ASSA Conference, San Francisco 3-5 january 2016

Can the Rise in Obesity in France be Blamed on the Food Environment? Evidence from French Urban Data

France Caillavet, Gayaneh Kyureghian*, Rudy Nayga, Pierre Chauvin

*Gayaneh Kyureghian – Corresponding Co-author

Department of Food and Resource Economics

Korea University

Anamdong-5-1, Seongbukgu

Seoul, 136-701, South Korea

E-mail: gkyuregh@korea.ac.kr

Phone: +82-2-2390-4623

Fax: +82-2-928-3584

France Caillavet is senior economist, French National Institute for Agricultural Research within the ALImentation et Sciences Sociales unit research (Unité de Recherche 1303, Food and Social Sciences) in Paris area. Gayaneh Kyureghian is assistant professor, Department of Food and Resource Economics, Korea University. Rodolfo M. Nayga, Jr. is Professor and Tyson Endowed Chair, Department of Agricultural Economics and Agribusiness, University of Arkansas, and adjunct professor, Korea University and Norwegian Agricultural Economics Research Institute. Pierre Chauvin is Research Director, French National Institute of Health and Medical Research (INSERM)

and head of the Department of Social Epidemiology, Pierre Louis Institute of Epidemiology and Public Health, INSERM & Sorbonne Universités UPMC, Paris, France. We acknowledge Coline Ferrant, graduate student at Sciences Po, Observatoire sociologique du changement, CNRS, for helpful assistance with data building. This research has been funded by the French National Research Agency (ANR) in the Framework of the Programme National de Recherche en Alimentation et nutrition, the French Institute of Public Health Research (IRESP) and the French Interministerial Committee of Urban Affairs. SIRS survey received the approval of the CNIL (Commission Nationale Informatique et Liberté): authorization No. 051024 following the request No. 904851 of the 17 June 2004.

Can the Rise in Obesity in France be Blamed on the Food Environment?

Evidence from French Urban Data

Abstract

We examine the effect of the food retail environment on consumers' body mass in France, a country that has experienced rising obesity rates in the last 15 years. We measure food availability in several ways – number of retail outlets, the combined shopping service area, the average surface area of retail outlets, as well as the retailer type mix expressed as the number of hypermarkets, number of supermarkets, and number of small grocery stores. We use a unique panel data set with information on body mass of a sample of approximately 1250 individuals at two points of time – 2005 and 2010, in the Parisian region. We also use different datasets to test the effect of different measures of food availability, and different geographic units or reference areas as well as different model estimations to test the robustness of our results. Our results do not generally suggest systematic or strong influence of any of the food environment measures on body mass. However, we robustly detect small negative effects of number of small grocery stores and small positive effects of number of hypermarkets on body mass. The findings imply that changing the food environment to counter the rise in obesity in France may be misguided.

Keywords: Food access, availability, obesity, endogeneity, French food retail.

JEL Codes: D12, L66, P46.

Introduction

There is evidence suggesting an increase of food retail availability internationally (e.g. for the case of the U.S. see Thomsen, Kyureghian and Nayga 2014). The French food retail industry seems to follow the suit as well (figure 1). Despite the fact that the French retail industry is highly regulated and in spite of the cessation of one of the prominent laws in the country – Raffarin, which subjected large retail openings to zoning committee regulations until 2008 (Caillavet *et al.* 2015), the trend is unmistakable given the increasing number of retail store openings in France. One caveat however is that due to the termination of the Raffarin law in France in 2008, the openings of large retail stores did not increase and remained essentially flat since 2008. Consequently, the average number of population per food outlet dropped by almost a half, while the same measure for large grocery stores stayed almost identical[i].

Yet another ubiquitous and also well documented trend is the rise of the body weight worldwide. Although the rise of obesity rates in France is not as worrisome as in other OECD countries (OECD 2015) or in the U.S., its increasing trend has been of great concern in France lately. It became a national priority with the launching of the Obesity Plan, launched in 2010 by the French Ministry of Health. Medical care expenditures and insurance costs attributed to obesity *stricto sensu* have been reported to reach 2.1 billion euros[ii]. Emery et al. (2007) estimate this cost to be approximately 4.5% of health expenditures.

ObEpi surveys, conducted every 3 years from 1997 on adults 18 years and over, showed that the obesity rate in 2012 is 15%. The mean BMI has been increasing from 24.3 kg/m2 in 1997 to 25.4 kg/m2 in 2012. Hence, the obesity rate in France has been rapidly increasing in the last 15 years (+76.5%) (see tables 1 and 2), and this is particularly accentuated in the Parisian region. Although the rate of increase of obesity prevalence has slowed down a little between 2009 and 2012, the goal of stopping this trend has not yet been reached. To make things even more worrisome, the people in lower socioeconomic status seem to be bearing the burden. Relative education and socio-economic status related inequalities in obesity and overweight are large in France, compared to other OECD countries (Devaux and Sassi 2011).

