Audit Report: Workflow `d1617aa16048423daa22ca7e29c3526c`

Overview

- Source JSON: `reports\workflow-report-d1617aa16048423daa22ca7e29c3526c.json`
- Run ID: `d1617aa16048423daa22ca7e29c3526c`
- Papers input: `papers\alcott food deserts.pdf`
- Started at: `2026-02-15T19:13:11.001652+00:00`
- Finished at: `2026-02-15T19:27:16.173575+00:00`
- Duration: `0:14:05.171923`

Effective Configuration

- Chat model: `gpt-5`
- Embedding model: `text-embedding-3-large`
- Top K: `10`
- Chunk words / overlap: `350 / 75`
- Batch size: `64`
- Database URL configured: `True`

Step Outcomes

- `prep: completed`
- `ingest: num_pdfs=1, num_papers=1`
- `enrich: openalex=0, citec=0`
- `econ_data: fetched`
- `agentic: completed`
- `index: skipped (reason: db_unreachable)`
- `report_store: skipped (reason: db_unreachable)`

Agentic Summary

- Status: `completed`
- Main question: What is the key contribution?
- Report question set: `both`
- Structured questions generated: `84`
- Confidence mean/median: `0.3041450240118245 / 0.2994701740878031`
- Confidence labels: low=65, medium=18, high=0

Final Answer

- Quantifies the drivers of the nutrition–income gradient, showing that roughly 90% is due to demand (preferences) and about 10% to supply (prices/availability), across robust specifications; assigns the unobserved term $\hat{\xi}$ to demand because it primarily governs demand elasticity with respect to price [page 41, words 275–411; page 38, words 275–377].

- Identifies supply-driven price variation using substantial geographic differences in the footprint of top retailers that set distinct relative prices across product groups, with market fixed effects to address endogeneity; validates the approach via a replicated "migration" effect for Coke vs. Pepsi (^ 0.16: a 10 pp higher county Coke share raises a household's Coke share by 1.6 pp) [page 32, words 275–431; page 27, words 275–482].

- Estimates nutrient-level preferences by income quartile and introduces a WTP for Health Index, $\sum_c \hat{\beta}_c s_c r_c$, that increases monotonically with income (Q1 0.429 → Q4 1.141), with higher-income households exhibiting more negative WTP for sodium/added sugar and more positive WTP for whole fruit/whole grains [page 36, words 0–349].

- Grounds the analysis in rich linked scanner data (Homescan + RMS), arguing packaged-goods focus reasonably represents produce purchasing in the magnet subsample and documenting broad store/channel coverage with national weighting and CPI-U deflation [page 8, words 0–349; page 8, words 275–393].

Sub-Answers

Sub-answer 1

- Question: ResponseTextConfig(format=ResponseFormatText(type='text'), verbosity='medium')

- Question tokens estimate: `19`

- Answer:

- Data sources and coverage

    - Homescan records packaged groceries; in 2004–2006, an 8,000-household "magnet" subsample also recorded nonpackaged items. About 60% of magnet households' produce calories are from packaged goods, with no statistical income differences, suggesting packaged focus is reasonably representative for produce and that produce is a small share of groceries overall (page 8, words 0–349).
    - Households report income (16 bins), composition, demographics; researchers compute daily calorie needs using Dietary Guidelines and scale children to adult equivalents. PanelViews add-on (2017) provides health importance, nutrition quiz, BMI, diabetes; results are nationally weighted (page 8, words 0–349).
    - RMS: weekly UPC-level prices/volumes at ~42,000 stores from 160 chains (35,000/year observed) for 2006–2016; excludes liquor and bulk produce. Channel coverage: 53% grocery, 32% mass merchandiser, 55% drug, 2% convenience. Prices/incomes deflated to 2010 dollars using CPI-U (page 8, words 275–393).

- Identification and market structure facts

- – Substantial geographic variation in the presence of the five largest retailers; these chains set different relative prices across product groups, yielding geographic price instruments. Example: produce predicted cheap on West Coast and expensive on East Coast, likely due to California production; market fixed effects address endogeneity from demand elasticity differences across markets (page 32, words 275–431).
  - – Within-household regressions: adding demographic controls changes results little and increases R^2 only slightly, implying possible unobserved within-household changes (Oster 2019). A migration-style thought experiment considers moving a household from low- to high-income retail environments (partial equilibrium) (page 27, words 275–482).
  - – Validation: replicates immediate brand choice migration effect for Coke vs. Pepsi with $\hat{}$ 0.16; a 10 pp higher county Coke share raises a household's Coke share by 1.6 pp (page 27, words 275–482).

- • Preference estimates (Table IV) and WTP for health

  - – GMM estimates of nutrient preference parameters by income quartile show:
    - ∗ Higher-income quartiles have more negative WTP for sodium and added sugar, and more positive WTP for whole fruit and whole grains; patterns are monotonic across quartiles (page 36, words 0–349).
    - ∗ "WTP for Health Index" rises with income: Q1 0.429, Q2 0.631, Q3 0.820, Q4 1.141 (SEs shown in table) (page 36, words 0–349).
  - – Units: sodium (g), whole/other fruit and dairy (cups), whole/refined grains and proteins (oz), added sugar (tsp), solid fats (calories). WTP for Health Index defined as $\sum_c \hat{\beta}_c s_c r_c$ (page 36, words 0–349).

- • Decomposition: supply vs. demand in nutrition–income gaps

  - – Across alternative specifications (including census-region-by–product group fixed effects and interactions with urban/rural and above-median income), supply explains 7%–12% and demand explains 88%–93% of the nutrition–income relationship; summary: about 90% demand, 10% supply (page 41, words 275–411; page 38, words 275–377).
  - – The unobserved characteristic $\hat{\xi}$ is attributed to demand because it primarily determines demand elasticity with respect to price (page 41, words 275–411).

- • Demographic controls used in outcomes regressions

  - – Controls include household size, age indicators, log years of education, race, marital status, employment, weekly work hours, plus add-

on variables such as the self-reported importance of staying healthy (page 42, words 275–356; page 8, words 0–349).

Structured Q&A Appendix

This section mirrors `agentic.report_questions` for audit traceability.

A01: What is the main research question of the paper?

- Category: `A) Research question / contribution`
- Confidence: `low` (0.33061339623330005)
- Retrieval method: `local`
- Evidence type: `direct textual evidence`
- Data source: Allcott et al., Quarterly Journal of Economics (2020)
- Assumption flag: `False`
- Answer:

The paper asks what causes nutritional inequality—why higher-income Americans eat more healthfully than lower-income Americans—and examines whether neighborhood food environments (supply) versus household demand explain this gap.

- Quote snippet:

  We study the causes of "nutritional inequality": why the wealthy eat more healthfully than the poor in the United States.

- Citation anchors:
  - page=1, words=35-70, section=Introduction, note=States the study's focus on causes of nutritional inequality and why the wealthy eat more healthfully than the poor.

A02: What is the paper's primary contribution relative to the existing literature?

- Category: `A) Research question / contribution`
- Confidence: `low` (0.3459851165680258)
- Retrieval method: `local`
- Evidence type: `direct`
- Data source: QJE 2020 article text (Allcott et al.)
- Assumption flag: `False`
- Answer:

The paper's primary contribution is to overturn the food-desert explanation for nutritional inequality by showing—using supermarket entry and migration event studies plus a structural demand model with a novel cross-market price instrument—that equalizing access to products and prices would reduce the gap by only about 10%, with roughly 90% driven by demand differences, implying limited impact of supply-side policies.

- Quote snippet:

  reduces nutritional inequality by only about 10%, while the remaining 90% is driven by differences in demand.

- Citation anchors:
  - page=2, words=0-349, section=I. INTRODUCTION, note=States 10% reduction from equalizing supply; 90% due to demand; counters supply-side policy arguments.
  - page=1, words=0-349, section=Abstract/Overview, note=Rejects meaningful neighborhood (food-desert) effects; introduces structural model and new instrument.
  - page=34, words=0-349, section=VI.C / Instrument validity, note=Describes the price instrument and notes it is novel in the literature.
  - page=42, words=0-349, section=VII.B, note=Reiterates that findings counter food-desert arguments; supply equalization cuts gap by ~10%.

A03: What is the central hypothesis being tested?

- Category: `A) Research question / contribution`
- Confidence: `low (0.29062503471444984)`
- Retrieval method: `local`
- Evidence type: `Within-household event study of movers (fixed-effects), complemented by structural demand model counterfactuals`
- Data source: Nielsen Homescan panel and RMS retail data
- Assumption flag: `False`
- Answer:

Whether local place/supply conditions—the neighborhood retail food environment ("food deserts")—causally change households' healthy grocery purchasing enough to explain the nutrition–income gap.

- Quote snippet:

  To test for place effects, we measure within-household changes in grocery purchases after the 20,031 cross–ZIP code moves and the 11,728 cross-county moves...

- Citation anchors:
  - page=23, words=275-408, section=Event study of place effects, note=Defines the test for place effects using within-household changes after moves.
  - page=4, words=275-423, section=Summary of move-based findings, note=Reports small convergence after moves and bounds place effects at 2%–3%.
  - page=42, words=0-349, section=Interpretation relative to food deserts, note=States the findings counter arguments that food

deserts drive nutritional inequality.

– page=48, words=0-323, section=Conclusion on place effects, note=Concludes place effects explain no more than ~3% of differences in healthy eating.

A04: What are the main outcomes of interest (dependent variables)?

- Category: `A) Research question / contribution`
- Confidence: `low (0.3047568666778385)`
- Retrieval method: `local`
- Evidence type: `Table and text from the paper`
- Data source: 2004–2016 Nielsen Homescan
- Table/Figure: Table II
- Assumption flag: `False`
- Answer:

Expenditure shares (including the entrant expenditure share and the share at grocery/supercenter/club stores) and the Health Index of household grocery purchases.

- Quote snippet:

  Expenditure shares are the share of total grocery expenditures recorded in Homescan, in units of percentage points.

- Citation anchors:
  – page=21, words=0-349, section=Table II notes, note=Defines expenditure shares and the Health Index and shows these as dependent variables in Panels A and B.
  – page=18, words=0-349, section=Figure IV description, note=Defines the entrant expenditure share outcome.

A05: What are the key treatment/exposure variables (independent variables)?

- Category: `A) Research question / contribution`
- Confidence: `low (0.24192713713243177)`
- Retrieval method: `local`
- Evidence type: `text`
- Data source: Nielsen Homescan panel and RMS store data
- Assumption flag: `False`
- Answer:

Two main exposures: (1) E_{bcqt}—an indicator for whether a supermarket entered within distance band b of a tract by q quarters after quarter t (event-study of entry); (2) H_m—the local (ZIP/county) Health Index of packaged groceries, used in the movers design to capture the retail nutrition environment.

- Quote snippet:

We define Ebcqt as an indicator variable denoting whether one supermarket had entered in distance band b of census tract c by q quarters after quarter t.

