Essay n°2 : Research Proposal for *Food Deserts and the Causes of Nutritional Inequality*, by Allcott et al. (2019)

Romain FERNEX

May 2026

1 Major Critiques

In the previous essay, I reviewed the main findings of the paper as well as the main assumptions the authors rely on. Amongst these modelling assumptions, three may be potentially problematic and are not addressed by the authors: the non-homotheticity of demand, the partial equilibrium assumption, and the choice of making the model static rather than dynamic. I shall expand upon the non-homotheticity critique in my proposition for an extension at the end of this paper.

The first major limitation of this model is that it does not account for potential general equilibrium responses on the retailers side or for spatial equilibrium effects, both of which might bias the policy estimates. The structural counterfactual, where the authors test for the effect of a food subsidy, holds store locations, product assortments and retailer prices fixed throughout the simulation. In a general equilibrium setting, a subsidy of this magnitude would also induce medium and long-run responses from retailers, who might expand their offer of healthy products in low-income markets, and from manufacturers who might adjust their pricing as the effective purchasing power of their customer base increases. This directly influences both how beneficial to the health of low-income households the policy might be, and how much a government would need to invest to sustain the nutritional health effects described by the author.

The bias is further amplified when one considers spatial equilibrium effects for a federal program such as SNAP. Moretti (2011) shows that local demand shocks do not stay local: they propagate across areas through population movements and factor price adjustments, so that a shock in

one location feeds back into wages and rents in connected markets.[1] Translated to the retail food setting, this suggests that even a local demand shift in favour of healthy products could propagate across markets and influence retail entry decisions and food prices in ways this model cannot capture. The endogenous spatial reallocation of retail activity might thus further bias, upward or downward, the estimated cost-effectiveness of such a program.

Another salient limitation of the proposed structural model is that it cannot efficiently distinguish deep-seated tastes from entrenched habits, due to its absence of a dynamic component. Past consumption does not enter current utility and structural preference parameter estimates do not evolve over time. If dietary habits are partly formed by cumulative exposure to a given food environment, as some studies suggest they are for children (Wadhera et al., 2015), a sustained change in this environment might have significant long-run effects on future preferences and consumption behaviour. This is analogous in spirit to De la Roca and Puga (2017), who show that city effects on individual productivity accumulate dynamically: a worker who has lived in a dense city for ten or more years has permanently higher productivity than one who arrives later in life, because learning from the urban environment compounds over time. The principle, though not the mechanism, is the same here: if exposure to a constrained food environment shapes preferences gradually, then the static model's estimated preference parameters will partly reflect the history of that exposure rather than intrinsic tastes. An interesting robustness check the authors could run to assess the magnitude of these dynamic effects would be to estimate preference parameters separately for long-tenure residents (same ZIP code for ten or more years) and recent movers, at fixed income levels. Systematic differences between the two groups would indicate that habit formation is contaminating the static preference estimates, although one should acknowledge that long-tenure residents in food deserts are likely to differ from recent movers along unobserved dimensions beyond habit formation, which complicates the interpretation.

Lastly, regarding homotheticity, unlike Handbury (2021), the Cobb-Douglas functional form the authors rely on implies that expenditure shares on each product group are invariant to total grocery expenditure at fixed relative implicit prices. This means that all income-related variation

[1] Specifically, Moretti (2011) models a two-region economy and derives that a productivity shock in one city raises nominal wages there, attracting workers from the other city, which in turn raises wages and housing costs in the sending city as well. The key result is that partial-equilibrium estimates of local shocks overstate true effects because they ignore these cross-market feedbacks.