Concerning the time cost of food access, one of the measures used in past literature is the distance to food outlets, with a distinction between healthy or unhealthy food (e.g., fast-food) providers. Several economic studies have been published on this topic with U.S. data (Dunn 2010; Michimi *et al.* 2010; Anderson and Matsa 2011; Courtemanche and Carden 2011, in the case of adults; Currie *et al.* 2010 and Thomsen *et al.* 2015 in the case of children). They found very different and inconsistent results. In France, associations between weight and food retail accessibility were studied only in a few epidemiologic works, with limited effects found. For example, a study on middle-school students in Eastern France found that the likelihood of being overweight was higher when spatial accessibility to general food outlets was low, for children of blue-collar workers only. No significant relationship was found with other food outlets such as bakeries or fast-food outlets (Casey *et al.* 2012). In another cohort study (Parisian region, 2007-08), participants shopping in hard discount supermarkets and in supermarkets whose catchment area comprised low education residents were found to have a higher BMI or waist circumference, after adjustment for individual and residential neighborhood characteristics (Chaix *et al.* 2012). Closer to our study, a study using the 2005 wave of *Santé, Inégalités et Ruptures Sociales* (SIRS) cohort in the Parisian region found associations between the obesity rate (BMI ≥30) and the residential environment, after adjusting for age, gender, education and income levels. Specifically, the distance to specialty stores (baker, butcher, fish shop), a higher proportion of fast-foods among restaurants environment, or a lower number of retail shops and services in the residence area (within a radius of 500m) were associated with a higher risk of obesity (Cadot *et al.* 2011).

There has also been much attention paid to the effects of large retail stores on diet quality and health outcomes. Economic theory predicts that the consumption of food, being an ordinary and normal good, should increase in response to more affordable prices and increased affluence, all else equal. Retailing activities on a large scale seem to endorse this, backed by a large body of literature on lower supercenter prices on most, including healthy, foods (Hausman and Leibtag 2007; Volpe, Okrent and Leibtag 2014; Courtemanche and Carden 2011), and hence the plausibility of supercenters and large retailer possibly being responsible for the increasing waistlines. The objective we share with these studies is to estimate the effect of food retail outlets, particularly the large ones, on weight. One of the contributions of this study is the focus on a country (i.e., France) known for its sophisticated food culture but has experienced increases in obesity rates in the last 15 years, as discussed above. We use a unique data set on a panel of approximately 1,250 individuals that, in addition to reporting a number of socio-demographic variables, also reported their height and weight in 2005 and 2010. We also account for the food retail environment, including street markets/stands and food away from home joints using data obtained from the business registry records at Institut National de la Statistique et des Etudes Economiques (INSEE). Finally, following Caillavet *et al.* (2015), we use alternative reference geographic areas for food environment.

A common obstacle to achieving our objective is the identification problem – typically it is hard to find a source of exogenous variation to disentangle the food environmental effects from other confounding effects. In our data, we are able to track the same individual twice at different times and so this allows us to factor out and control individual differences. Additionally, none of the panel members in our data has moved residence from 2005 to 2010, making the endogeneity of the retail environment less of a threat to identification of the food environmental effects. Nevertheless, we go a step farther and use instrumental variable technique to rule out potential endogeneity concerns. As a source of exogenous variation, we use the dispersion of new openings across administrative entities (arrondissements) within the administrative unit overseeing such decision whether or not to grant permission(departments). This pertains to the supply-side factors of food retail supply (Caillavet *et al.* 2015).

Our results indicate that even though there is weak evidence that the total number of retail outlets affect the weight, there seems to be systematic an inverse relation between the number of small outlets and weak and occasional positive association between the number of large outlets and body mass. Our results suggest that the combined retail surface area or the average store surface area variables are not associated with the changes in body mass.

Data

We draw information from several data sources. The data on weight and individual demographic information come from SIRS.

Height, Weight and Demographic Data

This study is based on the Health, Inequalities and Social Ruptures cohort of SIRS which contains a representative sample of over 3,000 French-speaking adults in the Paris metropolitan area. It covers 4 departments: Paris (75), Hauts-de-Seine (92), Seine St-Denis (93), Val-de-Marne (94), a region with a population of 6.5 million.

SIRS cohort has three waves. The 2005 data includes a 3-level random sample. First, 50 census blocks were randomly selected using a stratification based on their socioeconomic type and their being labelled or not labelled as “underprivileged areas”. Next, 60 households were randomly chosen from a complete list of households within each selected census block. Lastly, one adult was randomly selected from each household. The second wave took place in 2007 and the 3rd in 2009-2010, predominantly based on the 2005 population sample. In the latter wave, 47% of the initial respondents were re-interviewed (2.6% were deceased, 1.8% were too sick to answer questions, 2.7% were absent during the survey period, etc.). Their sex ratio and mean age were similar to those who were re-interviewed. The survey includes over 400 variables relating to among others, health and socio-economic variables. It also includes detailed and geocoded addresses of the respondents. The sample sizes in 2005 and 2010 were 1,430 and 3,006, respectively. Respondents with missing observations (e.g., missing height, weight, age) were removed from the sample, resulting in the final sample size of 1,248 individuals. Weight and height were self-reported in both 2005 and 2010 waves and so it is possible that these measures are misreported (see Cawley 2015).