- Citation anchors:
  - page=17, words=275-411, section=Event study: supermarket entry, note=Defines Ebcqt as the entry indicator by distance band and event time.
  - page=24, words=0-349, section=Place effects (movers) setup, note=Defines Hm (area Health Index) and its role as the key regressor in movers regression.
  - page=25, words=275-468, section=Place effects identification, note=Variation in local Health Index across moves identifies the effect (supports Hm as exposure).
  - page=21, words=275-503, section=Entry variants, note=Notes analyses considering entry by supercenters as a variant of the entry exposure.

A06: What setting/context does the paper study (country, market, period)?

- Category: `A) Research question / contribution`
- Confidence: `medium (0.3864847718005176)`
- Retrieval method: `local`
- Evidence type: `text`
- Data source: Nielsen Homescan and Nielsen Retail Measurement (RMS)
- Assumption flag: `False`
- Answer:

United States; the grocery/supermarket retail market (packaged groceries via Nielsen Homescan and RMS); period 2004–2016 for Homescan and 2006–2016 for RMS.

- Quote snippet:

  The Nielsen RMS data consist of weekly prices and sales vol- umes for each UPC sold at approximately 42,000 unique stores from 160 retail chains for 2006–2016

- Citation anchors:
  - page=8, words=0-349, section=Data description, note=RMS market coverage and period (grocery retail, 2006–2016).
  - page=26, words=0-282, section=Figure V caption, note=Homescan period stated as 2004–2016.
  - page=24, words=0-349, section=Place effects estimation setup, note=Uses U.S. geographic units (ZIP codes, counties), indicating U.S. setting.
  - page=1, words=275-483, section=Acknowledgments, note=References

to Nielsen Company (US) and USDA ERS, consistent with U.S. setting.

A07: What is the main mechanism proposed by the authors?

- Category: `A) Research question / contribution`
- Confidence: `low` (0.28020166763200416)
- Retrieval method: `local`
- Evidence type: `Event-study and structural demand model estimates`
- Data source: Nielsen Homescan household panel and RMS retail scanner data
- Table/Figure: Figure VII (Explaining the Relationship between Income and Healthy Grocery Demand)
- Assumption flag: `False`
- Answer:

Differences in consumer demand—driven by factors like preferences, education, and nutrition knowledge—rather than neighborhood food supply, are the main mechanism behind nutritional inequality (about 90% demand vs. 10% supply; place effects 3%).

- Quote snippet:

  Only about 10% of the nutrition–income relationship is driven by these differences in supply, while 90% of the relationship is driven by differences in demand.

- Citation anchors:
  - page=6, words=0-349, note=States ~10% of nutrition–income relationship is due to supply, ~90% due to demand.
  - page=42, words=0-349, section=VII.B, note=Emphasizes demand-side factors, not supply, are central.
  - page=48, words=0-323, section=Conclusion, note=Place effects explain no more than 3%; equalizing supply reduces gap by 10%.
  - page=43, words=0-349, section=VII.B, note=Education and nutrition knowledge help explain demand differences.

A08: What alternative mechanisms are discussed?

- Category: `A) Research question / contribution`
- Confidence: `low` (0.21443295838395354)
- Retrieval method: `local`
- Evidence type: `text_quote`
- Table/Figure: Figure VII
- Assumption flag: `False`
- Answer:

Demand-side mechanisms—especially education and nutrition knowledge—are discussed as alternatives to supply-side factors (like prices and availability) for

explaining differences in healthy eating.

- Quote snippet:

  Education explains the largest share of the relationship between demand for healthy groceries and income, at about 20%. Nutrition knowledge explains the second-largest share, at about 14%.

- Citation anchors:
  - page=43, words=0-349, section=VII.B, note=Education and nutrition knowledge cited as key contributors
  - page=38, words=275-377, section=VII.A, note=Decomposition into supply-side vs demand-side factors

A09: What are the main policy implications claimed by the paper?

- Category: `A) Research question / contribution`
- Confidence: `medium (0.394069770413192)`
- Retrieval method: `local`
- Evidence type: `Empirical analyses and model-based counterfactual simulations`
- Data source: Provided context excerpt from The Quarterly Journal of Economics (2020) article
- Assumption flag: `True`
- Assumption notes: Policy conclusions about subsidies rely on model-based simulations with assumptions; authors recommend randomized demonstrations and welfare analysis before adoption.
- Answer:

The paper argues that supply-side policies to eliminate "food deserts" (e.g., promoting supermarket entry or equalizing local supply conditions) have very limited impact on healthy eating and explain little of the income–nutrition gap. It recommends shifting policy focus toward demand-side tools—specifically, means-tested subsidies for healthy foods (potentially through SNAP)—which model-based simulations suggest could cost-effectively raise low-income households' healthy purchases to high-income levels. The authors also urge testing such subsidies via randomized demonstrations and note that any adoption should consider welfare analysis and practical constraints.

- Quote snippet:

  policy makers interested in reducing nutritional inequality might redirect efforts away from promoting access to healthy groceries and toward means-tested subsidies.

- Citation anchors:
  - page=2, words=0-349, section=Introduction, note=Only ~10% of nutritional inequality from supply; counters policies focused on increasing healthy grocery supply.

- page=22, words=0-349, section=Empirical results, note=Local supermarket access explains 1.5% of the Health Index gap.
- page=46, words=0-349, section=VII.C. Using Subsidies to Reduce Nutritional Inequality, note=Cost comparisons: $84M vs supermarket entry; $830M to equalize supply; $11B (~15% of SNAP) to reach top-quartile health level.
- page=47, words=0-349, section=VII.C and VIII. Conclusion, note=Recommendation to redirect efforts from access to means-tested subsidies; call for RCTs; note assumptions and considerations.
- page=48, words=0-323, section=Conclusion, note=Equalizing supply cuts gap by 10%; means-tested subsidy could close gap at ~15% of SNAP; eliminating food deserts yields little progress.

A10: What is the welfare interpretation (if any) of the results?

- Category: `A) Research question / contribution`
- Confidence: `medium (0.38966701151833294)`
- Retrieval method: `local`
- Evidence type: `textual_evidence`
- Assumption flag: `False`
- Answer:

The paper does not provide a welfare analysis. Its cost comparisons are illustrative, not full measures of social costs, and the authors state that welfare conclusions would require measuring market failures and conducting an optimal policy analysis in a welfare-maximization framework.

- Quote snippet:

  Before advocating for or against such a subsidy, one would need to measure the relevant market failures and study optimal policy in a principled welfare maximization framework.

- Citation anchors:
    - page=46, words=0-349, section=VII.C. Using Subsidies to Reduce Nutritional Inequality, note=Cautions that government expenditures are not a complete measure of social costs, indicating no full welfare evaluation.
    - page=48, words=0-323, section=VII.C. Using Subsidies to Reduce Nutritional Inequality / Summary, note=Explicitly states the need for a welfare maximization framework before advocating policies.

A11: What are the main limitations acknowledged by the authors?

- Category: `A) Research question / contribution`
- Confidence: `low (0.2796642781746991)`
- Retrieval method: `local`
- Evidence type: `paraphrase_with_quote`

- Assumption flag: `False`
- Answer:

The authors note several limitations: (1) their structural model relies on many assumptions and so the results should be viewed as suggestive rather than definitive, motivating randomized demonstration projects; (2) some key findings (e.g., contributions of education and nutrition knowledge) are correlational and not causal; (3) estimates of place effects are limited to the several-year time horizon allowed by the data and rely on an assumption that any endogeneity would bias effects upward, and reflect only partial-equilibrium impacts; and (4) before advocating subsidies, relevant market failures must be measured and there are substantial economic and political considerations around modifying SNAP.

- Quote snippet:

  Our model requires many assumptions, so we view these results as suggestive and a potential motivation for demonstration projects, ideally using randomized experiments...

- Citation anchors:
  - page=47, words=0-349, note=Model requires many assumptions; results suggestive; need randomized experiments; policy considerations around SNAP.
  - page=48, words=0-323, note=Limited several-year time horizon; assumption on endogeneity direction; partial equilibrium place effects; need to measure market failures before advocating subsidies.
  - page=43, words=0-349, note=Decomposition results are correlations, not causal effects.

A12: What does the paper claim is novel about its data or identification?

- Category: `A) Research question / contribution`
- Confidence: `low (0.3121536646348006)`
- Retrieval method: `local`
- Evidence type: `explicit statement in text`
- Data source: Nielsen Homescan PanelViews add-on survey (2017)
- Answer:

The paper introduces a novel identification strategy—a cross-sectional price instrument based on retailers' chain-specific relative pricing advantages and geographic presence—and leverages new data from a Homescan add-on survey that links scanner purchases to health preferences and nutrition knowledge.

- Quote snippet:

  adds a novel identification strategy and price instrument.

- Citation anchors:

    – page=7, words=0-349, section=Introduction, note=States the study adds a novel identification strategy and price instrument, and highlights the new Homescan add-on survey linking scanner data to health preferences/knowledge.

    – page=34, words=0-349, section=Identification, note=Explicitly says the instrument is novel in the literature and contrasts it with existing instruments.

B01: What is the identification strategy (in one sentence)?

- Category: B) Identification strategy / causal design
- Confidence: low (0.3493418440321413)
- Retrieval method: local
- Evidence type: quasi-experimental movers design + instrumental variables
- Data source: Nielsen Homescan panel; retailer price/availability by chain and market (RMS)
- Assumption flag: True
- Assumption notes: IV exclusion: chains' comparative advantages are uncorrelated with unobserved tastes (E[_ijt Pjmt]=0); movers design relies on within-household changes with controls for observable shocks.
- Answer:

Exploit within-household changes after cross–ZIP/county moves to identify place effects, and use an IV based on retail chains' comparative pricing advantages (from national UPC prices interacted with local chain presence), so identification comes from relative price differences across product groups within a market-year.

- Quote snippet:

  identification comes only from variation in the relative prices across product groups within a market in a given year.

- Citation anchors:
  - page=23, words=275-408, section=main text, note=Within-household changes after cross–ZIP/county moves to test place effects
  - page=32, words=0-349, section=main text, note=Identification comes from relative prices across product groups within a market-year
  - page=33, words=275-397, section=main text, note=Identifying assumption: instrument uncorrelated with idiosyncratic tastes (E[_ijt Pjmt]=0)
  - page=31, words=0-349, section=main text, note=Instrument intuition: chains have heterogeneous comparative advantages across products and markets

B02: Is the design experimental, quasi-experimental, or observational?

- Category: B) Identification strategy / causal design
- Confidence: low (0.27884377133987565)
- Retrieval method: local
- Evidence type: textual
- Data source: Nielsen Homescan; RMS; ZIP Code Business Patterns (ZBP)
- Answer:

Quasi-experimental

- Quote snippet:

  Before estimating equation (1), we show graphical results of the event study.

- Citation anchors:
  - page=17, words=275-350, section=Event study around entry, note=Mentions showing graphical results of the event study before estimating eq. (1)
  - page=18, words=0-180, section=Event study details, note=Defines event-study regression around supermarket entry
  - page=24, words=275-349, section=Movers event study, note=Again references graphical results of the event study before eq. (3)
  - page=23, words=275-360, section=Place effects and endogeneity, note=Acknowledges non-random moves and endogeneity concerns

B03: What is the source of exogenous variation used for identification?