in dietary composition is channelled entirely into differences in the estimated preference parameters across income quartiles. Consider two households with identical underlying preferences for whole grains but different incomes. The low-income household is more price-sensitive at the margin: when the implicit price of whole grains rises slightly, because a cheaper whole grain product disappears from the local store, it substitutes away from whole grains more strongly than the high-income household would. Under a flexible demand system this differential price sensitivity would show up as a difference in the price elasticity of the expenditure share, not as a difference in preferences. Under Cobb-Douglas, however, there is no room for this as both households have the same elasticity of substitution across product groups regardless of income. Thus, the observed difference in their dietary composition is entirely attributed to differences in preference parameters across income groups. The model therefore cannot distinguish between a household that genuinely values whole grains less and a household that values them equally but responds more strongly to price signals because it has less slack in its budget. It is important to note that this is fundamentally different from the endogeneity issue that the authors correct for using their price instrument: the instrument addresses the endogeneity of prices, but it does not relax the structural constraint that all income groups face the same elasticity of substitution. This is a direct callback to what Handbury (2021) observes with a non-homothetic CES framework, as they find that the elasticity of the food price index with respect to city income differs substantially across household income levels, with low-income households being more price-sensitive for quality-differentiated goods.

2 Minor Critiques

Due to data limitations, the authors do not extend their study to food away from home (restaurants, fast food chains, and similar outlets). This is problematic for three different reasons. First, the share of calorie consumption drawn from food away from home may differ based on household income or on whether a household is located in a food desert, and it is likely that low-income households with limited access to supermarkets turn to fast food chains more than their counterparts in better-supplied areas. Second, a 2018 USDA report (Saksena et al., 2018) has shown

that, irrespective of income, the share of calorie consumption drawn from food away from home has grown steadily over the past decade, which makes omitting this channel a genuine threat to their narrative. Third, the same report highlights significant discrepancies across income groups in the dietary quality of food obtained from away-from-home outlets: households who frequent full-service restaurants tend to be of higher socioeconomic status, and full-service restaurants are associated with substantially better diet quality than fast food outlets. To assess whether this is biasing their estimates, Allcott et al. could use FoodAPS data from the USDA, which covers both grocery and away-from-home food acquisition with ZIP-code level geographic identifiers, to estimate a descriptive decomposition of the income-nutrition gradient by channel and test whether the away-from-home gradient correlates with the same supply-side variables the paper uses. If the away-from-home nutritional gap is wider than the grocery gap, and correlated with the same supply-side variables, the paper's 90% demand attribution and its proposed subsidy cost estimate would both require substantial revision.

Lastly, the authors note that, in the Homescan data, even households in ZIP codes with no supermarkets still buy 85% of their groceries from supermarkets, and suggest this is consistent with the idea that physical access is not a major constraint. This is a plausible interpretation at the level of the full Homescan sample, but the statistic is not very informative about the specific population for whom access is most likely to bind. The figure is computed over a panel that heavily represents suburban, car-owning households. For such households, a supermarket located several miles away is both recorded as available in the ZIP code and practically accessible. For low-income, car-free households in genuine food deserts, by contrast, the distance, time and travel cost of reaching a supermarket are much more likely to be binding, and their share of groceries purchased at supermarkets is almost certainly lower than 85%, although this is not shown. In other words, the same aggregate number can be consistent with non-binding access constraints for the average Homescan household, yet still mask meaningful constraints for the small subset of households most relevant for food-desert policy. A simple descriptive extension reporting the supermarket share separately for food-desert households, and within that group for car-owners versus non-car-owners, would clarify whether this interpretation carries over to the policy-relevant

population using data that are already available within the Homescan panel.

3 Extension

As mentioned in my first essay, the paper concludes that 90% of the nutritional gap reflects demand-side factors, which it attributes to preference heterogeneity: low-income households place a lower value on the nutritional characteristics that define a healthy diet. This conclusion rests on the assumption, embedded in equation (7), that the sensitivity to the implicit price (the nominal price per calorie adjusted for the utility value of its nutritional characteristics) is the same across all product groups within an income group.

The model therefore cannot capture whether households are more sensitive to a given increase in this implicit price for some product groups than for others. Consider a low-income household facing a tight budget. When the implicit price of fresh produce rises, because the product group is expensive relative to the calories it provides, that household may cut back more sharply than a richer household would, not because it values produce less, but because it has less room to absorb cost increases. This distinction matters directly for the cost of the paper's proposed SNAP subsidy, since the subsidy works by reducing the implicit price of nutritionally valuable foods for low-income households. If those households are in fact more sensitive to changes in the implicit price for high-score product groups, each dollar of subsidy produces a larger dietary improvement than the preference-based model predicts, and the $11 billion cost estimate likely overstates the true cost. In this extension I propose a revised version of their model which could enable the authors to quantify this heterogeneity in implicit price elasticities at the product group level.