Finally, we use this data set for detailed information on respondents’ precise residential location and the distance to the closest metro station. We then use these radii to construct circles around the respondents’ residences and the metro stations as alternative reference areas to refrain from reliance on administrative geographic delineations, given that these has been rightfully criticized in the literature so often. The description and summary statistics of body mass and demographic variables are reported in table 3. The above mentioned reference areas are referred to resident- or metro station-centered circles. The mean BMI in our sample – 24.8, is close to the national mean reported here. In our panel, 38.7% are males and the majority of the sample has higher education attainment and is full or part-time employed. The mean household size is 2.2 persons, with a mean household income of approximately 3,000 euros/month. Finally, the mean age in the sample is 52.

Food Retail Data

The data on food retail environment come from two sources – Trade Dimensions (TDLinx) and SIRENE maintained by Institut National de la Statistique et des Etudes Economiques (INSEE). The former data set provides detailed information on most food retail outlets – complete addresses, down to longitude and latitude, type, surface area etc. We obtained two years of these data – 2007 and 2013, and using information on opening and closing dates, we were able to extrapolate the number of establishments – number of stores, total retail surface area, average store surface area, number of hypermarkets, number of supermarkets and number of superettes (*no stores*, *total surface area, average surface area, HP, SP and SUP* in table 3), reasonably closely. We also defined an alternative category of stores – *special*, to distinguish stores that offer specific services/product line, e.g. hard discount stores or stores selling exclusively frozen food. For a more detailed description of this data source we refer the readers to Caillavet *et al.* 2015.

It should be noted that TDLinx provides information for only 'major' retail store types – hypermarkets (hypermarche – stores with surface area more than 2,500m^2), supermarkets (supermarche - stores with surface area between 400 and 2,500m^2) and grocery stores (superettes – stores with surface area less than 400m^2).

Other types of retailing: street markets, stands, convenience and specialty stores, as well as restaurants, bars and canteens are not included in the data. Hence, we supplement this information from another source – Système Informatique pour le Répertoire des ENtreprises et des Etablissements (SIRENE).

The second source of data on retail outlets come from INSEE. INSEE registers all businesses present on the French territory and maintains an updated database called SIRENE. It includes more than 6.5 million businesses and 7.5 million retail establishments. In addition, it provides information on the industry code, opening date, on the number of employees, and detailed addresses. So this dataset can be considered as an exhaustive source on retail trade. Concerning food outlets, it includes supermarkets and small retailers, as well as all types of restaurants: fast food or full service. We used data available as of July 2015, with a stock of food retailers in the 4 departments covered by our study. Sorting outlets according to their opening date, we could estimate the number of new stores which were created between 2005 and 2010, corresponding to the 2 waves of SIRS data. However, we were not able to capture the stores which have been closed during this period. We were also able to construct retail presence variables for both at and away from home (*FAH* and *FAFH* in table 3) as the total number of retail establishments for food at home and away from home, respectively. We also obtained the number of hypermarkets, supermarkets and superettes by simply identifying the corresponding industry codes. These variables correspond to *HM*, *SM* and *SUP* in table 3.

Finally, we use this data source to construct instruments for our instrumental variable estimations. Following Caillavet *et al.*, 2015, we construct the instruments as the percentages of openings in arrondissements. To confine the retail outlet openings, as a measure of attractiveness of the area for business, to the supply-side effects exogenous of demand-side influences, we obtained data from INSEE on *all* retail openings (with the exception of automobile and motorcycle dealerships), not only food store openings. Since the choice of location decision rests with the enterprise, but is ultimately approved (if the surface area is greater than 300 m^2) by the department zoning committees, the distribution of such openings across arrondissements, but within the same department, could be reasonably expected to be entirely exogenous to food habits of individual respondents. These instruments are listed under the corresponding heading in Table 3.

Empirical Model

The objective of this study is to estimate the effect of food retail environment, measured by the number of stores, retail surface area, or the type of retail outlet, on body mass. Let i index individual, in region g at time t. Then

$$BMI_{i_g t} = \alpha + \beta Retail_{g,t} + \gamma X_{i_g,t} + \nu_{i_g t} \qquad (1)$$

where $BMI_{i_g t}$ is the body mass of individual i in neighborhood g at time t; $Retail_{g,t}$ are different measures of the retail environment; and $X_{i_g,t}$ is a series of socio-demographic controls of the individuals; α, β and γ are parameters and $\nu_{i_g t}$ is the idiosyncratic error term.

If the retail measures are indeed uncorrelated with the idiosyncratic error term $\nu_{i_g t}$, the parameter β is an unbiased measure of the magnitude of the role, if any, of the retail environments on BMI. If, however, this is not the case then (1) can be expressed as

$$BMI_{i_g t} = \alpha + \beta Retail_{g,t} + \gamma X_{i_g,t} + \mu_i + \varepsilon_{i_g t} \qquad (2)$$

where μ_i is unobserved and $\varepsilon_{i_g t}$ has all the attractive characteristics we wish to have in the error term. In other words, if these unobserved preferences are also influencing the choice of the residence neighborhood or retail environment, then this would bias our estimates. Since this survey is based on two, even though nonconsecutive, periods, we wish to utilize an estimation method that can control for or entirely get rid of the individual time-invariant preferences.