- Category: B) Identification strategy / causal design
- Confidence: medium (0.43571135442327524)
- Retrieval method: local
- Evidence type: textual
- Data source: Nielsen RMS (Retail Measurement Services) retail scanner data on chains' presence and pricing
- Table/Figure: Online Appendix Table A8; Figures A11, A14, A15
- Answer:

Cross-sectional variation in relative prices across product groups within a market-year, generated by the interaction of retail chains' heterogeneous comparative cost advantages and their differing geographic presence (the price instrument Pjmt).

- Quote snippet:

  The variation in our instrument comes from the interaction between retail chains' differing pricing advantages ... and their differing presence across geographic markets.

13

- Citation anchors:
  - page=32, words=0-349, section=VI.B. Price Endogeneity, note=Defines Pjmt and states that variation comes from chains' pricing advantages interacted with their market presence; identification is from relative prices across product groups within a market-year.
  - page=31, words=0-349, section=VI.B. Price Endogeneity, note=Introduces the instrument and explains chain-specific comparative advantages across products and geographic presence.

B04: What is the treatment definition and timing?

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.12914936559263318)`
- Retrieval method: `local`
- Evidence type: `text`
- Data source: 2004–2016 Nielsen Homescan household-by-quarter data
- Table/Figure: Table II
- Assumption flag: `False`
- Answer:

Treatment is defined as a supermarket entering within a 0–10 or 10–15 minute drive of a household's census-tract centroid. Timing is at the household-by-quarter level, using event-time indicators/counts relative to the entry quarter.

- Quote snippet:

  We consider the impact of supermarket entries that occur within a 0–10- or 10–15-minute drive of households' census tract centroids.

- Citation anchors:
  - page=17, words=0-349, section=Event study setup and regression specification, note=Defines treatment bands (0–10 and 10–15 minutes) and quarter-level data
  - page=17, words=275-411, section=Event-time indicators, note=Defines Ebcqt as indicator by q quarters after t
  - page=19, words=0-310, section=Figure IV description, note=Confirms use of 2004–2016 household-by-quarter Homescan data
  - page=21, words=0-349, section=Table II notes, note=Independent variables are counts of entries within 0–10 or 10–15 minutes as of quarter t

B05: What is the control/comparison group definition?

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.1973947657382083)`
- Retrieval method: `local`
- Evidence type: `methods description`
- Assumption flag: `True`

- Assumption notes: The text outlines event-study and within-household designs but does not explicitly label the control group; we infer it is the pre-event (pre-entry or pre-move) period for the same unit.
- Answer:

A pre-event period for the same unit serves as the comparison group in the event-study designs: for supermarket entry, census tracts before an entry within the specified distance band; for moves, the same households in the balanced pre-move window.

- Quote snippet:

  we measure within-household changes in grocery purchases after the 20,031 cross–ZIP code moves and the 11,728 cross-county moves… Bit be an indicator for whether observation it is part of a balanced panel around a move

- Citation anchors:
  - page=17, words=275-411, section=Event study of supermarket entry, note=Defines event-time indicator for entry within distance bands, implying pre-entry periods as comparison.
  - page=23, words=275-408, section=Within-household move analysis, note=States they measure within-household changes after moves, implying pre-move periods as comparison.
  - page=24, words=275-452, section=Event study around moves, note=Defines Bit for a balanced panel around a move, indicating pre- versus post-move comparison.

B06: What is the estimating equation / baseline regression specification?

- Category: B) Identification strategy / causal design
- Confidence: medium (0.39087819086762554)
- Retrieval method: local
- Evidence type: direct equation from text
- Assumption flag: False
- Answer:

Baseline estimating equation (annual calorie demand): $\ln Y\_ijt = -\ln \tilde{p}\_ijt - \Sigma\_{c\ 2} \tilde{}\_c\ \tilde{a}\_ijct - + \_j + \_m + \_t + \_ijt$, where Y_ijt is total calories purchased by household i in product group j in year t; $\tilde{p}\_ijt$ is the average price per calorie; $\tilde{a}\_ijct$ is the average amount of characteristic c per calorie; $\tilde{}\_c$ are money-metric marginal utilities; is an unobserved product characteristic; and $\_j, \_m, \_t$ are product-group, market, and year fixed effects.

- Quote snippet:

  $\ln Y_i\ jt = -\ln \tilde{p}_i\ jt - \tilde{}c\ \tilde{a}_i\ jct - + j + m + t + i\ jt$

- Citation anchors:

- page=30, words=0-349, section=VI.A. Empirical Model, note=Presents the annual calorie demand estimating equation (7) with fixed effects and variables defined.
- page=29, words=275-369, section=VI.A. Empirical Model, note=Leads into aggregating first-order conditions to form the estimating equation.

B07: What fixed effects are included (unit, time, two-way, higher dimensional)?

- Category: B) Identification strategy / causal design
- Confidence: low (0.3346980299233284)
- Retrieval method: local
- Evidence type: textual evidence from provided context
- Assumption flag: False
- Answer:

Included fixed effects are: household fixed effects (unit); year indicators (time); census-division-by-quarter fixed effects (time-by-region, two-way); household-by-census-tract fixed effects (interaction, higher dimensional). Additional specifications include county, product group, and year fixed effects in instrument tests, and county-by-quarter fixed effects in robustness checks.

- Quote snippet:

  ...census-division-by-quarter of sample indicators, and household-by-census-tract fixed effects.

- Citation anchors:
    - page=21, words=275-503, note=Mentions census-division-by-quarter indicators and household-by-census-tract fixed effects; also notes adding county-by-quarter fixed effects.
    - page=24, words=0-349, note=Defines year indicators (time FE) and household fixed effects (unit FE) for regression (3).
    - page=34, words=0-349, note=States use of county, product group, and year fixed effects in instrument validity tests.

B08: What standard errors are used (robust, clustered; at what level)?

- Category: B) Identification strategy / causal design
- Confidence: low (0.33384988660866577)
- Retrieval method: local
- Evidence type: textual
- Assumption flag: False
- Answer:

Robust standard errors with two-way clustering by household and census tract.

- Quote snippet:

  we use robust standard errors with two-way clustering by household
  and census tract.

- Citation anchors:
  - page=17, words=275-411, section=Event study estimation (equations 1 and 2), note=States use of robust SEs with two-way clustering by household and census tract.
  - page=21, words=275-503, section=Results/table notes, note=Confirms robust SEs clustered by household and census tract.

B09: What is the key identifying assumption (parallel trends, exclusion restriction, ignorability)?

- Category: `B) Identification strategy / causal design`
- Confidence: `medium (0.373532037182493)`
- Retrieval method: `local`
- Evidence type: `explicit statement in text`
- Assumption flag: `True`
- Assumption notes: Instrument exogeneity: chains' comparative pricing advantages are unrelated to unobserved consumer tastes within product groups.
- Answer:

Exclusion restriction: the price instrument must be uncorrelated with unobserved tastes (ijt).

- Quote snippet:

  Our identifying assumption is that household i's idiosyncratic preferences for product group j are uncorrelated with the price instrument Pjmt

- Citation anchors:
  - page=33, words=275-397, section=VI.B Price Endogeneity / Instrument, note=States the identifying assumption that ijt is uncorrelated with the instrument Pjmt.
  - page=34, words=0-349, section=VI.B Instrument validity tests, note=Discusses tests consistent with the exclusion restriction (no relation between instrument and predictable tastes).

B10: What evidence is provided to support the identifying assumption?

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.26630295512584257)`
- Retrieval method: `local`
- Evidence type: `instrument validity checks and robustness tests`

- Data source: Main text; references to Online Appendix Tables A8 and A9
- Assumption flag: `False`
- Answer:

They support the identifying assumption with three pieces of evidence: (1) most identifying variation in the price instrument comes from retailer-specific pricing advantages that are constant across counties (not from county-specific pricing), implying little scope for endogenous responses to local tastes; (2) two balance tests show no correlation between the instrument and predictable tastes—no association with purchases predicted by demographics and no systematic differences in high–Health Index product groups across low- vs. high-income counties; and (3) estimates are robust to adding/removing controls for preference variation, yielding very similar results.

- Quote snippet:

  We find that the instrument is not associated with predicted purchases conditional on our standard set of county, product group, and year fixed ef- fects.

- Citation anchors:
  - page=32, words=0-349, section=Instrument and identification, note=Variation mainly retailer-specific (constant across counties), limiting endogenous pricing concerns.
  - page=33, words=275-397, section=Identifying assumption, note=States the identifying assumption E[_ijt P_jmt]=0 and discusses testing strategy.
  - page=34, words=0-349, section=Instrument validity tests, note=Two tests: no association with predicted purchases; no systematic differences across income in high-Health Index groups.
  - page=38, words=275-377, section=Robustness to preference controls, note=Estimates very similar with added preference controls, reinforcing exclusion restriction plausibility.

B11: Are there event-study or pre-trend tests? What do they show?

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.32193505733248057)`
- Retrieval method: `local`
- Evidence type: `text`
- Data source: 2004–2016 Nielsen Homescan household data
- Table/Figure: Figure IV; Figure V
- Assumption flag: `False`
- Answer:

Yes. The paper runs event-study analyses around supermarket entry (eq. 2; Figure IV) and around household moves across counties (Figure V). After entry, entrant retailers' expenditure share rises by ~2 percentage points in a year (3%

18

in food deserts), with Health Index effects small/mostly insignificant and gains largely diverted from other supermarkets. For moves, there is no statistically significant post-move change in the Health Index when moving to healthier counties, though point estimates are positive. The entry event study uses a balanced panel with four pre-entry quarters and sets q = −1 as the omitted pre-period, allowing visual pre-trend assessment.

- Quote snippet:

  The top left panel shows that this entrant expenditure share increases by about 2 percentage points one year after entry.

- Citation anchors:
  - page=18, words=0-349, section=Event study of supermarket entry (eq. 2; Figure IV), note=Describes event-study setup, balanced panel, omitted q = −1, and ~2 pp increase post-entry
  - page=19, words=0-310, section=Figure IV description, note=Specifies outcomes including Health Index in event study panels
  - page=21, words=275-503, section=Entry effects interpretation, note=Health Index effects small/mostly insignificant; gains largely diverted from other supermarkets
  - page=26, words=0-282, section=Event study of moves (Figure V), note=Reports no statistically significant post-move Health Index change; point estimates positive
  - page=17, words=275-411, section=Event study setup details, note=Mentions showing graphical event-study results and defines key indicators

B12: What instruments are used (if IV)? Define instrument and first stage.

- Category: B) Identification strategy / causal design
- Confidence: low (0.18485188819713624)
- Retrieval method: local
- Evidence type: textual_evidence
- Table/Figure: Equation (9); Online Appendix Figure A17
- Assumption flag: True
- Assumption notes: Exclusion restriction: E[ijt Pjmt]=0; chains' comparative pricing advantages are unrelated to unobserved tastes. Market fixed effects address overall demand patterns.
- Answer:

Instrument: A chain-level, cross-market pricing instrument Pjmt. For each UPC k and chain r, compute the chain's cost advantage outside market m as $\Delta\ln(pkrt,−m) = \ln(pkrt,−m) − \ln(pkt,−m)$, where prices exclude market m. Then Pjmt is the weighted average of these cost advantages across UPCs in product group j, with weights based on the chain's presence in market m (Nrmt), chain-by-group sales weights (Njrt), and national UPC importance (Nkt) [Eq.

