint: Wald-test for joint significance

P-value Exog: *p*-value of the Wald test of exogeneity for the suspected endogenous variable;

ALN test (p-val): *p*-value of the Amemyia-Lee-Newey minimum distance chi-square statistic

F-stat (inst): *F*-statistic for test for joint significance of IVs coefficients in first stage equation low-income households, in particular that of ethnic profile. A similar diminished importance of the demographic variables seems also to be affecting adult FI among households with children.[25]

Before discussing FA's marginal effects, we provide an overview of the estimated food access parameters and validity of the identification strategies, summarized in Table 5. In the first place, we find no evidence of a statistically significant mitigating impact of Wal-Mart's presence on adult FI. Furthermore, the correction for endogeneity does not seem to be necessary in all the subsamples As the endogeneity tests performed are conditional on the identification strategy and instrument chosen, we experimented with different identifying assumptions with no major changes in outcome for the WMSC variable (see section 4.4 for a thorough discussion).

The estimated coefficients of GROC and SMALL obtained via IV-probit are negative and statistically significant. For these FA variables, the identification strategy adopted appears valid in the All HH and the HH with Children samples, although to a lesser extent in the latter: in the ALL HH subsamples the *p*-values of the Wald test of exogeneity and those of the ALN minimum distance chi-square statistic indicate the presence of endogeneity and provide evidence that the over-identifying instruments work appropriately to resolve the issue. In the HH with Children samples we find instead weaker evidence that correction for endogeneity is necessary, and in the low-income subsamples we also find little support that the instruments used fully resolve issues of spurious correlation that could bias the results of the model (the *p*-value of the ALN test are 0.010 for GROC and 0.085 for SMALL). That is, the estimated average marginal effects discussed below for the HH with Children subsample may be bias underestimated. Lastly, the values reports in Table 5 show that the estimated coefficients for GSCNV are positive and statistically significant in the low-income subsamples, where we also find support for the notion that this variable is endogenous (especially in the split samples); these results suggest that among low-income households, and in particular for households with children, the presence of a larger per-capita

---

[25]As in the case of the All HH sample, it should be pointed out that the behavior of the demographic variables' coefficients does not vary if other FA variables are used in place of SMALL.

Table 5
Summary of Results for FA Coefficients and Identification Strategy

Sample			WMSC	GROC	SMALL	GSCNV
All HH	Probit:	Coeff	0.0024	-0.0022	-0.0031	0.0123
N=36,887		(st.err.)	(0.0035)	(0.0124)	(0.0238)	(0.0169)
	IV-Probit:	Coeff	-0.0150	-0.1331**	-0.1319**	0.0377
		(st.err.)	(0.0173)	(0.0628)	(0.0523)	(0.0270)
	P-value Exog		0.3024	0.0331	0.0056	0.2286
	ALN test (p-val)		0.5879	0.3721	0.2836	0.1182
	F-stat (inst)		731.55	364.43	2415.94	4460.13
All HH	Probit:	Coeff	0.0101*	-0.0064	-0.0480	0.0628**
Low-Inc		(st.err.)	(0.0053)	(0.0180)	(0.0362)	(0.0255)
N=7,487	IV-Probit:	Coeff	-0.0251	-0.1633***	-0.3103***	0.1130***
		(st.err.)	(0.0302)	(0.0653)	(0.0949)	(0.0418)
	P-value Exog		0.2372	0.0119	0.0027	0.1277
	ALN test (p-val)		0.3589	0.1333	0.3694	0.4094
	F-stat (inst)		108.94	130.49	324.33	835.83
HH w/o Child	Probit:	Coeff	0.0036	-0.0188	0.0352	0.0136
N=22,196		(st.err.)	(0.0048)	(0.0169)	(0.0325)	(0.0232)
	IV-Probit:	Coeff	-0.0092	-0.1139	-0.0780	0.0001
		(st.err.)	(0.0237)	(0.0909)	(0.0711)	(0.0382)
	P-value Exog		0.5803	0.2864	0.0735	0.6212
	ALN test (p-val)		0.8126	0.4334	0.4078	0.2034
	F-stat (inst)		434.82	198.02	1465.85	2700.73
HH w/o Child	Probit:	Coeff	0.0153*	-0.0273	0.0274	0.0786**
Low-Inc		(st.err.)	(0.0079)	(0.0273)	(0.0549)	(0.0394)
N=3,406	IV-Probit:	Coeff	-0.0209	-0.1773	-0.2124*	0.0694
		(st.err.)	(0.0596)	(0.1132)	(0.1319)	(0.0643)
	P-value Exog		0.5390	0.1697	0.0447	0.8513
	ALN test (p-val)		0.1683	0.7907	0.7175	0.7275
	F-stat (inst)		28.12	44.14	178.75	386.04
HH with Child	Probit:	Coeff	-0.0025	0.0198	-0.0355	-0.0098
N=14,691		(st.err.)	(0.0053)	(0.0185)	(0.0356)	(0.0253)
	IV-Probit:	Coeff	-0.0160	-0.0828	-0.1317*	0.0425
		(st.err.)	(0.0256)	(0.0878)	(0.0791)	(0.0393)
	P-value Exog		0.5886	0.2312	0.1728	0.0819
	ALN test (p-val)		0.6228	0.1663	0.4370	0.4303
	F-stat (inst)		302.63	163.82	934.50	1727.80
HH with Child	Probit:	Coeff	0.0037	0.0153	-0.1146**	0.0452
Low-Inc		(st.err.)	(0.0072)	(0.0245)	(0.0489)	(0.0340)
N=4,081	IV-Probit:	Coeff	-0.0397	-0.1414*	-0.3863***	0.1406**
		(st.err.)	(0.0329)	(0.0810)	(0.1388)	(0.0557)
	P-value Exog		0.1751	0.0403	0.0351	0.0295
	ALN test (p-val)		0.8802	0.0010	0.0845	0.1132
	F-stat (inst)		97.20	86.28	146.40	447.29

*Note: *, **, and *** represent 10, 5 and 1% significance levels* – St.errors in parenthesis.
P-value Exog: *p*-value of the Wald test of exogeneity for the suspected endogenous variable.
ALN test (p-val): *p*-value of the Amemyia-Lee-Newey minimum distance chi-square statistic.
F-stat (inst): *F*-statistic for test of joint significance of IVs coefficients in first stage equations.

number of convenience stores attached to a gas station could exacerbate issues of adult FI, most likely due to a combination of non-ease of access, smaller assortments, and likely higher prices.