One way to proceed is to use an estimation method that would get rid of the μ_i in (2), such as differencing. We utilize three estimation methods – pooled OLS, random effect (RE) and fixed effect (FE)[iii] methods to deal with this. We estimate one model for each retail environment measure and each reference area. Hence, using retail measures constructed from TDLinx ,we estimate Model (1) with the number of stores (*No stores*) in resident- and metro-centered areas, Model (2) with the total surface area (*Total Surface Area*) in resident- and metro-centered areas, Model (3) with the average surface area (*Avg Surface Area*) in resident- and metro-centered areas, Model (4) with store types (*HM, SM and SUP*) in resident- and metro-centered areas, and Model (5) with the number of special stores (*Special*) in resident- and metro-centered areas. The estimation results are in table 4. Similarly, using retail measures constructed from SIRENE, we estimate Model (1′) with the combined number of all food at home retail establishments (*FAH*), Model (2′) with both the number of all food at home and away from home retail establishments (*FAH, FAFH*), Model (3′) with different store types (*HM, SU and SUP*). These estimation results are in table 5.

Results

Based on the results in table 4, there does not appear to be a strong or systematic pattern of association between food environment measures and BMI, with the exception of superettes. For example, based on results from our fixed effects and random effects models, the number of retail outlets (hypermarkets, supermarkets and superettes) is systematically, albeit weakly, associated with BMI. Translated loosely, at the average BMI level, an increase of one retail outlet per one square kilometer is associated with a reduction of 0.02 and 0.03 BMI points in resident- and metro-centered reference areas. The combined retail area neighborhood availability or the average store size appear to have no effect or consequence on BMI.

Although the combined number of food retail outlets is not associated with BMI, there appears to be heterogeneity in the effects of the different types of retail outlets. Specifically, superettes, the small (less than 400m^2) grocery stores, have a persistent negative effect on body mass across all reference areas and estimation methods. Although the effects are larger than the effect of combined number of all retail stores discussed above, by a factor of 3 and more, they are still relatively small in magnitude (see RE and FE estimates in line 1 of table 4). Interestingly, having a discount or specialty store nearby has no effect on BMI either.

The results in table 5 offer qualitatively different results from table 4. While there seems to be symmetric concordance in the results in table 4 across reference areas (columns (1) to (3) versus columns (4) to (6)) and estimation methods, this is lacking in the results depicted in table 5. For example, *FAH* (which includes all *HM*, *SM* and *SUP*, like *No stores*, but unlike *No stores* it also includes all other types of stores – bakery, wine, confectionery, butchery, deli, fish, fruit and vegetable stores, etc.) is negatively associated with body mass in metro-centered reference areas, but not in resident-centered areas. Similarly, the number of food away from home outlets and *HMs* are negatively and positively associated with BMI, respectively.

Endogeneity

In estimating the effects of the retail environment on BMI, we would like to ascertain that the effects are in fact identifiable. Typically it is unrealistic to assume that the lifestyle choices and the choice of surrounding environments, such as food environment among others, are independent. As we mentioned in the Data section, we have a panel data which allows us to follow an individual through time. The random effects and fixed effects models discussed above would control for the time invariant unobservable factors. Secondly, none of the panel members in our data moved during the five year period between the two waves of the SIRS survey. Therefore, it could be argued that, under some assumptions, the survey respondents did not get to choose/change the environment they live in. In other words, the neighborhoods did not evolve or change during those 5 years, other than food availability. If so, then by simply including area fixed effects, we could assume away all the area specific differences and get to measure the only factor that is changing – food availability. The results of this approach were presented in tables 4 and 5 and discussed in the previous section.

It could, however, be argued that endogeneity could still be a threat due to time varying unobservable factors. We address this issue by performing instrumental variable estimation on some of our models. To identify the effects of food environment, we use supply-side factors following Caillavet *et al.* (2015).

As mentioned above, France has one of the most regulated retail industries. Starting 1973, a law called "Loi Royer" was introduced to regulate the opening of new or expanding existing retail outlets over $1000m^2$ in area (with the exception of hotels and restaurants). The regulation was done through zoning committees at the department (the medium size geographic units in France) level. So even though the zoning decisions are made at the department level, the firms choose the area within the department where to start the business. Therefore, we exploit this allocation variation within the departments as a source of exogenous variation. For this purpose, we obtained data on *all* retail openings (with the exception of car dealerships and such), and then expressed the openings in each arrondissement as a percentage of the whole (or department).

A potential pitfall of our instruments could be that after Loi Royer, a stricter regulation – Raffarin, was adopted by the country in 1996. This regulation required approval of openings of 300m^2 or larger stores and was the enforced law until 2008, which is about half way the gap in the two SIRS waves. This fact makes it difficult to treat zoning committee approvals or even applications before and after the Raffarin law was dropped. Consequently, we constructed the instruments in a way that exploits the cross-sectional variation rather than levels to safeguard against such external shocks that may obscure the true picture.

Following this logic, we constructed a number of instruments: percentage of all openings in the arrondissement (*dist*), percentage of openings of stores with 20 – 49 employees in the arrondissement (*E_20*), percentage of openings of stores with 50 – 99 employees in the arrondissement (*E_50*), and percentage of openings of stores with more than 100 employees in the arrondissement (*E_100*). The summary statistics of these instruments are presented in table 3.