(9)]. First stage: Regress log prices on Pjmt with market and product-group fixed effects; there is a robust linear relationship, with first-stage F-statistics around 243–260 across income groups.

- Quote snippet:

  The price instrument Pjmt is the weighted average cost advantage
  that chains in market m have for UPCs in product group j:

- Citation anchors:
  - page=page 31, words=275-432, note=Defines ln(pkrt,−m), ln(pkt,−m) and exclusion of market m to avoid local demand contamination.
  - page=page 32, words=0-349, note=Defines Pjmt (Eq. 9), weighting scheme, and variation source; mentions fixed effects context.
  - page=page 33, words=0-349, note=States robust linear first-stage and F-statistics for instrument predicting log prices with FE.
  - page=page 33, words=275-397, note=Identifying assumption E[ijt Pjmt]=0; economic content of exclusion restriction.

B13: What is the first-stage strength (F-stat, partial R^2, relevance evidence)?

- Category: B) Identification strategy / causal design
- Confidence: low (0.30479736690202236)
- Retrieval method: local
- Evidence type: Qualitative description of instrument variation and cited appendix checks; no numeric first-stage stats in provided text
- Table/Figure: Online Appendix Table A8; Figures A11, A14, A15 (referenced)
- Assumption flag: False
- Answer:

Numeric first-stage strength (F-stat or partial R^2) is not reported in the provided excerpts. The relevance evidence given is qualitative: variation comes from retailer-specific pricing advantages and retailers' differing geographic presence, with identification from relative price differences across product groups within a market-year; appendix evidence indicates most within–product-group variation is retailer- (not county-) specific.

- Quote snippet:

  identification comes only from variation in the relative prices across
  product groups within a market in a given year.

- Citation anchors:
  - page=32, words=0-349, section=Instrument construction and variation, note=Defines instrument; states where variation and identi-

20

fication come from; references appendix showing variation largely retailer-specific.
  – page=34, words=0-349, section=Instrument validity checks, note=Reports tests showing instrument not associated with predicted tastes (exogeneity checks).
  – page=33, words=275-397, section=Identifying assumption, note=States exclusion restriction; context for instrument validity (not first-stage strength).

B14: If RDD: what is the running variable and cutoff? bandwidth choice?

- Category: B) Identification strategy / causal design
- Confidence: low (0.26722121815091543)
- Retrieval method: local
- Evidence type: textual
- Data source: Nielsen Homescan household panel; Nielsen RMS store data; Google Maps API driving times
- Assumption flag: False
- Answer:

Not applicable. The study does not use an RDD; it relies on event-study regressions around supermarket entries and household moves (and an IV for prices). Therefore, there is no running variable, cutoff, or bandwidth.

- Quote snippet:

  Before estimating equation (1), we show graphical results of the event study.

- Citation anchors:
  – page=17, words=0-349, section=Event study around supermarket entries, note=Describes event-study setup and regression (1), not an RDD.
  – page=24, words=0-349, section=Event study around household moves, note=Describes event-study design and regression (3), not an RDD.
  – page=32, words=0-349, section=Instrumental variables for prices, note=Introduces a cross-sectional IV; again not an RDD.

B15: If DiD: what is the timing variation (staggered adoption)? estimator used?

- Category: B) Identification strategy / causal design
- Confidence: low (0.3211861386487479)
- Retrieval method: local
- Evidence type: textual_excerpts
- Data source: Homescan household-by-quarter data

- Table/Figure: Figure IV
- Assumption flag: `False`
- Answer:

Staggered timing comes from supermarket openings occurring in different quarters across census tracts within 0–10 or 10–15 minutes' drive; event time is defined around the quarter of entry (balanced panels: 4 pre, 8 post). The estimator is a two-way fixed-effects DiD/event-study in household-by-quarter data: Yict = b Sbct + Xit + d(c)t + ic + ict and, for event time, Yict = Bbit Σq bq Ebcqt + b + Xit + d(c)t + ic + ict, with q = −1 omitted and robust SE two-way clustered by household and census tract.

- Quote snippet:

  We run the following regression in household-by-quarter Homescan data: Yict = b Sbct + Xit + d(c)t + ic + ict … We run the following regression… Yict = Bbit bq Ebcqt + b … The omitted category is q = −1.

- Citation anchors:
  - page=17, words=0-349, section=Equation (1) DiD setup, note=Defines distance bands, Sbct, and TWFE regression with division-by-quarter and household-by-tract FE; mentions two-way clustering.
  - page=18, words=0-349, section=Equation (2) event study, note=Defines balanced panels around entry, event-time regression, and that q = −1 is the omitted category.
  - page=17, words=275-411, section=Event study definitions; clustering, note=Introduces Ebcqt, Bbit and notes two-way clustering by household and census tract.

C01: What dataset(s) are used? (name sources explicitly)

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.323291429229872)`
- Retrieval method: `local`
- Evidence type: `textual_evidence`
- Data source: Allcott et al. (2020), The Quarterly Journal of Economics article text (provided context)
- Assumption flag: `False`
- Answer:

Nielsen Homescan household panel (including the 2004–2006 "magnet" subsample and the 2017 PanelViews add-on survey); Nielsen Retail Measurement Services (RMS); Nielsen TDLinx (store opening dates/locations); ZIP Code Business Patterns (ZBP); USDA Food and Nutrient Database for Dietary Studies (FNDDS) and National Nutrient Database for Standard Reference (SR); Consumer Price Index for Urban Consumers (CPI-U) for deflation.

- Quote snippet:

  Our nutrition facts are from the Food and Nutrient Database for Dietary Studies and the National Nutrient Database for Standard Reference

- Citation anchors:
  - page=8, words=1-200, section=Data description, note=Nielsen Homescan; 8,000-household 'magnet' subsample; PanelViews add-on survey
  - page=8, words=275-393, section=Data description, note=Nielsen RMS; CPI-U deflator
  - page=9, words=275-389, section=II.B. Grocery Retail Establishments, note=ZIP Code Business Patterns (ZBP)
  - page=10, words=1-180, section=II.B. Grocery Retail Establishments, note=Nielsen TDLinx store openings data
  - page=10, words=180-349, section=II.C. Nutrition Facts and the Health Index, note=USDA FNDDS and SR nutrition databases; USDA crosswalks

C02: What is the unit of observation (individual, household, firm, county, transaction, product)?

- Category: `C) Data, sample, and measurement`
- Confidence: `medium (0.40760063179281236)`
- Retrieval method: `local`
- Evidence type: `explicit statement in text`
- Data source: Nielsen Homescan
- Assumption flag: `False`
- Answer:

Household-by–product group–by-year

- Quote snippet:

  Data are at the household-by–product group–by-year level.

- Citation anchors:
  - page=34, words=275-385, section=VI.C. Method of Moments Estimation, note=States the estimation dataset unit
  - page=9, words=275-389, note=Notes collapsing to household (or store)-by-year for analyses

C03: What is the sample period and geographic coverage?

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.2994701740878031)`
- Retrieval method: `local`
- Evidence type: `direct_author_text`

- Data source: Nielsen Homescan; Nielsen TDLinx; ZIP Code Business Patterns
- Assumption flag: `False`
- Answer:

2004–2016; United States (nationwide). The Homescan panel is nationally representative, ZIP Code Business Patterns cover every ZIP code, and supermarket entries come from a census of all U.S. food retailers.

- Quote snippet:

  Homescan data include 731,994 household-by-year observations for 2004–2016 and are weighted for national representativeness.

- Citation anchors:
  - page=9, words=0-349, section=Table I notes / Data description, note=Homescan period 2004–2016; nationally representative.
  - page=10, words=0-349, section=II.B. Grocery Retail Establishments, note=TDLinx is a census of all U.S. food retailers; 2004–2016 entries.
  - page=10, words=0-349, section=II.B. Grocery Retail Establishments, note=ZBP provides counts for every ZIP code.
  - page=9, words=275-389, section=II.B. Grocery Retail Establishments, note=ZBP data span 2004–2016.

C04: What are the sample restrictions / inclusion criteria?

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.2702960096992867)`
- Retrieval method: `local`
- Evidence type: `direct_excerpts`
- Data source: Nielsen Homescan; Nielsen Retail Scanner (RMS); ZIP Code Business Patterns (ZBP)
- Assumption flag: `False`
- Answer:

The analyses use Nielsen Homescan households and Nielsen RMS stores, with the following core sample restrictions/inclusion criteria: (1) RMS excludes liquor stores; (2) both Homescan and RMS exclude bulk produce and other nonpackaged items (focus on packaged groceries); (3) for most empirical analyses (except the supermarket-entry event study), data are collapsed to the household- or store-by-year level; (4) entry and heterogeneity analyses sometimes restrict to subsamples—e.g., households in the bottom income quartile (defined by residualized income) and households in food deserts; (5) the household-move design includes only observations in a balanced panel around a move (Bit = 1) and imposes a 50% local shopping restriction to ensure exposure to the local retail environment; (6) some estimations stratify the sample into four income quartiles for separate datasets.

- Quote snippet:

As with Homescan, RMS does not include sales of bulk produce and other nonpackaged items.

- Citation anchors:
  - page=8, words=275-393, section=RMS sample definition, note=Excludes liquor stores; no bulk produce/nonpackaged items
  - page=9, words=275-389, section=Data construction, note=Collapse to household- or store-by-year level
  - page=20, words=275-432, section=Subsample analyses, note=Limits to bottom income quartile and to food deserts; definition of bottom quartile
  - page=24, words=275-452, section=Move design sample, note=Balanced panel around a move (Bit definition)
  - page=25, words=275-468, section=Move design restriction, note=50% local shopping restriction
  - page=34, words=275-385, section=Estimation stratification, note=Separate datasets for four household income quartiles

C05: What is the sample size (N) in the main analysis?

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.28304756286944066)`
- Retrieval method: `local`
- Evidence type: `Table note (descriptive statistics)`
- Data source: Nielsen Homescan data, 2004–2016
- Table/Figure: Table I (Notes)
- Assumption flag: `True`
- Assumption notes: Interpreted 'main analysis' as the Homescan household-by-year sample used throughout; ZIP code counts (508,951) also reported but pertain to establishment data, not the main household analysis.
- Answer:

731,994 household-by-year observations (Homescan, 2004–2016).

- Quote snippet:

  Homescan data include 731,994 household-by-year observations for 2004–2016

- Citation anchors:
  - page=9, words=0-349, section=Table I notes (Panel A), note=States total Homescan household-by-year observations for 2004–2016.

C06: How is the key outcome measured? Any transformations (logs, z-scores, indices)?

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.3488158062475002)`

- Retrieval method: `local`
- Evidence type: `textual evidence from methods/notes`
- Data source: Nielsen Homescan (2004–2016)
- Table/Figure: Table III
- Assumption flag: `False`
- Answer:

The key outcome is the Health Index—an overall, linearized Healthy Eating Index of grocery purchases. It is standardized to mean 0 and SD 1 across households (z-score). In some counterfactual figures it is further renormalized so the base high–low income gap equals 1.