4.2 Marginal Effects, and Dollar Equivalent of Food Access

The average marginal effects of the food access variables obtained using the 2-SRI estimates are reported in Table 6; since the estimated FA coefficients for the samples of HH without children are not statistically significant, and the identification strategies are less effective, their discussion is omitted.

The marginal effects reported in Table 6 indicate that adding one additional Wal-Mart Supercenter per 1,000,000 individuals could, on average lead to a modest benefit in terms of reduction of adult FI likelihood. Such marginal effects, which vary from -0.2 percent and -1.24 percent for the All HH sample and the Low-Inc HH with Children sample, are however not statistically significant, and therefore not discussed further.

A marginal increase of one grocery store per 100,000 people leads to a decrease in the likelihood of being food insecure of circa -1.67 percent in the All HH sample, -1.17 percent in the HH with Children sample (not statistically different than zero), reaching values of -4.74 percent among low-income households, and -4.2 percent among low-income households with children; as this last effect may still be biased downward (as discussed previously, the p-value of the ALN test in this case is too small to warrant unbiasedness), the effect of GROC may be even larger among these households. These results indicate that an increase in the number of large food stores can have a substantial impact in reducing the likelihood of observing adult FI, especially among low-income households: the marginal effects described above represent the effect of approximately a 16 percent increase in the density of medium-large food stores. That is, if one could *double* the density of GROC, the likelihood of experiencing adult FI among low-income households could be reduced, on average, by more than 25 percent.

Increasing the number of proximity stores per 10,000 individuals (i.e., a one-unit increase of SMALL) leads to a considerable reduction in the probability of having experienced adult FI in the last 12 months, with values ranging from -1.67 percent (All HH sample) to -11.27 percent (HH with Children,

Table 6
Average Marginal Effects of FA Variables on Adult Food Insecurity (2SRI-probit)

Sample	All HH		HH with Children	
	Full	Low-Inc	Full	Low-Inc
WMSC	-0.0021	-0.0082	-0.0024	-0.0124
	(0.0026)	(0.0089)	(0.0039)	(0.0089)
GROC	-0.0167**	-0.0474***	-0.0117	-0.0420**
	(0.0079)	(0.0131)	(0.0107)	(0.0209)
SMALL	-0.0167***	-0.0896***	-0.0187*	-0.1127***
	(0.0062)	(0.0275)	(0.0105)	(0.0426)
GSCNV	0.0048	0.0332***	0.0061	0.0420**
	(0.0039)	(0.0105)	(0.0054)	(0.0202)

*Note: *, **, and *** represent 10, 5 and 1% significance levels* – Standard errors in parenthesis, obtained using the delta method.

Low-Income sample). As the average value of SMALL in the sample is approximately 2.35, such a marginal increase corresponds to a 42.5 percent variation in the number of these outlets per capita. Repeating the same calculation performed for GROC, doubling the number of proximity stores, the likelihood of adult FI would decrease, on average by -3.85 percent, while for low-income households with children, it could be lowered as much as approximately -26 percent. Lastly, an increase in one convenience store attached to gas stations (GSCNV) per 100,000 people would lead to a statistically significant increase in the likelihood of experiencing adult FI among low-income households by 3.3 percent (All HH sample) to 4.2 percent (HH with Children).

Table 7 reports the ratio of food access and income marginal effects calculated from the 2SRI-Probit estimates, capturing the dollar equivalent of the change in the likelihood of experiencing FI for a marginal change in food stores presence (measured in annual, per capita $ thousands). Another interpretation of this ratio is the amount of money that each member of a household should receive in order to maintain the same likelihood of experiencing adult FI if the number of per-capita stores of a given type decreases (or in the case of GSCNV increases) by one unit. No results for the "HH without Children" samples are discussed.

It should be noted that the estimated marginal effects of the per-capita income proxy on adult FI are very close when the same subsample of households is used, due to the high stability of the parameters highlighted in Table 3. These marginal effects indicate that increasing annual per-capita household income by an amount of $ 1,000, would lead to a small decrease of the likelihood of adult food insecurity among the full sample, equal to 0.3 percent (All HH Full sample), 0.7 percent (All HH, Low-income sample), 0.74 percent (HH with Children, full sample) and -1.3 percent circa (HH with Children, low-income sample).

The effect of a marginal increase in Wal-Mart Supercenters density on adult FI may (as the coefficients for WMSC are not statistically significant) be attached a limited annual monetary value, which varies from $328 to $1.162 per year (circa). The "value" of the reduction in adult FI from improving access to medium-large food stores per 100,000 people by one unit (an approximate 16 percent

Table 7
Average Marginal Effects of Food Access and Income; Monetary Value of Food Access' FI Reducing Effect (RATIO)

Sample	All HH Full			All HH Low-Inc			HH with Children Full			HH with Children Low-Inc		
	FA	Income	RATIO	FA	Income	RATIO	FA	Income	RATIO	FA	Income	RATIO
WMSC	-0.0021	-0.0032***	0.6477	-0.0082	-0.0070***	1.1623	-0.0024	-0.0074***	0.3279	-0.0124	-0.0143***	0.8689
	(0.0026)	(0.0002)		(0.0089)	(0.0012)		(0.0039)	(0.0004)		(0.0089)	(0.0023)	
GROC	-0.0167**	-0.0032***	5.2430	-0.0474***	-0.0067***	7.1156	-0.0117	-0.0073***	1.5954	-0.0420**	-0.0129***	3.2555
	(0.0079)	(0.0002)		(0.0131)	(0.0016)		(0.0107)	(0.0005)		(0.0209)	(0.0029)	
SMALL	-0.0167***	-0.0032***	5.2661	-0.0896***	-0.0070***	12.7410	-0.0187*	-0.0073***	2.5669	-0.1127***	-0.0135***	8.3731
	(0.0062)	(0.0002)		(0.0275)	(0.0013)		(0.0105)	(0.0006)		(0.0426)	(0.0031)	
GSCNV	0.0048	-0.0032***	-1.4998	0.0332***	-0.0068***	-4.8862	0.0061	-0.0073***	-0.8447	0.0420**	-0.0131***	-3.1934
	(0.0039)	(0.0002)		(0.0105)	(0.0016)		(0.0054)	(0.0005)		(0.0202)	(0.0017)	

*Note: *, **, and *** represent 10, 5 and 1% significance levels* – Standard errors in parenthesis, obtained using the delta method.
† The ratio of the food access and income marginal effects measures the increase in number of stores/population that will result in a benefit (in terms of reduction in food insecurity) equivalent to that of an increase in 1,000 $ of per-capita HH income.

increase in access) varies between $1.59 (HH with children, Full sample) and $7.11 thousands (All HH, low-income sample) per household member / per year.