Since some of our models have more than one food environment measurement variable, we checked the correlation between the instruments to make sure they can be used in the same models. The correlation matrix is presented in table 6. As can be seen, there is no strong correlation between these instruments and so they could be used in the same model, if necessary.

The results of the IV estimations are presented in tables 7 and 8. As can be seen from the first stage results, the instruments are strong. All the F statistics are sufficiently large and significant compared to Stock and Yogo critical values (Stock and Yogo 2005) and all the weak identification or orthogonality tests were rejected in the models.

The magnitudes of the estimates in tables 4 and 7 are comparable. *Avg surface area* appears to be significant at 10% level, indicating that larger size stores are associated with increased BMI in both reference areas, although the effect is very small. Interestingly, the IV parameter of HM came out to be very large and significant at 5%. In accordance with the results in table 4, the effect of Superettes is negative and statistically significant. The results in tables 5 and 8 are not consistent however with the model using TDLinx data.

Concluding Remarks

In this study, we set to estimate the effect of food retail environment on body mass in Paris, France. France is an interesting case to study since it is a country with a strong food culture. The "French Paradox" has also been of great interest to many consumers in the US and elsewhere. For example, it has been reported that the French people have low incidence of cardiovascular related diseases despite consuming high caloric and fatty foods. The bad news for the French, however, is that their obesity rate has been increasing in the last 15 years and this has become of great concern in France. So the question of interest to many is why are obesity rates increasing in France? Is the food environment partly to blame?

While Caillavet *et al.* (2015) has examined the effect of the food environment on fruits and vegetable consumption in France, there is scant information on the effect of food access and availability on obesity in France. In this study, we exploited a panel data set of approximately 1,250 respondents who reported their weight and height. We also used two additional data sources to construct our food environment measures to allow us to measure the effect of retail food availability on BMI. We used different datasets, different measures of food availability, different geographic units or reference areas, and different model estimations to test the robustness of our results.

Our results do not suggest any systematic or even strong influence of any of the food environment measures on BMI. Remarkably, the only consistent effects we detect are for superettes which have a negative effect and for hypermarkets which have a positive effect on BMI regardless of reference area or estimation method. The magnitude of these effects is relatively small, however. Hence, the results in this study should be taken only as suggestive. We basically cannot definitively say that the food environment in a French urban setting has an economically important effect on people's body mass in France. Therefore, given our findings, public intervention to change the food environment in view of countering the rise in obesity in France is, perhaps, misguided.

It is then possible that other factors are to blame for France's rising obesity rates. Our study, however, is focused only on one part of the country (i.e., Paris region) and so future studies should test the robustness of our findings using data from other regions or data for the whole country. Given that we did not find meaningful effects of the food environment on BMI, future studies should also attempt at examining other types of factors (e.g., behavioral factors) in an effort to decipher the reasons behind the recent surge in obesity in France.

References

Anderson, M.L., and D.A. Matsa. 2011. “Are Restaurants Supersizing America?” *American Economic Journal: Applied Economics* 3: 152-188.

Cadot E, J. Martin and P. Chauvin. 2011. Inégalités sociales et territoriales de santé: l’exemple de l’obésité dans la cohorte SIRS, agglomération parisienne, 2005. *Bulletin Epidémiologique Hebdomadaire* (8-9): 88-91.

Caillavet, F., G. Kyureghian, R.M. Nayga Jr., C. Ferrant and P. Chauvin. 2015. Does Healthy Food Access Matter in a French Urban Setting? The Role of Food Retail Structure. *American Journal of Agricultural Economics* 97(5): 1400-1416.

Casey R., B. Chaix, C. Weber, B. Schweitzer, H. Charreire, P. Salze, D. Badariotti, A. Banos *et al.* 2012. Spatial accessibility to physical activity facilities and to food outlets and overweight in French youth. *International Journal of Obesity* (36): 914-919.

Cawley J. 2015. An economy of scales: a selective review of obesity’s economic causes, consequences, and solutions. *Journal of Health Economics* http://dx.doi.org/10.1016/j.jhealeco.2015.03.001

Chaix, B., K. Bean, M. Daniel, S.N. Zek, Y. Kestens, H. Charreire, *et al.* 2012. ”Associations of Supermarket Characteristics with Weight Status and Body Fat: A Multilevel Analysis of Individuals within Supermarkets (RECORD Study). *Plos One* 7(4), April.

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

Currie, J., E. Moretti, S. Della Vigna, and V. Pathania. 2010.“The Effect of Fast Food Restaurants on Obesity and Weight Gain.” *American Economic Journal: Economic Policy* 2.3: 32-63.

Mis en forme : Anglais (États Unis)

Devaux M., Sassi F. 2011. Social inequalities in obesity and overweight in 11 OECD countries. *European Journal of Public Health* 23(3): 464-69.

Dunn, R.A. 2010. "The Effect of Fast-Food Availability on Obesity: An Analysis by Gender, Race, and Residential Location." *American Journal of Agricultural Economics* 92(4): 1149-1164.

Emery, C., J. Dinet, A. Lafuma, C. Sermet, B. Khoshnood, and F. Fagnani. 2007. Evaluation of the cost associated with obesity in France. *Presse Médicale* 6 juin 2007, p.832-840.