- Quote snippet:

  The Health Index is our overall measure of the healthfulness of grocery purchases and is normalized to have a mean of 0 and a standard deviation of 1 across households.

- Citation anchors:
  - page=27, words=0-349, section=Table III notes, note=Defines Health Index and states it is normalized to mean 0, SD 1 across households.
  - page=11, words=275-399, section=Methods/Measurement, note=Describes linearized HEI and its correlation with the true HEI; notes omission of fatty acid ratio.
  - page=40, words=0-285, section=Figure VI caption, note=States the Health Index is renormalized so the base difference between highest and lowest income groups equals 1.
  - page=36, words=275-508, section=Estimation Results/HEI components, note=Clarifies HEI components and units used in constructing the index.

C07: How is treatment/exposure measured? Any constructed variables?

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.2926683796776228)`
- Retrieval method: `local`
- Evidence type: `textual description of methods and variable construction`
- Data source: Nielsen Homescan and RMS; ZBP for store-type distribution adjustment
- Assumption flag: `False`
- Answer:

Exposure is measured as: (1) supermarket entry near households via Ebcqt—an indicator for whether one supermarket entered within distance band b of a census tract by q quarters after quarter t (with an additional indicator Bbit);

and (2) the local retail environment's healthfulness Hm—the Health Index of packaged groceries in a ZIP code or county, constructed as the calorie-weighted average Health Index of groceries sold in RMS stores and regression-adjusted to match the true distribution of store channel types from ZBP. For mover event studies, Bit flags observations in a balanced panel around a move; they also apply a 50% local shopping restriction and use the realized change in the local Health Index across moves. Constructed variables include Ebcqt, Bbit/Bit, the area Health Index Hm (calorie-weighted and store-mix adjusted), and a linearized Healthy Eating Index (HEI) used to summarize dietary quality.

- Quote snippet:

  We define Ebcqt as an indicator variable denoting whether one supermarket had entered in distance band b of census tract c by q quarters after quarter t.

- Citation anchors:
  - page=17, words=275-411, section=Event study design (store entry), note=Defines Ebcqt (entry indicator by distance band and time); mentions Bbit indicator.
  - page=24, words=0-349, section=Place effects specification, note=Defines Hm (area Health Index) and Bit (balanced panel around a move).
  - page=24, words=275-452, section=Construction of area Health Index, note=Details calorie-weighted averaging and regression adjustment to match ZBP store-type distribution.
  - page=25, words=275-468, section=Mover exposure measurement, note=Shows change in local Health Index across moves; mentions 50% local shopping restriction and trips shifting to the new county.
  - page=11, words=275-399, section=Outcome/quality measure construction, note=Describes the linearized HEI and its components.

C08: Are there key covariates/controls? Which ones are always included?

- Category: `C) Data, sample, and measurement`
- Confidence: `medium (0.35539971550701116)`
- Retrieval method: `local`
- Evidence type: `textual`
- Assumption flag: `False`
- Answer:

Yes. A standard demographic vector Xit is routinely included: log income, log years of education, age indicators, household size, race indicators, marital status, employment status, and weekly work hours. Entry regressions also include census-division-by-quarter indicators and household-by-census-tract fixed effects. Mover/event-study regressions include year indicators and household

fixed effects (with the same Xit), and additionally control for changes in income, job responsibilities, household composition, and marriage status.

- Quote snippet:

  Xit includes the natural log of income, natural log of years of education, indicators for each integer age from 23–90, household size, race indicators,

- Citation anchors:
  - page=17, words=275-411, section=Covariates (Xit), note=Lists Xit demographics: log income, log education, age dummies, household size, race, marital status, employment status, weekly hours.
  - page=21, words=275-360, section=Fixed effects in entry regressions, note=States inclusion of census-division-by-quarter indicators and household-by-census-tract fixed effects.
  - page=24, words=0-200, section=Mover/event-study specification, note=Defines regression with year indicators and household fixed effects and reuses Xit.
  - page=23, words=300-390, section=Additional controls in mover analysis, note=Mentions adding controls for changes in income, job responsibilities, household composition, and marriage status.

C09: How are missing data handled (dropping, imputation, weighting)?

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.3056674984716571)`
- Retrieval method: `local`
- Evidence type: `textual evidence`
- Assumption flag: `False`
- Answer:

Missing data are handled by dropping zero-purchase cells and using survey weights. Specifically, they drop 10.6% of household–product group–year observations with zero purchases, and Homescan results are generally weighted for national representativeness unless noted; some analyses are explicitly unweighted. No imputation is described in the provided context.

- Quote snippet:

  We drop 10.6% of observations at the household-by–product group–by-year level because they have zero purchases.

- Citation anchors:
  - page=35, words=275-428, section=Main text, note=Dropping zero-purchase observations (10.6%) and discussion of missing observations
  - page=8, words=275-393, section=Main text, note=Homescan results weighted for national representativeness

– page=21, words=275-503, section=Main text, note=Example noting observations are not weighted in a particular analysis

C10: Are weights used (survey weights, propensity weights)? How?

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.3415259310295733)`
- Retrieval method: `local`
- Evidence type: `text from provided context`
- Data source: Nielsen Homescan (consumer panel) and Nielsen RMS (retail scanner) data
- Assumption flag: `False`
- Answer:

Yes. Homescan survey/sample weights are used to make estimates nationally representative—"unless otherwise stated, all Homescan results are weighted." They explicitly note using Homescan sample weights in an analysis, while some regressions are explicitly not weighted for national representativeness. Additionally, certain constructs (e.g., area Health Index) use calorie-weighted averages of store sales. No propensity weights are mentioned in the provided text.

- Quote snippet:

  Unless otherwise stated, all Homescan results are weighted for national representativeness.

- Citation anchors:
  - page=8, words=275-393, section=main text, note=States that Homescan results are weighted for national representativeness unless otherwise noted.
  - page=17, words=275-411, section=main text, note=Explicitly mentions using Homescan sample weights for the analysis.
  - page=21, words=275-503, section=main text, note=Notes an instance where observations are not weighted for national representativeness.
  - page=24, words=275-452, section=main text, note=Describes using calorie-weighted averages when constructing area Health Index measures.

C11: Are data linked/merged across sources? How is linkage performed?

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.2555935989172092)`
- Retrieval method: `local`
- Evidence type: `methods description`
- Data source: Nielsen Homescan, Nielsen RMS, Nielsen TDLinx, USDA FNDDS/Standard Reference, ZIP Code Business Patterns, Nielsen Pan-

elViews
- Assumption flag: `False`
- Answer:

Yes. Several cross-source linkages are used: (1) USDA nutrition facts are matched to product UPCs via USDA-provided crosswalks; (2) TDLinx supermarket entries are matched to Nielsen Homescan retailer codes (with filters to drop spurious entries); (3) Area Health Indexes from Nielsen RMS are regression-adjusted using ZIP Code Business Patterns' store-type distributions; (4) Nielsen PanelViews survey responses are merged to Homescan households; (5) Household calorie needs combine U.S. Dietary Guidelines with Homescan household composition.

- Quote snippet:

  We match these nutrition facts to UPCs using crosswalks developed by the USDA (Carlson et al. 2019).

- Citation anchors:
  - page=10, words=0-349, section=II.B. Grocery Retail Establishments, note=TDLinx entries matched to Homescan retailer codes
  - page=10, words=275-424, section=II.C. Nutrition Facts and the Health Index, note=USDA nutrition facts matched to UPCs via USDA crosswalks; TDLinx entry cleaning
  - page=24, words=275-452, section=Event study methodology, note=RMS area Health Index regression-adjusted using ZBP store-type distributions
  - page=8, words=0-349, section=Data description, note=PanelViews survey linked to Homescan; calorie needs combined with Homescan composition
  - page=9, words=275-389, section=Data sources, note=ZBP establishment counts used for ZIP-by-year linkage context

C12: What summary statistics are reported for main variables?

- Category: `C) Data, sample, and measurement`
- Confidence: `medium (0.35556736097602754)`
- Retrieval method: `local`
- Evidence type: `table`
- Data source: Nielsen Homescan households and ZIP Code Business Patterns (2004–2016)
- Table/Figure: Table I: Descriptive Statistics (Panels A and B)
- Assumption flag: `False`
- Answer:

Means and standard deviations are reported (Table I) for Homescan household variables—income, education, age, household size, race shares (White/Black), marital and employment status, weekly work hours, household daily calorie need,

health importance, nutrition knowledge, BMI, diabetic—and for ZIP code establishment counts—grocery, large grocery, supercenters/club stores, drug stores, convenience stores, and meat/fish/produce stores.

- Quote snippet:

  TABLE I DESCRIPTIVE STATISTICS Standard Variable Mean deviation

- Citation anchors:
  - page=9, words=0-349, section=Table I (Panels A and B), note=Table lists variables with columns for Mean and Standard deviation, covering Homescan household variables and ZIP code establishment counts with data-source notes.

C13: Are there descriptive figures/maps that establish baseline patterns?

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.3249774887968107)`
- Retrieval method: `local`
- Evidence type: `figures/maps (descriptive)`
- Data source: Nielsen RMS (2006–2016)
- Table/Figure: Figure II; Online Appendix Figures A11, A15
- Assumption flag: `False`
- Answer:

Yes. The paper includes descriptive figures and maps that establish baseline patterns—for example, Figure II plots store healthfulness against ZIP code median income, and Online Appendix Figure A11 maps the geographic presence of major retailers (with A15 illustrating geographic price variation).

- Quote snippet:

  Online Appendix Figure A11 maps the geographic presence of the five largest retailers in RMS

- Citation anchors:
  - page=14, words=0-230, section=III.B. Supply Disparities by Neighborhood Income, note=Description of Figure II showing relationships between income and healthfulness offered in stores.
  - page=32, words=275-431, section=Appendix figures discussion, note=Mentions Online Appendix Figure A11 (maps retailer presence) and Figure A15 (geographic variation).

D01: What is the headline main effect estimate (sign and magnitude)?

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `medium (0.39317042595932583)`
- Retrieval method: `local`

- Evidence type: `textual`
- Data source: Nielsen Homescan (as referenced elsewhere for these estimates)
- Table/Figure: Table V (text; referencing Table II, Panel B, col (2))
- Assumption flag: `False`
- Answer:

Negative 0.9%: one supermarket entry within a 10-minute drive reduces the top–minus–bottom income Health Index gap by about 0.9%.

- Quote snippet:

  one supermarket entry within a 10-minute drive of a bottom income quartile household reduces the top minus bottom quartile Health Index difference by 0.005 0.56   0.9%.

- Citation anchors:
  - page=45, words=275-394, note=States the point estimate of supermarket entry's impact on the Health Index gap (~0.9% reduction).

D02: What is the preferred specification and why is it preferred?