Similarly, a marginal increase in convenience and proximity stores would result in a beneficial effect of a reduction of the likelihood of experiencing adult FI quantifiable in a range of values from $ 2.57 (HH with children, full sample) and $ 12.7 thousand in the All HH sample, low-income household. Although this amount may appear much larger than that obtained for GROC, it should be noted that a marginal, one-unit increase in SMALL correspond in a 42.5 percent increase in the number of small stores, while the estimated value of a marginal increase in GROC refers to a 16 percent increase in access. Accounting for these differences, one has that increasing SMALL by the same relative magnitude of a marginal increase in GROC (16 percent) would result in smaller monetary values (ranging approximately between $ 1,000 and $ 4,800). Lastly, the value associated with an increase in the probability of experiencing adult FI among low-income households for every additional convenience store attached to gas station per 100,000 people varies between -$3.19 and -4.9 thousands per household member per year.

4.3 Alternative Estimates and Robustness Checks

In this section we summarize the results of alternative estimation methods, and robustness checks that focus on model specifications, and identification strategies.

First, the model was re-estimated using a discrete adult FI indicator (Food secure and Marginal food secure = 0; Low Food Security = 1 and Very Low food Insecurity = 2) as dependent variables, in order to evaluate whether food access impacts differently adult FI likelihood across severity levels. We approached this task by relaxing the normality assumption and instead considering the probability of observing a given FI outcome (conditional on the explanatory variables), taking a logistic distribution[26] and the model estimated via fully unconstrained generalized-ordered logit (Williams, 2006) where the

[26] See footnote 9 for more details.

independent variables (i.e., FA) are allowed to impact the probability of observing different FI status in a non-proportional way (i.e., relaxing the proportional odds assumptions characterizing ordered probit and ordered logit estimators).[27] Since the model failed to converge for some subsamples, a constrained generalized ordered-logit was used instead. This estimator allows, via a backward stepwise selection process, to impose sequential constraints on the estimated coefficients that do not violate the proportional odds property, leaving the other ones unconstrained (Long and Freese, 2006). Across samples and FA measures, Wald-tests on the estimated parameters did not support the rejection of the null that the parameters jointly satisfy the proportional odds assumption.[28] In other words, we found no evidence that the impact of the covariates (in particular food access) would vary across adult FI status in a non-proportional way. Thus, the estimation of an ordered logit was performed: in Table 8 we present the average marginal effects of the four FA variables on the likelihood of a household showing increasing levels of adult FI obtained using a 2-SRI approach[29] to correct for the food access variables potential endogeneity (estimated coefficients and summary statistics are omitted for brevity).[30]

The values in Table 8 show that in the All HH full sample the marginal effect of the FA variables on the probability of observing VLFS is approximately halved compared to the effect on LFS. The same pattern in marginal effects is not observed for the low-income samples, where a decline is on average 27 percent. This suggests that, everything else constant, food access may have a relatively larger impact on the likelihood of observing VLFS for low-income households than on that of other households. Also, it

---

[27]The parameters of the ordered probit and logits are constrained to satisfy the proportional odds assumption, which can be quite restrictive if the independent variables' impact on the different levels of the outcome of the dependent variable in analysis are not proportional.

[28]Brant (1990) tests, comparing slope coefficients of a number of logits equal to the number of categories in the dependent variable minus one, were also attempted, using the "**brant**" STATA routine (Long and Freese, 2006); as in some cases the fully unconstrained generalized ordered logits failed to converge, it was impossible to systematically test the validity of the parallel regression assumptions using such approach.

[29]Terza, Basu, and Rathouz (2008) show that the 2SRI method produces unbiased and consistent estimates for a broad family of non-linear estimators, including the ordered logit.

[30]The values of the Pseudo R^2 are approximately 0.19 across All HH, HH with Children, and HH without Children samples, while they are around 0.05 in the Low-Income samples.

Table 8
Ordered Logit: Average Marginal Effects of Food Access on Different Levels of Adult Food Insecurity (2SRI-Ologit; ALL HH sample)

	WMSC		GROC		SMALL		GSCNV	
Sample	Full	Low-Inc	Full	Low-Inc	Full	Low-Inc	Full	Low-Inc
FI=0	0.0011	0.0092	0.0102**	0.0518***	0.0105***	0.0970***	-0.0030	-0.0351***
(*FS*)	(0.0013)	(0.0093)	(0.0045)	(0.0188)	(0.0038)	(0.0276)	(0.0020)	(0.0121)
FI=1	-0.0007	-0.0053	-0.0069**	-0.0299***	-0.0071***	-0.0561***	0.0020	0.0203***
(*LFS*)	(0.0009)	(0.0054)	(0.0030)	(0.0109)	(0.0026)	(0.0161)	(0.0013)	(0.0070)
FI=2	-0.0004	-0.0039	-0.0033**	-0.0219***	-0.0034***	-0.0409***	0.0010	0.0148***
(*VLFS*)	(0.0004)	(0.0039)	(0.0015)	(0.0080)	(0.0013)	(0.0117)	(0.0006)	(0.0051)

*Note: *, **, and *** represent 10, 5 and 1% significance levels* – Standard errors in parenthesis, obtained using the delta method.

should be noted that for the low-income sample in most cases the estimated marginal effects are of the FA variables on the outcome "FI =0", that is, that of a household being food secure, is very close in magnitude (but showing opposite sign) to those reported in Table 6. Given that the average marginal effects for the full samples in Table 8 are on average one-third smaller than those reported in Table 6, one could see the marginal effects of the FA variables on the "FS" outcome in Table 8 as more conservative estimates of such effects.

Second, as the food access measures used in this analysis are MSA-level aggregates, one could suspect that the estimated effects of FA could be capturing other factors affecting food insecurity (presence of public transportation; differences in infrastructure, etc.). Furthermore, as MSAs show heterogeneous distribution in terms of size, one could suspect that the effect of food access on FI could vary considerably depending upon the size of the aggregate considered (see, e.g., Ver Ploeg et al., 2009, for an illustration of how measures of food deserts differ across rural and nonrural areas and according to different contextualizations).