Following the paper's own Health Index construction, I define the nutritional score of a product group j as $s_j = \sum_c \omega_c \bar{a}_{jc}$, where $\bar{a}_{jc}$ is the sample-average amount of characteristic c per calorie in group j and ω_c are the linearised HEI weights from Online Appendix Table A1. Groups in the upper half of this distribution form the high-score set $\mathcal{H}$ (typically vegetables and fruits), while groups in the lower half constitute the low-score set $\mathcal{L}$ (soft drinks, biscuits). This classification uses no external data and applies the paper's own weights to the paper's own average-

characteristic data.

3.1 New Demand Equation

Following the paper's notation, the baseline demand equation for income group g is (eq. 7):

$$\ln Y_{igjt} = \alpha_{gj} + \mu_m + \tau_t - \lambda_g\Big(p_{ijt} - \sum_c \tilde{\beta}_{gc}\,\tilde{a}_{ijct}\Big) + \varepsilon_{igjt}, \tag{1}$$

where λ_g is the slope on the implicit price for income group g (unnamed in the paper). I modify this demand equation in the following way:

$$\ln Y_{igjt} = \alpha_{gj} + \mu_m + \tau_t - \Big(\lambda_g + \rho_g \cdot \mathbf{1}[j \in \mathcal{H}]\Big)\Big(p_{ijt} - \sum_c \tilde{\beta}_{gc}\,\tilde{a}_{ijct}\Big) + \varepsilon_{igjt}, \tag{2}$$

where ρ_g captures the additional implicit-price sensitivity of income group g for high-score product groups relative to low-score ones.[2] The modification adds exactly four parameters $\{\rho_g\}_{g=1}^4$ to the specification.

3.2 Identification Strategy

The paper's chain instrument generates price variation across product groups but does not distinguish whether this variation falls on high- or low-score groups. Identifying ρ_g therefore requires augmenting the instrument with chain-by-score-group interactions: the relevant variation is whether a chain has a comparative pricing advantage specifically within $\mathcal{H}$ versus $\mathcal{L}$ products, conditional on its overall pricing. This variation can be obtained directly from the RMS data, provided the chain-level pricing advantage varies sufficiently between high- and low-score groups, which can be verified directly before estimation.

3.3 Interpretation

If $\hat{\rho}_g$ is negative and larger in absolute value for low-income quartiles than for high-income ones, then low-income households are more sensitive to changes in the implicit price of nutritionally valuable foods, holding preferences fixed. Part of what the paper labels as a preference gap would then be better understood as a price-sensitivity gap: low-income households are not intrinsically

[2] All other parameters retain their original interpretation.

less interested in healthy food, but respond more sharply to its implicit cost. In that case, the SNAP healthy-food subsidy would produce a larger dietary improvement per dollar spent than the preference-based model predicts, and the $11 billion cost estimate would overstate the true cost. Conversely, if $\hat{\rho}_g$ does not vary meaningfully across income groups, the preference interpretation stands and the cost estimate is validated. Either outcome sharpens the paper's decomposition and directly informs the calibration of its main policy recommendation.

References

De la Roca, Jorge and Diego Puga (2017). "Learning by Working in Big Cities". In: *Review of Economic Studies* 84.1, pp. 106–142.

Handbury, Jessie (2021). "Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across US Cities". In: *Econometrica* 89.6, pp. 1679–2715.

Moretti, Enrico (2011). "Local Labor Markets". In: *Handbook of Labor Economics*. Ed. by Orley Ashenfelter and David Card. Vol. 4B. Amsterdam: North-Holland, pp. 1237–1313.

Saksena, Michelle J. et al. (2018). *America's Eating Habits: Food Away From Home*. Economic Information Bulletin 187. U.S. Department of Agriculture, Economic Research Service.

Wadhera, Devina et al. (June 2015). "Perceived recollection of frequent exposure to foods in childhood is associated with adulthood liking". In: *Appetite* 89, pp. 22–32. DOI: 10.1016/j.appet.2015.01.011.