- Category: D) Results, magnitudes, heterogeneity, robustness
- Confidence: `low (0.22637865371004126)`
- Retrieval method: `local`
- Evidence type: `textual`
- Data source: QJE article: Food Deserts and Nutritional Inequality (excerpts provided)
- Assumption flag: `True`
- Assumption notes: Interprets 'preferred specification' as the IV model using the chain-based price instrument; the text does not explicitly label a 'preferred specification'.
- Answer:

The preferred specification is the IV demand model that instruments for product-group prices using grocery chains' comparative-advantage-driven price variation across markets (Pjmt). It is preferred because it addresses price endogeneity (E[, p]   0), relies on the identifying assumption E[, Pjmt] = 0, has a very strong first stage, and delivers estimates that are robust to adding geography/income and other preference controls.

- Quote snippet:

  we introduce a new instrument using price variation from chains' comparative advantages... The instrument has a very strong first stage and is uncorrelated with observably predicted tastes.

- Citation anchors:

- page=5, words=0-349, section=Instrument introduction and assumptions, note=Introduces chain-based price instrument; key assumption; strong first stage; robustness to controls.
- page=30, words=275-393, section=VI.B. Price Endogeneity, note=Motivates need for IV due to E[, p] 0.
- page=33, words=275-397, section=Identification assumption, note=States E[ijt, Pjmt] = 0 and discusses exclusion logic.
- page=38, words=275-377, section=Robustness to controls, note=Estimates similar with added controls, supporting instrument validity.

D03: How economically meaningful is the effect (percent change, elasticity, dollars)?

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `medium (0.44032523075622043)`
- Retrieval method: `local`
- Evidence type: `Reported effect sizes and WTP estimates from main text, figures, and tables`
- Data source: Nielsen Homescan household-by-quarter data; ZBP
- Table/Figure: Figure IV; Table II, Panel B
- Assumption flag: `False`
- Answer:

Economically small. Supermarket entry raises entrant chains' expenditure share by about 2 percentage points (~\$2 per \$100; ~3 pp in food deserts), but effects on healthy consumption are tiny: one entry increases the Health Index by no more than ~0.02 SD and explains at most ~1.5% of the income–nutrition gap. Preferences translate to modest willingness to pay for healthier bundles: roughly 0.43–1.14 per 1,000 calories (or 0.20–0.63 with added controls).

- Quote snippet:

  entrant expenditure share increases by about 2 percentage points one year after entry.

- Citation anchors:
    - page=18, words=0-349, section=Main text (Figure IV discussion), note=Entrant expenditure share increases ~2 pp; ~\$2 per \$100; ~3% in food deserts.
    - page=22, words=0-349, section=Main text (Table II, Panel B), note=Effects on Health Index economically small; upper bound ~0.02 SD; access explains ~1.5% of gap.
    - page=37, words=275-433, section=Main text (WTP for healthy bundle), note=WTP: \$0.43 (lowest income) to \$1.14 (highest) per 1,000 calories.
    - page=38, words=0-349, section=Main text (WTP with added controls), note=With shelf-life/convenience controls, WTP: \$0.20 to

$0.63 per 1,000 calories.

D04: What are the key robustness checks and do results survive them?

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `low (0.33043848420519606)`
- Retrieval method: `local`
- Evidence type: `textual evidence from article excerpts`
- Data source: QJE article excerpts (pages 21, 27, 34, 38)
- Table/Figure: Online Appendix Tables A5, A9, A11
- Assumption flag: `False`
- Answer:

Key checks include: (1) redefining entry to only supercenters—expenditure-share effects get larger but Health Index effects remain economically small and mostly insignificant; (2) adding county-by-quarter fixed effects and controlling for convenience store entry—results do not qualitatively change; (3) in mover/event-study regressions, adding demographic controls has very little impact on estimates; (4) instrument validity tests—no correlation with tastes predicted by demographics and no systematic differences across high- vs low-income counties' high-Health Index categories; (5) adding/removing preference controls (e.g., county median income above national) yields very similar estimates. Overall, the main conclusions survive these checks.

- Quote snippet:

  Adding county-by-quarter fixed effects or a control for convenience store entry does not qualitatively change the results.

- Citation anchors:
  - page=21, words=275-503, section=Entry effects robustness checks, note=Supercenter-only entry; adding county-by-quarter fixed effects; controlling for convenience store entry; Health Index effects small.
  - page=27, words=275-400, section=Event-study controls, note=Including demographics has very little impact; R^2 only slightly increases.
  - page=34, words=0-349, section=Instrument validity tests, note=Instrument not associated with predicted purchases; not systematically different across income-linked health categories.
  - page=38, words=275-377, section=Exclusion restriction robustness, note=Estimates very similar when adding/removing controls like county median income above national.

D05: What placebo tests are run and what do they show?

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `low (0.2010298943364759)`
- Retrieval method: `local`

- Evidence type: `text_quote`
- Assumption flag: `True`
- Assumption notes: The text does not label these as 'placebo' explicitly; we interpret the pre-trend check and the instrument–tastes null-correlation tests as placebo-style validations.
- Answer:

They run two main placebo-style checks. First, an event-study pre-trend test around household moves: before a move, there is no trend in healthy eating; after moving, purchases converge only slightly toward the new area's pattern. Second, two instrument-validity placebos: (i) using demographics to predict purchases at the household/county level and testing whether the price instrument correlates with those predicted tastes—there is no association; (ii) testing whether the instrument is systematically different in high–Health Index product groups between low- and high-income counties—again, no systematic difference. As an additional validation, they replicate the known immediate Coke-vs-Pepsi brand-choice effect using the same strategy, confirming the design detects expected effects.

- Quote snippet:

  We find that the instrument is not associated with predicted purchases conditional on our standard set of county, product group, and year fixed effects.

- Citation anchors:
  - page=4, words=275-423, section=Event study around moves, note=No pre-move trend in healthy eating; small post-move convergence.
  - page=34, words=0-349, section=Instrument validity tests, note=Instrument not associated with predicted purchases; not systematically different across income in high-Health Index groups.
  - page=27, words=275-482, section=Validation via Coke vs. Pepsi effect, note=Replicates known brand-choice effect as a design check.

D06: What falsification outcomes are tested (unaffected outcomes)?

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `low (0.25114935550222117)`
- Retrieval method: `local`
- Evidence type: `Placebo/falsification (instrument validity) tests`
- Table/Figure: Online Appendix Table A9
- Answer:

They run placebo-style validity checks showing the instrument is unrelated to outcomes it should not affect: (1) predicted purchases/tastes based on household and county demographics, and (2) whether the instrument is systematically

different in high–Health Index product groups across low- versus high-income counties. In both cases, there is no relationship.

- Quote snippet:

  We find that the instrument is not associated with predicted purchases... We find that the instrument is not systematically different in high–Health Index product groups in low- versus high-income counties.

- Citation anchors:
  - page=34, words=0-349, section=Instrument validity tests, note=Two tests: instrument unrelated to predicted purchases from demographics; not systematically different in high–Health Index product groups across income levels.

D07: What heterogeneity results are reported (by income, size, baseline exposure, region)?

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `medium (0.4340082483189159)`
- Retrieval method: `local`
- Evidence type: `text`
- Data source: Nielsen Homescan (2004–2016)
- Table/Figure: Figure VII
- Assumption flag: `False`
- Answer:

Heterogeneity shows: (1) by income—expenditure share changes are generally larger for low-income households; (2) by size—when focusing on supercenter entry (larger stores), expenditure share effects are larger, while Health Index effects remain economically small and mostly insignificant; (3) by baseline exposure—expenditure share changes are larger for households in food deserts; (4) by region—adding census-region-by–product-group (and related) controls leaves estimates very similar, and the role of Census divisions is accounted for in the decomposition.

- Quote snippet:

  the expenditure share changes are generally larger for low-income households and households in food deserts.

- Citation anchors:
  - page=21, words=275-430, section=Entry effects and heterogeneity, note=Larger expenditure share changes for low-income households and those in food deserts.
  - page=21, words=430-503, section=Entry effects and heterogeneity (supercenters), note=Supercenter-only entry yields larger expenditure share effects; Health Index effects small/mostly insignificant.

- page=38, words=0-220, section=Regional controls robustness, note=Adding census-region-by–product group and interactions yields very similar estimates.
- page=44, words=0-150, section=Decomposition figure note, note=Figure VII includes a Census division bar reflecting regional contribution.

D08: What mechanism tests are performed and what do they imply?

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `low (0.16578383654914944)`
- Retrieval method: `local`
- Evidence type: `textual`
- Data source: Article text (pages 33–34)
- Table/Figure: Online Appendix Table A9
- Assumption flag: `False`
- Answer:

They run two predictive-taste tests of the price instrument: (1) predict product-group purchases using household and county demographics and test whether the instrument correlates with these predicted purchases—finding no association; and (2) use the nutrition–income gradient to test whether the instrument is systematically different for high–Health Index product groups in low- vs. high-income counties—finding no systematic differences. These results imply the instrument is not correlated with predictable tastes, supporting the exclusion restriction and suggesting it captures supply-side comparative pricing rather than demand-driven preferences.

- Quote snippet:

  We implement two tests, each of which predicts tastes in a different way... We find that the instrument is not associated with predicted purchases... not systematically different in high–Health Index product groups...

- Citation anchors:
  - page=34, words=0-349, note=Describes two tests and null findings on association with predicted tastes and income-based healthy product groups.
  - page=33, words=275-397, note=States the identifying assumption and motivation for tests against predictable tastes.

D09: How sensitive are results to alternative samples/bandwidths/controls?

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `medium (0.37156554303415845)`
- Retrieval method: `local`
- Evidence type: `text`

- Data source: Provided QJE article excerpts (pp. 21, 27, 34, 38)
- Table/Figure: Online Appendix Table A5; Online Appendix Table A11
- Assumption flag: `False`
- Answer:

Results appear robust to alternative controls and samples. Adding county-by-quarter fixed effects and a convenience-store-entry control does not qualitatively change results; adding household demographic controls has very little impact (only slightly increases R^2). Restricting to supercenter-only entry yields larger expenditure-share effects but Health Index effects remain small and mostly insignificant. Instrument-validity checks show no association with predictable tastes, and estimates are very similar when adding/removing preference controls. The provided excerpts do not report bandwidth-sensitivity checks.

- Quote snippet:

  Adding county-by-quarter fixed effects or a control for convenience store entry does not qualitatively change the results.

- Citation anchors:
  - page=21, words=275-503, section=Entry effects and robustness, note=County-by-quarter FE and convenience-store-entry control; supercenter-only robustness; main effects unchanged
  - page=27, words=275-482, section=Place-effects regressions, note=Including demographic controls has very little impact; R^2 increases only slightly
  - page=34, words=0-349, section=Instrument validity tests, note=Instrument unrelated to predictable tastes based on demographics and income-linked Health Index patterns
  - page=38, words=275-377, section=Additional instrument robustness, note=Estimates very similar with alternative controls (Online Appendix Table A11)

D10: What are the main takeaways in the conclusion (bullet summary)?

- Category: `D) Results, magnitudes, heterogeneity, robustness`

- Confidence: `low (0.2529475933271664)`

- Retrieval method: `local`

- Evidence type: `Direct quotations and paraphrase from the article's conclusion and supporting results`

- Assumption flag: `False`

- Answer:

- Supermarket entry has economically small effects on healthy grocery purchases; differences in local supermarket density explain no more than

about 1.5% of the income–nutrition gap.