To account for this possibility we first attempted to estimate the model for the ALL HH samples separating the households in subsamples according to MSA-size indicators, using the classifications reported in the CPS.[31] Because of the limited variation in the MSA-level variables, especially for the large MSA-size subsamples, this option was discarded. A second attempt was that of using MSA-size level fixed effects in place of state-level ones; in this case the remaining unobserved heterogeneity in the data invalidated the identification strategies of the FA coefficients across samples. Third, the estimation was re-executed controlling for (1) MSA-size level fixed effects in addition to the state-level ones, and (2) interacting MSA-size dummies with the state-level number privately owned automobiles. In both cases the results were substantially unchanged as the MSA-size indicators, and their interactions with cars per

---

[31]Using the CPS classification of MSAs we obtained six MSA-Size class indicators identified by population brackets: *SizMSA1*: 100,000 – 249,999; *SizMSA2*: 250,000 – 499,999; *SizMSA3*: 500,000 – 999,999; *SizMSA4*: 1,000,000 – 2,499,999; *SizMSA5*: 2,500,000 – 4,999,999; *SizMSA6:* 5,000,000+

capita where (in their respective models) mostly not statistically significant. Also, attempts to use population density directly in the model (instead of using it as instrument for SMALL) resulted unfruitful.

Lastly, we evaluated the sensitivity of our results obtained via IV-probit and 2SRI-probit to instrument choice. Although no detailed results will be analyzed in this section, the full sets of results are available upon request from the corresponding author. Due to the unsatisfactory performance of the identification strategy for WMSC, most of the efforts were dedicated to explore alternative empirical approaches to appropriately identify the effect of this variable on adult FI.

Following the same logic illustrated in section 3.2, one can use information on the distance from food distribution centers as a predictor of the presence of Wal-Mart Supercenters in a given area. Therefore, we replicated the results discussed above in section 4.1, which used distance from Benton County as instrument for WMSC, with an inverse weighted average distance from the closest food distribution center,[32] as in Basker and Noel (2009), using the information contained in Holmes (2010) database (see Bonanno, 2010, for more details), along with the with the three-year lagged number of per-capita DSs. The results obtained were almost identical to those discussed above, in terms of power, validity of the identification strategy (or lack thereof), magnitude, and sign of the estimated coefficients. We then created 100-mile rings of distance from Benton County (as suggested by Neumark, Zhang and Ciccarella [2008] and Courtemanche and Carden [2011]) and used them in place of the "simple" distance from Benton County, also failing to observe the sign of the estimated parameter "switching" from positive to negative. Lastly, we adopted Basker and Noel's (2009) approach, described in the appendix of their paper and replaced the 3-year lagged value of per-capita number of DSs, with *historical* number of per-capita DSs (i.e., the value in years as early as 1989, when Wal-Mart started differentiating into food retailing): using this alternative method did not result in an improvement of the identification strategy.

[32]Holmes (2011) shows that a major driver of the company's location decision is the proximity of a distribution center, which allows the company to capitalize from economies of density.

As for the other FA variables, we attempted to use different combinations of retail cost variables; in place of the average Housing Price Index from the FHFA, we used the more detailed, state-level Home Price Index by Morris and Heathcote (2007) available from the Lincoln Institute of Land Policy; also, we replaced the electricity price for commercial use with the (division-level) Producer Price Index for Electricity (Commercial Use) from the BLS. The results were substantially unchanged in terms of magnitude and significance of the IV-probit coefficients although it appears that using the PPI for electricity results in stronger support towards GROC being endogenous. Lastly, as some states set very low (as low as zero) minimum corporate tax rates to attract small businesses, the minimum corporate tax rate for the lowest net income level was used as additional instrument for SMALL; similarly, as large companies will be discouraged from operating in areas where the business tax rates for large operations are higher, the corporate tax rate for the highest net income level is used as an additional instrument for GROC (both variables come from the Tax Foundation of the U.S. Bureau of Census); the results were basically unchanged when these variables were used.

4.4 Simultaneous Effects of Food Access on FI and Wal-Mart's Indirect Effects

The results discussed above are obtained including each FA variable one at the time in the FI equation. As illustrated at the beginning of section 4, this choice was forced by the limited variation in the FA data (at the MSA-level) and the state-level nature of most of the instrumental variables used. In this section we present the result of a specification of our model, estimated using the All HH full sample, where we account for the presence of all stores simultaneously and, at the same time, show one possible explanation for the counterintuitive result obtained above that the presence of per-capita Wal-Mart Supercenters was found to have no impact on FI.

First, we naively included all the FA variables in the model, without accounting for their endogeneity. The results in the first column of Table 9 show that the estimated probit coefficients for the FA variables are all statistically insignificant. We then attempted to correct for food access endogeneity using two-step IV-probit and two different approaches; we first used the full set of instruments described

Table 9
Selected Estimated Parameters of Model Specifications Including all the Food Access Variables (All HH sample)

	Probit	IV-Probit	IV-Probit2	2SRI-Probit
WMS	0.0014	-0.0180	-0.0272	-0.0603*
	(0.0047)	(0.0175)	(0.0283)	(0.0328)
SMALL	0.0032	-0.1280	-0.1565	-0.1576**
	(0.0258)	(0.0801)	(0.0933)	(0.0783)
GROC	-0.0005	-0.0419	-0.0278	-0.1241
	(0.0131)	(0.0650)	(0.0566)	(0.0986)
GSCNV	0.0090	0.0542	-0.0148	-0.0960*
	(0.0218)	(0.0512)	(0.0779)	(0.0588)
WMS_RES			0.0242	0.0624*
			(0.0325)	(0.0335)
GROC_RES				0.1283
				(0.1003)
SMALL_RES				0.1921**
				(0.0824)
GSCNV_RES				0.0729
				(0.0576)
Wald Joint	3858.3	2806.2	2806.7	3868.4
Pseudo R2	0.1813			0.1817
P-value Exog		0.0474	0.0324	
ALN test (p-val)		0.4147	0.3411	

*Note: *, **, and *** represent 10, 5 and 1% significance levels* – Standard errors in parenthesis. State-level fixed effects and household characteristics' coefficients omitted for brevity
Wald Joint: Wald-test for joint significance.
P-value Exog: p-value of the Wald test of exogeneity for the suspected endogenous variable.
ALN test (p-val): p-value of the Amemyia-Lee-Newey minimum distance chi-square statistic.

above to correct for the endogeneity of all the food access variables; then we included the residual of a first-stage regression of WMSC on its own instruments and the other exogenous variables in the FI equation (the results for this regression are reported in the first column of Table 10) while the endogeneity of the other variables is treated in the second stage equation. The main difference between the two approaches is that in the second Wal-Mart's presence (appropriately instrumented) is treated as a determinant of other stores' presence. The second and third columns of Table 9 contain the estimated parameters for the estimated models using the approached just illustrated. Although in both models the identification strategy (the *p*-value of the ALN test bare 0.4147 and 0.3441 in the two estimators, respectively), the lack of statistical significance of either all, or most, of the estimated FA coefficients (with the exception of SMALL using the second approach) prevents any reasonable inference. As already mentioned above, one possible reason for this outcome is the likely correlation of the different state-level variables used to account for the endogeneity of the different food access measures; also another likely cause is that some of the variables used as instrument for one store type may not work satisfactorily for others.[33]