Hausman, J., and E. Leibtag. 2007. "Consumer Benefits from Increased Competition in Shopping Outlets: Measuring the Effect of Wal-Mart." *Journal of Applied Econometrics* 22(7): 1157-1177.

Michimi, A., and M.C. Wimberly. 2010. "Association of Supermarket Accessibility with Obesity and Fruit and Vegetable Consumption in the Conterminous United States." *International Journal of Health Geographics* 9:49.

ObEpi 2012. Enquête épidémiologique nationale sur le surpoids et l'obésité. http://www.roche.fr/content/dam/corporate/roche_fr/doc/obepi_2012.pdf

Organization for Economic Co-operation and Development (OECD), Obesity and the Economics of Prevention: Fit not Fat- France Key Facts. Accessed at http://www.oecd.org/els/health-systems/obesityandtheeconomicsofpreventionfitnotfat-francekeyfacts.htm on November 13, 2015.

Stock, J.H., and Yogo, M. 2005. "Testing for Weak Instruments in Linear IV Regression." In D.W.K. Andrews and J.H. Stock, eds. Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg. Cambridge: Cambridge University Press: 80-108.

Thomsen M.R., Kyureghian G., and Nayga Jr. R.M. Food Marketing. In: Neal Van Alfen, editor-in-chief. Encyclopedia of Agriculture and Food Systems, Vol. 3, San Diego: Elsevier; 2014. pp. 196-212.

Volpe, R., A. Okrent, and E. Leibtag. 2013. "The Effect of Supercenter-format Stores on the Healthfulness of Consumers' Grocery Purchses." *American Journal of Agricultural Economics* 95(3): 568-589.

Table 1. Overweight and Obesity by Sex at the National Level, France

Years	Overweight Men	Overweight Women	Obesity Men	Obesity Women
1997	45.7	31.6	8.8	8.3
2000	48.6	33.5	10.3	10.0
2003	51.1	36.4	12.0	11.9
2006	50.0	37.8	12.5	13.6
2009	52.4	41.1	13.9	15.1
2012	53.1	42.0	14.3	15.7
97-2012	+16.2	+32.9	+62.5	+89.2

Source: ObEpi 2012, adults >= 18years

Table 2. Overweight and Obesity Rates at the National Level and in the Parisian Region

Years	Overweight national	Obesity rate national	Rate of increase	Obesity rate Parisian region	Rate of increase
1997	38.3	8.5		7.8	
2000	40.7	10.1	+18.8	8.9	+14.1
2003	43.4	11.9	+17.8	11.9	+33.7
2006	43.7	13.1	+10.1	12.1	+1.7
2009	46.4	14.5	+10.7	13.2	+9.0
2012	47.3	15.0	+3.4	14.4	+9.1
97-2012	+23.5	+76.5		+84.6%	

Source: ObEpi 2012, adults >= 18years

Table 3. Variable Descriptions and Summary Statistics.

Variable Name	Variable Description	Mean (Std Dev)
Dependent Variable		
BMI	Body Mass Index = kg/m^2	24.8103 (4.5523)
Demographic Variables		
HHDSize	The number of the household members	2.2412 (1.4699)
Gender	Binary variable = 1 if respondent is male; = 0 otherwise	0.3870 (0.4872)
Education	Binary variable = 1 if respondent has college or higher education; = 0 if high school or less	0.5381 (0.4986)
Employed	Binary variable = 1 if respondent employed full- or part-time; = 0 otherwise	0.5585 (0.4967)
Income	Monthly household income, in Euros	2822.17 (2699.73)
Age	Age of respondent in years	51.77 (15.91)
Instruments		
Dist	The proportional distribution (%) of *all* retail openings in the department in each arrondissement	27.5800 (15.5431)
E_20	Number of openings of retail enterprises with the number of employees from 20 to 49	30.3776 (26.0699)
E_50	Number of openings of retail enterprises with the number of employees from 50 to 99	19.3877 (26.8270)
E_100	Number of openings of retail enterprises with the number of employees more than 100	14.7436 (29.4739)

Table 3. Continued

Variable Name	Variable Description	Mean (Std Dev) Circle around Residence	Mean (Std Dev) Circle around Metro
*Food Availability Measures**			
TDLinx			
No of Stores	Number of all food stores in the circle around residence or metro	5.3435 (7.1390)	4.4219 (5.6964)
HM	Number of Hypermarket stores in the circle around residence or metro	0.2152 (0.4568)	0.1741 (0.3995)
SM	Number of Supermarket stores in the circle around residence or metro	2.3060 (3.2035)	1.9659 (2.6043)
SUP	Number of Superettes in the circle around residence or metro	2.8223 (4.5808)	2.2819 (3.4268)
Special	Number of stores specializing in either the foods or price	1.6032 (2.5980)	1.4880 (2.6742)
Total Surface Area	Combined area (in m^2) devoted to sales in all food retail stores in the circle around residence or metro	4786.83 (5911.10)	3445.47 (4014.51)
Avg Surface Area	Average area (in m^2) devoted to sales in all food retail stores in the circle around residence or metro	1253.26 (1095.74)	1166.37 (1365.56)
SIRENE			
FAH	Number of all grocery foods combined – not only HM, SM, SUP	19.0959 (26.3020)	26.8082 (36.8590)
FAFH	Number of all food away from home outlets	34.5667 (63.6339)	45.4305 (74.2677)
HM	Number of Hypermarket stores in the circle around residence or metro	0.0477 (0.1708)	0.1155 (0.3898)
SM	Number of Supermarket stores in the circle around residence or metro	0.9032 (1.5761)	1.2337 (2.7067)
SUP	Number of Superettes in the circle around residence or metro	0.5495 (1.3061)	1.2099 (2.1184)

* All availability measures are normalized to be per square kilometer by dividing by area of the reference area/circle.