- Consumers travel far for groceries; even residents of food deserts get most groceries from supermarkets, so new nearby stores mostly divert purchases from other supermarkets.

- The common "food desert" framing is misleading if it ignores shoppers' willingness to travel.

- Broader place effects (e.g., neighborhood influences observed via moves) explain no more than about 3% of the gap over several years.

- Equalizing supply conditions (prices/availability) would reduce the gap by no more than about 10%; most of the gap reflects differences in demand/preferences.

- A demand-side, means-tested subsidy for healthy groceries could raise low-income households' healthy eating to high-income levels at a cost of about 15% of the current SNAP budget.

- Therefore, policies focused on eliminating food deserts are unlikely to substantially reduce nutritional inequality; policy attention should shift toward demand-side tools (e.g., targeted subsidies), with further welfare analysis and testing.

- Quote snippet:

  policies aimed at eliminating food deserts likely generate little progress toward a goal of reducing nutritional inequality.

- Citation anchors:
  - page=47, words=0-349, section=VIII. CONCLUSION, note=Small impact of supermarket entry; 1.5% explained; travel behavior and diversion; reframing food deserts
  - page=48, words=0-323, section=VIII. CONCLUSION (cont.), note=Place effects 3%; equalizing supply 10%; means-tested subsidy ~15% of SNAP; limited gains from eliminating food deserts
  - page=22, words=0-349, section=Event study results, note=Bound showing supermarket access explains 1.5% of Health Index gap
  - page=41, words=0-349, section=Decomposition of supply vs. demand, note=Supply explains ~7–12% while demand explains ~88–93% of the relationship

E01: What are the most important prior papers cited and why are they central here?

- Category: `E) Citations and related literature`
- Confidence: `low (0.29819296174365445)`
- Retrieval method: `local`
- Evidence type: `direct quote and cited references`

- Data source: Allcott et al., QJE 2019, Food Deserts and the Causes of Nutritional Inequality
- Assumption flag: `False`
- Answer:

The pivotal prior work is Bitler and Haider (2011), which argues that much existing research assumes supply-side (retail access) causes food deserts and concludes there is insufficient evidence that food deserts systematically cause unhealthy diets. This frames the central question the paper tackles by testing place effects and supply-side explanations. In addition, a core set of earlier food-access studies (Alwitt & Donley 1997; Horowitz et al. 2004; Jetter & Cassady 2005; Algert, Agrawal & Lewis 2006; Baker et al. 2006; Powell et al. 2007; Larson, Story & Nelson 2009; Sharkey, Horel & Dean 2010) are central as they embody the supply-side perspective that this paper evaluates and ultimately finds limited for explaining nutritional inequality.

- Quote snippet:

  "it appears that much of the existing research implicitly assumes that supply-side factors cause any food deserts that exist."

- Citation anchors:
  - page=2, words=300-415, note=Bitler and Haider (2011) critique: assumption of supply-side causality and insufficient evidence
  - page=2, words=275-305, note=List of prominent prior food-access studies often cited in the literature
  - page=48, words=0-150, note=Authors' conclusion that place/supply-side policies (eliminating food deserts) yield little progress

E02: Which papers does this work most directly build on or extend?

- Category: `E) Citations and related literature`
- Confidence: `low (0.26231349582309116)`
- Retrieval method: `local`
- Evidence type: `direct textual statements in the article`
- Data source: Quarterly Journal of Economics article text (provided context)
- Assumption flag: `False`
- Answer:

Most directly, it subsumes and replaces Handbury, Rahkovsky, and Schnell (2015) and Allcott, Diamond, and Dubé (2017); it also extends migration-based studies such as Bronnenberg, Dubé, and Gentzkow (2012), Atkin (2016), Eid et al. (2008), and Finkelstein, Gentzkow, and Williams (2016, 2018a, 2018b) and Molitor (2018).

- Quote snippet:

  This article subsumes and replaces our previous work, Handbury,

Rahkovsky, and Schnell (2015) and Allcott, Diamond, and Dubé (2017).

- Citation anchors:
  - page=1, words=275-483, section=Front matter/notes, note=States the article subsumes and replaces Handbury et al. (2015) and Allcott et al. (2017).
  - page=6, words=275-429, section=Related literature/migration event study, note=Says the migration event study adds a nutritional aspect to prior migration-based research (Bronnenberg et al. 2012; Atkin 2016; Eid et al. 2008; Finkelstein et al. 2016, 2018a,b; Molitor 2018).

E03: Which papers are used as benchmarks or comparisons in the results?

- Category: `E) Citations and related literature`
- Confidence: `low (0.2675972693069443)`
- Retrieval method: `local`
- Evidence type: `textual`
- Answer:

DellaVigna and Gentzkow (2019) and Hausman (1996).

- Quote snippet:

  DellaVigna and Gentzkow (2019), for example, use price variation from individual stores' short-term promotions... Hausman (1996) uses variation in prices over time in other markets

- Citation anchors:
  - page=34, words=0-349, section=Instrument validity and comparisons, note=Paragraph contrasts their instrument with approaches in DellaVigna & Gentzkow (2019) and Hausman (1996).

E04: What data sources or datasets are cited and how are they used?

- Category: `E) Citations and related literature`

- Confidence: `low (0.318231478069409)`

- Retrieval method: `local`

- Evidence type: `textual evidence from article body`

- Data source: QJE article text (provided excerpts: pp. 1, 8–9)

- Assumption flag: `False`

- Answer:

- Nielsen Homescan household panel: used to measure household grocery purchases and demographics; focus on packaged goods, with a 2004–2006

8,000-household "magnet" subsample recording nonpackaged items to validate representativeness of packaged produce.

- Nielsen PanelViews (2017 add-on): collected health attitudes, a nutrition knowledge quiz, BMI, and diabetes status for Homescan households.

- U.S. Dietary Guidelines: used to map age- and gender-specific calorie needs to households and construct adult-equivalent household size and daily calorie need.

- Nielsen Retail Measurement Services (RMS): weekly UPC-level prices and sales for ~42,000 stores in 160 chains (2006–2016); used for store-by-year analyses; liquor stores excluded.

- ZIP Code Business Patterns (2004–2016): used to construct ZIP code establishment counts (508,951 ZIP-year observations).

- General Nutrition Knowledge Questionnaire (Kliemann et al. 2016): source of 28 questions used in the PanelViews survey.

- InfoUSA and National Establishment Time Series (NETS): noted as potentially useful for cross-sections but not reliable for store opening dates needed for entry analyses.

- Nielsen Company datasets via the Kilts Center for Marketing (UChicago Booth): primary source of scanner and marketing data underlying analyses.

- Quote snippet:

  The Nielsen RMS data consist of weekly prices and sales volumes for each UPC sold at approximately 42,000 unique stores from 160 retail chains for 2006–2016

- Citation anchors:
  - page=8, words=0-349, section=Data description (Homescan, PanelViews, Dietary Guidelines, RMS), note=Describes Homescan, magnet subsample, demographics, calorie needs via Dietary Guidelines, PanelViews survey, and RMS coverage/period.
  - page=9, words=275-389, section=Additional data sources, note=Mentions 28 GNKQ questions, ZIP Code Business Patterns counts, data collapse, and comments on InfoUSA/NETS not recording openings.
  - page=1, words=275-483, section=Acknowledgments/data provenance, note=States analyses are based in part on Nielsen Company data via the Kilts Center marketing databases.

E05: What methodological or econometric references are cited (e.g., DiD, IV, RDD methods)?

- Category: `E) Citations and related literature`

- Confidence: `medium (0.436374008181675)`
- Retrieval method: `local`
- Evidence type: `direct_citation`
- Assumption flag: `False`
- Answer:

Methodological/econometric references cited include: (1) DellaVigna and Gentzkow (2019) on using store-level price variation (residual demand vs consumer demand identification); (2) Hausman (1996) on cross-market price variation as an instrumental variables approach; (3) Method of Moments Estimation (VI.C.); (4) Gelbach (2016) on covariate contribution/decomposition; (5) Dubois, Griffith, and Nevo (2014) on structural demand using prices and attributes; (6) Kling, Liebman, and Katz (2007) on experimental analysis of neighborhood effects.

- Quote snippet:

  DellaVigna and Gentzkow (2019)... identify a store's residual demand elasticity... Hausman (1996) uses variation in prices over time in other markets. By contrast, our instrument generates cross-sectional identification...

- Citation anchors:
  - page=34, words=0-349, section=Instrument validity and comparisons, note=Main-text comparison to DellaVigna & Gentzkow (2019) and Hausman (1996), and discussion of instrument
  - page=34, words=275-385, section=VI.C. Method of Moments Estimation, note=Section header indicating Method of Moments Estimation
  - page=50, words=0-349, section=References, note=Lists DellaVigna & Gentzkow (2019); Dubois, Griffith, and Nevo (2014)
  - page=51, words=0-349, section=References, note=Lists Gelbach (2016); Hausman (1996); Kling, Liebman, and Katz (2007)

E06: Are there any seminal or classic references the paper positions itself against?

- Category: `E) Citations and related literature`
- Confidence: `low (0.2711114475094952)`
- Retrieval method: `local`
- Evidence type: `textual`
- Assumption flag: `False`
- Answer:

Yes. The paper pushes back against the classic "food desert" causal narrative—i.e., that supply-side access drives unhealthy eating—associated with early, frequently cited studies documenting access disparities (e.g., Alwitt & Donley 1997; Horowitz et al. 2004; Jetter & Cassady 2005; Algert, Agrawal, & Lewis 2006;

Baker et al. 2006; Powell et al. 2007; Larson, Story, & Nelson 2009; Sharkey, Horel, & Dean 2010) and related policy arguments that increasing healthy-food supply would meaningfully reduce nutritional inequality. The authors explicitly state their findings counter that argument.

- Quote snippet:

  These findings counter the argument that policies to increase the supply of healthy groceries could play an important role in reducing nutritional inequality.

- Citation anchors:
  - page=2, words=0-349, section=Introduction, note=States findings counter the argument that increasing healthy-food supply would reduce nutritional inequality and references the large literature on food deserts.
  - page=2, words=275-415, section=Introduction (footnotes/references), note=Lists classic food-desert studies (Alwitt & Donley 1997; Horowitz et al. 2004; Jetter & Cassady 2005; Algert et al. 2006; Baker et al. 2006; Powell et al. 2007; Larson et al. 2009; Sharkey et al. 2010) and contrasts with Bitler & Haider (2011).
  - page=48, words=0-323, section=Conclusion, note=Concludes that eliminating food deserts would generate little progress toward reducing nutritional inequality.

E07: Are there citations to code, data repositories, or appendices that are essential to the claims?

- Category: `E) Citations and related literature`
- Confidence: `low (0.3281837968846545)`
- Retrieval method: `local`
- Evidence type: `supplementary material and references`
- Data source: Harvard Dataverse (doi:10.7910/DVN/MSOBYI); Online Appendix; Nielsen Datasets at the Kilts Center
- Assumption flag: `False`
- Answer:

Yes. The article cites an Online Appendix and a Harvard Dataverse repository with replication code/data (doi:10.7910/DVN/MSOBYI), and notes use of Nielsen datasets via the Kilts Center.