As the two-step IV-probit routine used does not allow us to use separate sets of instruments for each one of the potentially endogenous variables in the model, we attempted an indirect approach. Each food access measure other than WMCS was regressed on all the exogenous variables in the FI equation, its instruments, and the predicted value of WMSC from the first-stage regression described above. The residuals from these regressions (including those of the WMSC first stage regression) were used along with the actual values in a 2SRI-probit. The detailed results of the first stage regression results are reported in Table 10; although they will not be discussed in detail, it is worthwhile to notice the negative and statistically significant coefficient of the predicted WMSC in the last three columns of Table 10. Namely, every additional Wal-Mart supercenter per 1,000,000 individuals (i.e., an increase of 22 percent

---

[33]This is, for example, the case of population density which, when included as instrument for GROC, did not result to be orthogonal to the error terms.

Table 10
First-Stage Regressions Using Instrumented WMSC (WMSC_HAT)

	WMSC	GROC	SMALL	GSCNV
WMSC_HAT		-0.2726***	-0.1285***	-0.0067
		(0.0069)	(0.0036)	(0.0044)
NDS_LAG	-0.2852***			
	(0.0073)			
DIST_BC	-0.0009***			
	(0.0002)			
AV_HPI		-0.0079***	-0.0001	0.0005*
		(0.0010)	(0.0002)	(0.0003)
PER_INC		0.0413***	0.0139***	-0.0522***
		(0.0009)	(0.0005)	(0.0006)
P_ELECT		0.0692***	-0.0110	0.0213**
		(0.0181)	(0.0084)	(0.0106)
LAND_SHARE		4.7589***		
		(0.5393)		
P_DIESEL		-2.5673***		
		(0.4873)		
POPDEN			0.3500***	
			(0.0044)	
CARS_PC				0.0215***
				(0.0003)
P_GAS				-0.3304**
				(0.1623)
AGE	-0.0034**	-0.0011***	-0.0000	-0.0007***
	(0.0013)	(0.0004)	(0.0002)	(0.0002)
MALE	0.0254	0.0056	-0.0048	-0.0001
	(0.0317)	(0.0085)	(0.0043)	(0.0053)
HIGHSC	-0.0409	-0.0272	-0.0040	-0.0034
	(0.0651)	(0.0175)	(0.0088)	(0.0109)
SOMCOL	-0.1497**	-0.0691***	-0.0123	-0.0165
	(0.0650)	(0.0175)	(0.0088)	(0.0109)
COLMOR	-0.3784***	-0.1164***	-0.0479***	-0.0237**
	(0.0664)	(0.0181)	(0.0091)	(0.0113)
HISP	-0.7849***	-0.3697***	-0.0977***	-0.1507***
	(0.0580)	(0.0165)	(0.0084)	(0.0103)
ASIAN	-0.4327***	-0.1836***	-0.0506***	0.0072
	(0.0843)	(0.0229)	(0.0115)	(0.0143)
BLACK	-0.5552***	-0.2198***	-0.0402***	-0.0399***
	(0.0520)	(0.0145)	(0.0073)	(0.0091)
HMOWN	0.1469***	0.0561***	0.0033	0.0170***
	(0.0389)	(0.0105)	(0.0053)	(0.0065)
INC_PR	-0.0107***	-0.0036***	-0.0013***	-0.0013***
	(0.0008)	(0.0002)	(0.0001)	(0.0001)
SINGLEH	-0.0843*	-0.0084	-0.0162***	-0.0062
	(0.0461)	(0.0124)	(0.0062)	(0.0077)
SINGLUN	0.0123	0.0438***	-0.0053	0.0166**
	(0.0419)	(0.0113)	(0.0057)	(0.0070)

(table continues)

Table 10, continued

	WMSC	GROC	SMALL	GSCNV
CHILD	-0.0538***	-0.0179***	-0.0094***	-0.0114***
	(0.0171)	(0.0046)	(0.0023)	(0.0029)
NONCITIZ	-0.5044***	-0.1298***	-0.0619***	-0.0784***
	(0.0524)	(0.0145)	(0.0073)	(0.0090)
UNEMPL	-0.0465	-0.0118	0.0024	-0.0085
	(0.0637)	(0.0171)	(0.0086)	(0.0107)
DISABL	0.1646	-0.0042	0.0554*	-0.0114
	(0.2130)	(0.0573)	(0.0288)	(0.0357)
CPI	-0.0259***	-0.0070***	0.0067***	-0.0120***
	(0.0024)	(0.0006)	(0.0003)	(0.0004)
CONSTANT	17.5397***	16.5630***	1.3713***	8.6841***
	(0.5780)	(1.2628)	(0.1359)	(0.3353)
Adj R-squared	0.6613	0.6572	0.8494	0.8386

in the density of this outlet), is associated with a decrease in the number of medium and large grocery stores per 100,000 people by 0.27 units (or 4.3 percent) while the number of proximity food stores per 10,000 people decreases by -0.1285 units (or 5.33 percent). WMSC shows no statistically significant impact on GSCNV.

The indirect 2SRI probit estimated coefficients, reported in the last column of Table 9, show a negative and statistically significant (at the 10 percent level) effect of WMSC on adult FI. Under the assumption that the indirect estimation approach used here is isolating the exogenous variations in the number of stores per capita effectively, while accounting for the indirect impact of Wal-Mart's presence on other stores, the results point to a statistically significant *direct* effect of the company's presence on mitigating adult FI. Also, the estimated parameter of SMALL is negative, statistically significant, and close in magnitude to that reported in the first column of Table 4; that of GROC shows a magnitude similar to the IV-probit result reported in Table 3a, although not statistically significant. Surprisingly, the results in this model indicate that GSCNV can have a beneficial (mitigating) effect on adult food insecurity. However, as the residuals of both GROC and GSCNV are not statistically different than zero, the results presented here should be considered only for illustration purposes.