Table 4. Estimation results using TDLinx data

Food Access Variable	Resident-Centered Area Pooled OLS (1)	Resident-Centered Area Random Effect (2)	Resident-Centered Area Fixed Effect (3)	Metro-Station-Centered Area Pooled OLS (4)	Metro-Station-Centered Area Random Effect (5)	Metro-Station-Centered Area Fixed Effect (6)
Model (1)						
No Stores	-0.0009	-0.0008*	-0.0007*	-0.0001	-0.0012*	-0.0013*
	(0.0009)	(0.0004)	(0.0004)	(0.0013)	(0.0007)	(0.0007)
Model (2)						
Total Surface Area	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Model (3)						
Avg Surface Area	0.0000	0.0000	0.0000	0.0000*	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Model (4)						
Hypermarket	-0.0052	0.0021	0.0031	0.0086	0.0065	0.0068
	(0.0178)	(0.0111)	(0.0111)	(0.0214)	(0.0109)	(0.0109)
Supermarket	0.0027	0.0012	0.0011	0.0045	0.0017	0.0013
	(0.0023)	(0.0013)	(0.0014)	(0.0032)	(0.0015)	(0.0016)
Superette	-0.0025*	-0.0013***	-0.0012**	-0.0032	-0.0024***	-0.0023***
	(0.0013)	(0.0005)	(0.0005)	(0.0021)	(0.0008)	(0.0008)
Model (5)						
Special	-0.0016	-0.0021*	-0.0020	0.0018	-0.0003	-0.0014
	(0.0026)	(0.0013)	(0.0013)	(0.0021)	(0.0013)	(0.0019)

Note: Data source – Trade Dimensions on TDLinx by the Nielsen Co. The dependent variable in each model is the natural log of BMI. Robust standard errors in parentheses. All estimations were clustered at the individual level. The sample size is 2,496, the number of clusters is 1,248. All estimations include a full array of demographic variables area fixed effect, except for the FE method. Variable definitions and summary statistics are available in the Data section. *, **, and *** signify significance at 10%, 5% and 1% levels.

Table 5. Estimation results using SIRENE data

Food Access Variable	Resident-Centered Area			Metro-Station-Centered Area		
	Pooled OLS (1)	Random Effect (2)	Fixed Effect (3)	Pooled OLS (4)	Random Effect (5)	Fixed Effect (6)
Model (1′)						
FAH	0.0002	-0.0002	-0.0003	0.0000	-0.0002*	-0.0003**
	(0.0004)	(0.0002)	(0.0002)	(0.0003)	(0.0001)	(0.0001)
Model (2′)						
FAH	0.0006	0.0006	0.0005	0.0002	-0.0001	-0.0002
	(0.0005)	(0.0004)	(0.0005)	(0.0003)	(0.0002)	(0.0003)
FAFH	-0.0003	-0.0003**	-0.0003	-0.0002	-0.0001	0.0000
	(0.0002)	(0.0001)	(0.0002)	(0.0002)	(0.0001)	(0.0001)
Model (3′)						
Hypermarket	0.1425	-0.0593	-0.0784	0.0030	0.0099**	0.0112
	(0.2875)	(0.1545)	(0.1498)	(0.0099)	(0.0049)	(0.0049)
Supermarket	-0.0013	-0.0019	-0.0018	0.0026	0.0025	0.0033
	(0.0036)	(0.0017)	(0.0018)	(0.0025)	(0.0018)	(0.0030)
Superette	-0.0102	-0.0009	0.0022	-0.0062	-0.0048	-0.0043
	(0.0086)	(0.0046)	(0.0050)	(0.0049)	(0.0020)	(0.0021)

Note: Data source – Institut National de la Statistique et des Etudes Economiques or INSEE. The dependent variable in each model is the natural log of BMI. Robust standard errors in parentheses. All estimations were clustered at the individual level. The sample size is 2,496, the number of clusters is 1,248. All estimations include a full array of demographic variables area fixed effect, except for the FE method. Variable definitions and summary statistics are available in the Data section. *, **, and *** signify significance at 10%, 5% and 1% levels.

Table 6. Pearson Correlation Coefficients between Instrumental Variables

	Dist	E_20	E_50	E_100
Dist	1.0000	0.4669	0.1056	0.4165
		(0.0439)	(0.6672)	(0.0721)
E_20			0.1398	0.4114
			(0.5682)	(0.0801)
E_50				0.4719
				(0.0413)
E_100				1.0000

Note: p-values in parentheses. N = 19.

Table 7. Instrumental Variable Estimations with Food Access Data from TDLinx.