Combining the average marginal effects calculated from the estimated parameters of the 2SRI-probit and the WMSC_HAT parameters from the first stage regressions of the other food access variables (see Table 10), one can observe the compounded direct and indirect effects of a marginal change in the per-capita number of Wal-Mart Supercenters as reported in Table 11.[34] In the first place, an additional Wal-Mart Supercenters for 1,000,000 people would *directly* reduce the likelihood of adult FI by 0.77

---

[34]Formally, one could express the sum of the direct and indirect effect of WMSC on adult FI probability as $\frac{\partial \Pr(FI_{il}=1 \mid \mathbf{Z},\hat{r})}{\partial WMSC} = \left.\frac{\partial \Pr(FI_{il}=1 \mid \mathbf{Z},\hat{r})}{\partial WMSC}\right|_{FA_j} + \sum_{j \neq WMSC} \frac{\partial \Pr(FI_{il}=1 \mid \mathbf{Z},\hat{r})}{\partial FA_j} \frac{\partial FA_j}{\partial WMSC}$, where the terms $\partial \Pr(FI_{il}=1 \mid \mathbf{Z},\hat{r}) \big/ \partial WMSC$ and $\partial \Pr(FI_{il}=1 \mid \mathbf{Z},\hat{r}) \big/ \partial FA_j$ represent WMSC direct marginal effect and the marginal effects of the other food access variable, respectively, and $\partial FA_j / \partial WMSC$ is the impact of WMSC on the other FA measures, from the first-stage regressions.

Table 11
Average Marginal Effects of Food Access on Adult Food Insecurity; Direct, Indirect and total Marginal Effects of WMSC (2SRI-Probit; ALL HH Sample)

	GROC	SMALL	GSCNV	WMSC
Direct FA	-0.0159	-0.0202**	-0.0123*	-0.0077*
Effect	(0.0126)	(0.0100)	(0.0075)	(0.0042)
Ind. WM Effect	0.0043	0.0026	0.0001	
Total WM Effect				-0.0007

*Note: *, **, and *** represent 10, 5 and 1% significance levels* – Standard errors in parenthesis, obtained using the delta method.

percent. Also, as a marginal increase in income results in a decrease in adult FI likelihood by -0.33 percent (not reported in Table 11), the value associated with a marginal increase in WMSC would be of $2.33 thousand per household member per year. The second row in Table 11 shows the product of each of the FA marginal effects times the respective WMSC_HAT coefficient from Table 9, all of them being positive (although that of GROC trivial in magnitude), which, when added to the direct marginal effects of WMSC, leads to a small cumulative average marginal effect of -0.07 percent (value in the bottom tight corner of Table 11). Thus, the cumulative (direct and indirect) marginal effects illustrated in this section suggest that the lack of statistical significance of the WMSC effects found above may be attributed to two opposite effects: a direct, FI mitigating effect of the company through lower prices and larger assortment, and an indirect, FI worsening effect, coming from the company's negative impact on other food outlets that show an FI easing effect.

5. CONCLUDING REMARKS

Although food insecurity afflicts a sizable portion of the U.S. population, especially among low-income individuals, and plentiful evidence exists pointing to the lack of adequate food access among the disadvantaged population, no empirical work had so far formally addressed (and quantified) whether a relationship exists between access to different types of food outlets and food insecurity.

Our empirical results indicate that improved food access helps mitigate the likelihood of adult food insecurity, especially among low-income households and those with children. In other words, food environments play an important role in determining food insecurity for those households that face hardships due to lack or resources (low income levels) or who have to obtain food and nourishment for their children. In either case, improved food access seems to help mitigating direct (prices) or indirect (transport, search) cost of food sourcing. Our results indicate that, among the different food stores considered, those with the largest impact are medium and large sized traditional food stores and proximity stores.

Although considerable efforts were made to eliminate sources of spurious correlation, the causality of such effects is conditional on the validity of the identification strategy used. If the adult food insecurity mitigating effect we find is in fact causal, our findings indicate that the renewed public interest (especially at the local level) to strengthen food systems and to improve food access for low-income individuals could, in fact, lead to reduction in food insecurity levels. However, in light of the different impacts across store types, one could envision that the development of policies aimed to increase access to large stores (via, for example, improvements in public transportation systems, less stringent zoning laws) could be an effective way to stimulate food security especially among low-income households.

References

Ailawadi, K. L., J. Zhang, A. Krishna, and M. W. Kruger. 2010. "When Wal-Mart Enters: How Incumbent Retailers React and How This Affects Their Sales Outcomes." *Journal of Marketing Research 47*(4):577–593.

Alvarez, R. M., and G. Glasgow. 1999. "Two-Stage Estimation of Nonrecursive Choice Models." *Political Analysis 8*:147–165.

Alwitt, L. F., and T. D. Donley. 1997. "Retail Stores in Poor Urban Neighbourhoods." *The Journal of Consumer Affairs 31*:139–163.

Amemiya, T. 1978. "The Estimation of a Simultaneous Equation Generalized probit model." *Econometrica 46*(5): 1193–1205.

Asplund, M., and R. Sandin. 1999. "The Number of Firms and Production Capacity in Relation to Market Size." *Journal of Industrial Economics* *47*(1): 69–85.

Ball, K., A. Timperio, and D. Crawford. 2008. "Neighborhood Socioeconomic Inequalities in Food Access and Affordability." *Health and Place 15*(2):578–585.

Bartfeld, J., and R. Dunifon. 2006. "State-Level Predictors of Food Insecurity among Households with Children." *Journal of Policy Analysis and Management 25*(4):921–942.

Basker, E. 2005. "Selling a Cheaper Mousetrap: Wal-Mart's Effect on Retail Prices." *Journal of Urban Economics 58*: 203–229.

Basker, E., and M. D. Noel. 2009. "The Evolving Food Chain: Competitive Effects of Walmart's Entry into the Supermarket Industry." *Journal of Economics and Management Strategy 18*(4):977–1009.

Baum, C. F., M. E. Schaffer, S. Stillman, and V. Wiggins. 2006. "Overid: Stata Module to Calculate Tests of Overidentifying Restrictions after ivreg, ivreg2, ivprobit, ivtobit, reg3." Retrieved December 6, 2011, from http://ideas.repec.org/c/boc/bocode/s396802.html

Berry, S. T. 1992. "Estimation of a Model of Entry in the Airline Industry." *Econometrica 60*: 889–917.

Bonanno, A. 2010. "An Empirical Investigation of Wal-Mart's Expansion into Food Retailing." *Agribusiness: An International Journal 26*(2):220–242.

Bonanno, A., and R. A. Lopez. 2009. "Competition Effects of Supermarket Services." *American Journal of Agricultural Economics 91*(3): 555–568.

Borjas, G. 2004. "Food Insecurity and Public Assistance." *Journal of Public Economics 88* (7): 1421–1423.

Brant, R. 1990. "Assessing Proportionality in the Proportional Odds Model for the Ordinal Logistic Regression." *Biometrics 46*: 1171–1178.

Bresnahan, T. F., and P. C. Reiss. 1991. "Entry and Competition in Concentrated Markets." *Journal of Political Economy 99*(5):977–1009.