Food Access Variable	Resident-Centered Area			Metro-Station-Centered Area		
	1st Stage		2nd Stage	1st Stage		2nd Stage
	Instrument	F-stat (p-value)	Parameter Estimate (St.Err)	Instrument	F-stat (p-value)	Parameter Estimate (St.Err)
Model (1)						
No Stores	E_20	67.33 (0.0000)	0.0013 (0.0025)	E_20	144.63 (0.0000)	0.0015 (0.0028)
Model (2)						
Total Surface Area	E_20	48.33 (0.0000)	0.0000 (0.0000)	E_20	40.61 (0.0000)	0.0000 (0.0028)
Model (3)						
Avg Surface Area	E_50	41.06 (0.0000)	0.0000* (0.0000)	E_50	95.18 (0.0000)	0.0000* (0.0000)
Model (4)						
Hypermarket	Distr	24.55 (0.0000)	0.1526** (0.0767)	Distr	21.55 (0.0000)	0.0917 (0.0570)
Supermarket	E_20	16.55 (0.0000)	0.0316 (0.0299)	E_20	44.54 (0.0000)	0.0087 (0.0062)
Superette	E_50	110.45 (0.0000)	-0.0162 (0.0101)	E_50	48.93 (0.0000)	-0.0061* (0.0033)

Note: Data source – Trade Dimensions on TDLinx by the Nielsen Co. Robust standard errors. All estimations were clustered at the individual level. The sample size is 2,496, the number of clusters is 1,248. Variable definitions and summary statistics are available in the Data section. *, **, and *** signify significance at 10%, 5% and 1% levels.

Table 8. Instrumental Variable Estimations with Food Access Data from SIRENE

Food Access Variable	Resident-Centered Area			Metro-Station-Centered Area		
	1st Stage		2nd Stage	1st Stage		2nd Stage
	Instrument	F-stat (p-value)	Parameter Estimate (St.Err)	Instrument	F-stat (p-value)	Parameter Estimate (St.Err)
Model (1)						
FAH	E_20	90.22	0.0006	E_20	132.30	0.0003
		(0.0000)	(0.0011)		(0.0000)	(0.0007)
Model (2)						
FAH	Distr	36.96	0.0578	Distr	66.55	-0.0033
		(0.0000)	(0.0631)		(0.0000)	(0.0022)
FAFH	E_20	38.62	-0.0214	E_20	42.63	0.0018*
		(0.0000)	(0.0238)		(0.0000)	(0.0010)
Model (3)						
Hypermarket	Distr	24.42	-1.3305	Distr	29.89	0.0358
		(0.0000)	(1.0076)		(0.0000)	(0.0315)
Supermarket	E_20	28.02	0.0126	E_20	23.05	0.0105
		(0.0000)	(0.0226)		(0.0000)	(0.0201)
Superette	E_100	69.82	-0.0247	E_100	41.84	0.0099
		(0.0000)	(0.0252)		(0.0000)	(0.0132)

Note: Data source – Institut National de la Statistique et des Etudes Economiques or INSEE. Robust standard errors in parentheses. All estimations were clustered at the individual level. The sample size is 2,496, the number of clusters is 1,248. Variable definitions and summary statistics are available in the Data section. Underidentification and instrument orthogonality conditions were met at all first stage estimations (statistics are not reported here, but are available upon request). *, **, and *** signify significance at 10%, 5% and 1% levels.

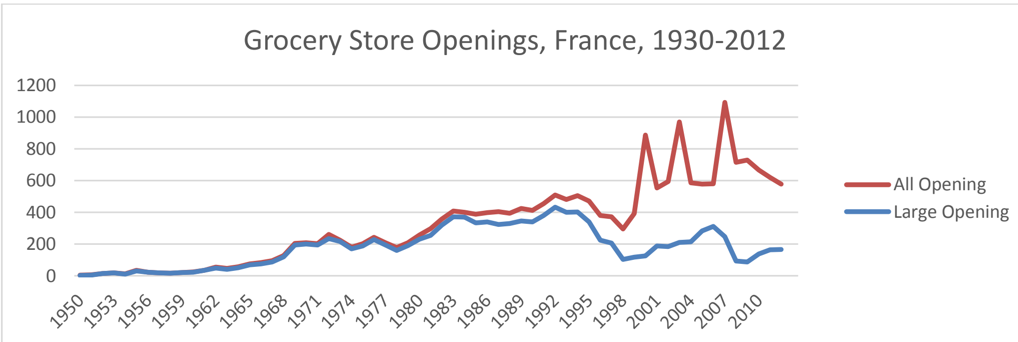

Figure 1. Grocery store openings since 1950’s in France.

Note: Data source for the openings is Trade Dimensions. The openings from 2007 to 2013 are estimates obtained by dividing the difference of the numbers of stores in the two periods by the number of the years, and adjusting for the closings in each year.

Endnotes

[i] The data about the store numbers were obtained from Trade Dimension from 2007 and 2013; the population numbers in France for the two periods were obtained from INSEE.

[ii] IGF/Igas report for the Senate, http://www.senat.fr/rap/r13-399/r13-3993.html

[iii] Conveniently, the first differencing and fixed effect methods yield identical results in the case of two periods.