Cleary, R., and R. A. Lopez. 2011. "Supermarket Responses to Wal-Mart Expansion: A Structural Approach." Selected Presentation, International Industrial Organization Conference, April 8 to 10, Boston, MA.

Cotterill, R. W., and A. W. Franklin. 1995. "The Urban Grocery Store Gap." Food Marketing Policy Issue Paper 8, Food Marketing Policy Center, University of Connecticut.

Courtemanche, C., and A. Carden. 2011. "Supersizing Supercenters? The Impact of Wal-Mart Supercenters on Body Mass Index and Obesity." *Journal of Urban Economics 69*(2): 165–181.

Cummins, S., and S. Macintyre. 2002. "'Food Deserts' – Evidence and Assumption in Health Policy Making. " *BMJ (Clinical Research Ed.) 325*(7361): 436–438.

Daponte, B. O., and M. Stephens. 2004. "The Relationship between Food Assistance, the Value of Food Acquired, and Household Food Security." Working paper, Joint Center for Poverty Research, University of Chicago, Chicago, IL.

Ellickson, P. B. 2006. "Quality Competition in Retailing: A Structural Analysis." *International Journal of Industrial Organization 24*(3): 521–540.

Ellickson, P. B. 2007. "Does Sutton Apply to Supermarkets?" *RAND Journal of Economics 38*(1): 43–59.

Food Marketing Institute. 2007. Top U.S. Supermarket and Grocery Chains. Retrieved December 7, 2007, from http://www.fmi.org/docs/facts_figs/faq/top_retailers.pdf

Gibson-David, C. M., and E. M. Foster. 2006. "A Cautionary Tale: Using Propensity Scores to Estimate the Effect of Food Stamps on Food Insecurity." *Social Service Review 80*: 93–126.

Goetz, S. J., and H. Swaminathan. 2006. "Wal-Mart and Family Poverty in US Counties." *Social Science Quarterly 83*: 211–225.

Gundersen, C., and V. Oliveira. 2001. "The Food Stamp and Food Insufficiency." *American Journal of Agricultural Economics 83*(4): 875–887.

Haering, S. A. and S. B. Syed. 2009. "Community Food Security in the United States Cities: A Survey of the Relevant Scientific Literature." Center for a Livable Future, Johns Hopkins Bloomberg School of Public Health. Retrieved September 29, 2011, from http://www.jhsph.edu/bin/s/c/FS_Literature%20Booklet.pdf

Haltiwanger, J., R. Jarmin, and C. J. Krizan. 2010. "Mom-and-Pop Meet Big-Box: Complements or Substitutes?" *Journal of Urban Economics 67:* 116–134.

Hausman, J. A. 1978. "Specification Tests in Econometrics." *Econometrica 46*: 1251–1271.

Hausman, J. A., and E. S. Leibtag. 2007. "Consumer Benefits from Increased Competition in Shopping Outlets: Measuring the Effect of Wal-Mart." *Journal of Applied Econometrics 22*: 1157–1177.

Holmes, T. J. 2010. "Opening Dates of Wal-Mart Stores and Supercenters, 1962-Jan 31, 2006." Retrieved from http://www.econ.umn.edu/~holmes/data/WalMart/index.html

Holmes, T. J. 2011. "The Diffusion of Wal-Mart and Economies of Density." *Econometrica 79*: 253–302.

Jensen, H. H. 2002. "Food Insecurity and the Food Stamp Program." *American Journal of Agricultural Economics 84*(5): 1215–1228.

Jia, P. 2008. "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retail Industry." *Econometrica 76*(6): 1263–1316.

Kabbani, N. S., and M. Y. Kmeid. 2005. "The Role of Food Assistance in Helping Food Insecure Households Escape Hunger." *Review of Agricultural Economics 27*: 439–445.

King, R. P., E. S. Leibtag, and A. S. Behl. 2004. *Supermarket Characteristics and Operating Costs in Low-Income Areas*. USDA Economics Research Service Agricultural Economic Report No. 839.

Lee, L. F. 1992. "Amemiya's Generalized Least Squares and Tests of Overidentification in Simultaneous Equation Models with Qualitative or Limited Dependent Variables." *Econometric Reviews 11*(3): 319–328.

Leibtag, E. L., and P. R. Kaufman. 2003. "Exploring Food Purchase Behavior of Low-Income Households How Do They Economize?" USDA Economics Research Service, Agriculture Information Bulletin No. 747-07. Retrieved from www.ers.usda.gov/publications/aib747/aib74707.pdf

Lincoln Institute of Land Policy. Land and Property Values in the U.S.: Land Prices by State Database. Retrieved July 30, 2011, from http://www.lincolninst.edu/subcenters/land-values/land-prices-by-state.asp

Long, J. S., and J. Freese. 2006. *Regression Models for Categorical Dependent Variables Using Stata* (2nd ed.). College Station, TX: Stata Press.

Mazzeo, M. J. 2002. "Product Choice and Oligopoly Market Structure." *The RAND Journal of Economics 33*:221–242.

Moore, L. V., and A. V. Diez Roux. 2006. "Associations of Neighborhood Characteristics with the Location and Type of Food Stores." *American Journal of Public Health 96* (2): 325–331.

Morland, K., S. Wing, and A. Diez Roux. 2002. "The Contextual Effect of the Local Food Environment on Residents' Diets: The Atherosclerosis Risk in Communities (ARIC) Study." *American Journal of Public Health 92*: 1761–1767.

Newey, W. K. 1987. "Efficient Estimation of Limited Dependent Variable Models with Endogeneous Explanatory Variables." *Journal of Econometrics 36*(3): 231–250.

Nelson, M. 2000. "Childhood Nutrition and Poverty." *The Proceedings of the Nutrition Society 59*(2): 307–315.

Neumark, D., J. Zhang, and S. Ciccarella. 2008. "The Effects of Wal-Mart on Local Labor Markets." *Journal of Urban Economics 63*: 405–430.

Nord, M. 2002. "A 30-Day Food Security Scale for CPS Food Security Supplement Data" E-FAN (No. 02-015).Retrieved January 10, 2010, from http://www.ers.usda.gov/publications/efan02015/efan02015.pdf

Nord, M., M. Andrews, and S. Carlson. 2004. *Household Food Security in the United States, 2003*. Food Assistance and Nutrition Research Report No. 42, Food and Rural Economics Division, U.S. Department of Agriculture Economics Research Service.

Nord, M., A. Coleman-Jensen, M. Andrews, and S. Carlson. 2011. *Household Food Security in the United States, 2010*. Economic Research Report No. (ERR-127). Retrieved September 20, 2011, from http://www.ers.usda.gov/Publications/err125/

Nord, M., A. Coleman-Jensen, M. Andrews, and S. Carlson. 2010. *Household Food Security in the United States, 2009*. Economic Research Report No. (ERR-108). Retrieved December 2, 2011, from http://www.ers.usda.gov/Publications/err108/

Nord, M., and A. M. Golla. 2009. *Does SNAP Decrease Food Insecurity? Untangling the Self-Selection Effect*. U.S. Department of Agriculture Economic Research Service, Report N.85. Retrieved January 10, 2010, from http://www.ers.usda.gov/Publications/ERR85/ERR85.pdf

Petrin, A., and K. Train. 2010. "A Control Function Approach to Endogeneity in Consumer Choice Models." *Journal of Marketing Research 47*(1): 3–13.

Powell, L. M., S. Slater, D. Mirtcheva, Y. Bao, and F. J. Chaloupka. 2007. "Food Store Availability and Neighborhood Characteristics in the United States." *Preventive Medicine 44*: 189–195.

Ratcliffe, C., S. McKernan, and S. Zhang. 2011. "How Much Does the Supplemental Nutrition Assistance Program Reduce Food Insecurity?" *American Journal of Agricultural Ecomomics 93*(4): 1082–1098.

Rivers, D., and H. Vuong. 1988. "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models." *Journal of Econometrics 39*: 347–366.

Rose, D., and R. Richards. 2004. "Food Store Access and Household Fruit and Vegetable Use among Participants in the US Food Stamp Program." *Public Health Nutrition 7*(8): 1081–1088.

Rose, D., G. Gundersen, and V. Oliveira. 1998. "Socio-Economic Determinants of Food Insecurity in the United States: Evidence from the SIPP and CSFII Datasets." Technical Bulletin No.1869, Food and Rural Economics Division, U.S. Department of Agriculture Economic Research Service.

Scott Kantor, L. 2001. "Community Food Security Programs Improve Food Access." *Food Review 24*(1): 20–26.

Seim, K. 2006. "An Empirical Model of Firm Entry with Endogenous Product – Type Choices." *The RAND Journal of Economics 37*: 619–640.

Singh, V. P., K. T. Hansen, and R. C. Blattberg. 2006. "Market Entry and Consumer Behavior: An Investigation of a Wal-Mart Supercenter." *Marketing Science 25*(5): 457–476.

Staiger, D., and J. H. Stock. 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica 65*(3): 557-586.

Shaked, A., and J. Sutton. 1987. "Product Differentiation and Industrial Structure." *Journal of Industrial Economics 36*(2): 131–146.

Sutton, J. 1991. *Sunk Cost and Market Structure: Price Competition, Advertising, and the Evolution of Concentration*. Cambridge, MA: MIT Press.

Terza, J. V., A. Basu, and P. J. Rathouz. 2008. "Two-Stage Residual Inclusion Estimation: Addressing Endogeneity in Health Econometric Modeling." *Journal of Health Economics 27*: 531–543.

U.S. Bureau of Census and U. S. Bureau of Labor Statistics, Current Population Survey - Food Security Supplement 2004, 2005. Retrieved from http://www.bls.gov/cps/

U.S. Bureau of Census. Population Estimates Program: 2004, 2005.

U.S. Bureau of Census. County Business Pattern 2004, 2005.

U.S. Bureau of Labor Statistics. Consumer Price Index. Retrieved from http://www.bls.gov/cpi/

U.S. Bureau of Census. 2001. *Gazetteer of Counties*: Year 2000. Retrieved April 1, 2006 from http://www.census.gov/geo/www/gazetteer/places2k.html

U.S. Department of Energy. Energy Information Administration. Refiner Gasoline Prices by Grade and Sales Type – Regular.

U.S. Department of Energy. Current and Historical Monthly Retail Sales, Revenues and Average Revenue per KWH by State and by Sector (Form EIA-826) - commercial prices.

U.S. Department of Energy. Weekly Retail Gasoline and Diesel Prices: Diesel - All Types. Retrieved from http://www.eia.doe.gov/dnav/pet/pet_pri_gnd_a_epd2d_pte_dpgal_a.htm

U.S. Department of Transportation, Federal Highway Administration, Highway Statistics: Section II, Motor Vehicles. State Motor-vehicle Registrations. (2004 and 2005). Retrieved November 10, 2006, from http://www.fhwa.dot.gov/policyinformation/statistics.cfm .

U.S. Federal Housing Financing Agency. House Price Index data. Purchase Only Indexes. Retrieved March 29, 2011, from http://www.fhfa.gov/Default.aspx?Page=87

U.S. Tax Foundation. Corporate Income Tax Rates 2000-2011. Retrieved June 1, 2010, from http://www.taxfoundation.org/taxdata/show/230.html

Ver Ploeg, M., V. Breneman, T. Farrigan, K. Hamrick, D. Hopkins, P. Kaufman, B. Lin, M. Nord, T. Smith, R. Williams, K. Kinnison, C. Olander, A. Sing, and E. Tuckermanty. 2009. "Access to Affordable and Nutritious Food—Measuring and Understanding Food Deserts and Their Consequences: Report to Congress." Administrative Publication No. (AP-036), U.S. Department of Agriculture, Economic Research Service. Retrieved from http://www.ers.usda.gov/Publications/AP/AP036/

Volpe III, R. J., and N. Lavoie. 2008. "The Effect of Wal-Mart Supercenters on Grocery Prices in New England, Richard." *Review of Agricultural Economics 30*(1): 4–26.

Wal-Mart Stores Inc. 2011. 2010 Annual Report. Retrieved March 29, 2011, from http://walmartstores.com/sites/annualreport/2010/

Walton, S., and J. Huey. 1992. *Sam Walton, Made in America: My Story*. New York City, NY: Doubleday Publisher.

Williams, R. 2006. “Generalized Ordered Logit/Partial Proportional Odds Models for Ordinal Dependent Variables.” *The Stata Journal 6*: 58–82.

Wooldridge J. M. 2002. *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: MIT Press.

Yen, S. T., M. Andrews, Z. Chen, and D. B. Eastwood. 2008. “Food Stamp Program Participation and Food Insecurity: An Instrumental Variables Approach.” *American Journal of Agricultural Economics 90*(1): 117–132.

Zenk, S., A. J. Schulz, B. A. Israel, S. A. James, S. Bao, M. L. Wilson. 2005. “Neighborhood Racial Composition, Neighborhood Poverty, and the Spatial Accessibility of Supermarkets in Metropolitan Detroit.” *American Journal of Public Health 95*(4): 660–667.