- Quote snippet:

  Code replicating tables and fig- ures in this article can be found in Allcott et al. (2019), in the Harvard Dataverse, doi: 10.7910/DVN/MSOBYI.

- Citation anchors:

- page=48, words=0-323, section=Supplementary Material, note=Mentions Online Appendix and replication code in Harvard Dataverse with DOI.
- page=49, words=0-349, section=References, note=Lists the Harvard Dataverse replication data entry with DOI.
- page=1, words=275-483, section=Acknowledgments/Data, note=Describes use of Nielsen datasets via the Kilts Center data repository.

E08: What gaps in the literature do the authors say these citations leave open?

- Category: E) Citations and related literature
- Confidence: low (0.3164521631976586)
- Retrieval method: local
- Evidence type: direct quote from source
- Data source: QJE article text (Allcott et al. 2019)
- Assumption flag: False
- Answer:

The cited literature largely assumes supply-side access drives outcomes but leaves a key gap: there is not sufficient causal evidence that food deserts (neighborhood food environments) are systematically the cause of unhealthy eating among low-income households.

- Quote snippet:

  "we do not have sufficient evidence to determine whether food deserts are systematically the cause" of unhealthy eating by low-income people.

- Citation anchors:
  - page=2, words=275-415, note=Bitler and Haider (2011) noted insufficient evidence that food deserts systematically cause unhealthy eating; prior work often assumes supply-side causation.

F01: Are replication files or code provided? If so, where?

- Category: F) Replication and transparency
- Confidence: low (0.19556468707292782)
- Retrieval method: local
- Evidence type: text
- Data source: Harvard Dataverse DOI:10.7910/DVN/MSOBYI
- Assumption flag: False
- Answer:

Yes. Replication code and data are provided in the Harvard Dataverse (Allcott et al., 2019), DOI 10.7910/DVN/MSOBYI.

- Quote snippet:

  Code replicating tables and figures in this article can be found in Allcott et al. (2019), in the Harvard Dataverse, doi: 10.7910/DVN/MSOBYI.

- Citation anchors:
  - page=48, words=0-323, section=SUPPLEMENTARY MATERIAL, note=States that code replicating tables and figures is available with DOI.
  - page=49, words=0-349, section=References, note=Reference entry titled 'Replication Data...' with the same Harvard Dataverse DOI.

F02: Is there a pre-analysis plan or registered trial? Provide details if mentioned.

- Category: `F) Replication and transparency`
- Confidence: `low (0.24517154549579873)`
- Retrieval method: `local`
- Evidence type: `textual (no mention in provided excerpts)`
- Data source: QJE (2020) Food Deserts and Nutritional Inequality – provided excerpts
- Assumption flag: `False`
- Answer:

No. The provided excerpts do not mention any pre-analysis plan or registered trial.

- Quote snippet:

  An Online Appendix for this article can be found at The Quarterly Journal of Economics online. Code replicating tables and fig-ures in this article can be found in Allcott et al. (2019)

- Citation anchors:
  - page=48, words=200-320, section=SUPPLEMENTARY MATE-RIAL, note=Mentions online appendix and replication code; no preregistration or pre-analysis plan is indicated.

F03: Are data access constraints disclosed (restricted access, propri-etary data, NDAs)?

- Category: `F) Replication and transparency`
- Confidence: `low (0.2872801877515374)`
- Retrieval method: `local`
- Evidence type: `article text (acknowledgments and data description)`
- Data source: Nielsen datasets via Kilts Center (Homescan/RMS, TDLinx); USDA nutrition databases; Harvard Dataverse replication DOI:10.7910/DVN/MSOBYI

- Assumption flag: `True`
- Assumption notes: Proprietary/restricted access is inferred from the Nielsen/Kilts Center sourcing and standard disclaimer language; no explicit mention of access restrictions or NDAs in the provided text.
- Answer:

Partially. The paper discloses use of proprietary Nielsen datasets (Homescan/RMS and TDLinx) provided via the Kilts Center and includes required non-involvement disclaimers; it also cites a Harvard Dataverse replication package. However, explicit statements about restricted access terms or NDAs are not detailed.

- Quote snippet:

  data from the Nielsen Company (US), LLC, and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at the University of Chicago Booth

- Citation anchors:
  - page=1, words=275-483, section=Acknowledgments/Disclosure, note=Nielsen data source via Kilts Center; non-involvement disclaimer; USDA dissemination note.
  - page=10, words=0-200, section=Data description, note=Use of Nielsen's proprietary TDLinx data set.
  - page=49, words=0-349, section=References, note=Harvard Dataverse replication data citation (DOI provided).

F04: Are key steps in data cleaning and construction documented?

- Category: `F) Replication and transparency`
- Confidence: `low (0.2245506611258558)`
- Retrieval method: `local`
- Evidence type: `textual`
- Data source: Nielsen Homescan and RMS; Nielsen TDLinx; USDA FNDDS and SR; ZIP Code Business Patterns
- Table/Figure: Table I (Panels A and B)
- Assumption flag: `False`
- Answer:

Yes. The article documents multiple key steps in data cleaning and construction, including excluding liquor stores and nonpackaged items, deflating prices/incomes to 2010 dollars, collapsing data to household/store-by-year, validating and restricting supermarket entries (and dropping spurious ones), and detailing variable construction and weighting.

- Quote snippet:

  we use a list of specific TDLinx stores transferred through mergers

and acquisitions to exclude spurious "entrants" ... and we further drop potentially spurious entries

- Citation anchors:
  - page=8, words=275-393, section=Data description (RMS/Homescan), note=Excludes liquor stores; excludes bulk/nonpackaged items; deflates to 2010 dollars.
  - page=9, words=0-349, section=Descriptive statistics and construction notes, note=Weights for representativeness; construction of household size and normalized variables; collapsing to household/store-by-year.
  - page=10, words=0-349, section=Store entry measurement (TDLinx), note=Validation of opening dates; inclusion/exclusion criteria for store types.
  - page=10, words=275-424, section=Entry data cleaning, note=Excludes spurious entrants from M&A transfers; drops entries with same subtype in same census block in prior year.

F05: Are robustness and sensitivity analyses fully reported or partially omitted?

- Category: `F) Replication and transparency`
- Confidence: `low (0.28247126071816303)`
- Retrieval method: `local`
- Evidence type: `direct_text`
- Data source: Quarterly Journal of Economics article text (provided context)
- Table/Figure: Online Appendix Tables A4, A5, A9, A11
- Assumption flag: `False`
- Answer:

Fully reported

- Quote snippet:

As a robustness check, Online Appendix Table A11 repeats this decomposition using our alternative estimates...

- Citation anchors:
  - page=41, words=275-411, section=Main text, note=Explicit robustness check using alternative specifications (Table A11).
  - page=34, words=0-349, section=Main text, note=Two validity tests for the instrument; details in Online Appendix Table A9.
  - page=21, words=275-503, section=Main text, note=Additional estimates and sensitivity (Online Appendix Tables A4, A5); discusses anomaly and added controls.
  - page=38, words=275-377, section=Main text, note=Results robust across alternative controls; references Table A11.

   – page=7, words=0-349, section=Data/Overview, note=States all appendix material is available in the Online Appendix.

G01: What populations or settings are most likely to generalize from this study?

- Category: `G) External validity and generalization`
- Confidence: `low (0.29237732075827916)`
- Retrieval method: `local`
- Evidence type: `Study population and design scope`
- Data source: Nielsen Homescan Panel (2004–2016) and Nielsen Retail Scanner (RMS) data
- Table/Figure: Figure V
- Assumption flag: `False`
- Answer:

Findings most directly generalize to U.S. households' at-home, packaged grocery purchases in retail scanner–covered markets (RMS stores), analyzed at the ZIP code or county level over a medium/several-year horizon. They are less applicable to away-from-home food consumption, nonpackaged foods, or nationally representative mover populations.

- Quote snippet:

  We therefore focus on explaining income-related differences in the take-home market (i.e., grocery purchases) instead of overall diets.

- Citation anchors:
  - page=7, words=0-349, section=II.A. Nielsen Homescan and Retail Scanner Data, note=Homescan scope; packaged UPC foods; excludes away-from-home; focus on take-home market
  - page=24, words=0-349, section=II. DATA / Event study setup, note=Health Index defined for packaged groceries in RMS stores; geography is ZIP/county; movers not nationally representative; medium-term estimation
  - page=48, words=0-323, section=Conclusion, note=Effects assessed over the several-year time horizon allowed by the data

G02: What populations or settings are least likely to generalize?

- Category: `G) External validity and generalization`
- Confidence: `low (0.24059967921698125)`
- Retrieval method: `local`
- Evidence type: `textual evidence from article excerpts`
- Data source: Nielsen Homescan Panel (2004–2016)
- Assumption flag: `False`
- Answer:

The study's findings are least likely to generalize to: (1) away-from-home eating settings (e.g., restaurants), (2) nonpackaged foods without UPCs (e.g., bulk/fresh items), (3) nationally representative inferences without weighting, and (4) contexts involving very recent store entries or randomized neighborhood assignment (given imputation issues and nonrandom migration).

- Quote snippet:

  Homescan does not include data on food purchased away from home in establishments like restaurants. We therefore focus on explaining income-related differences in the take-home market

- Citation anchors:
  - page=7, words=0-349, section=II.A. Nielsen Homescan and Retail Scanner Data, note=Homescan excludes away-from-home food and mostly nonpackaged items.
  - page=21, words=275-503, section=Estimation notes, note=Observations are not weighted for national representativeness.
  - page=20, words=275-432, section=Store entry analysis, note=Imputation is unreliable for recently entered stores.
  - page=23, words=275-408, section=Migration/place effects, note=Moves are not randomly assigned; endogeneity concerns limit generalization to randomized settings.

G03: Do the authors discuss boundary conditions or scope limits?

- Category: `G) External validity and generalization`
- Confidence: `low (0.2423924986550409)`
- Retrieval method: `local`
- Evidence type: `direct textual evidence`
- Data source: Nielsen Homescan panel; Google Maps API (driving times)
- Assumption flag: `True`
- Assumption notes: Partial-equilibrium framework; possible upward bias from mover endogeneity; effects measured only over several years; supermarket exposure defined by 0–10/10–15 minute drive bands assuming no congestion; annual-only location reporting.
- Answer:

Yes. They bound their analysis to medium-term, partial-equilibrium effects over a several-year horizon; interpret estimates as likely upper bounds due to potential mover endogeneity; restrict exposure to supermarket entries within 0–10 and 10–15 minutes (assuming no congestion); and use annual location data because panel geography is only reported at year-end.

- Quote snippet:

  We therefore interpret the results in this section as likely upper bounds on true place effects.

- Citation anchors:
  - page=4, words=275-423, section=Movers event study; scope, note=Panel not long enough; bounds medium-term, partial-equilibrium effects of place.
  - page=24, words=0-349, section=Movers event study; limitations, note=Interpret results as likely upper bounds; annual location reporting; medium-term only.
  - page=17, words=0-349, section=Supermarket entry design, note=Define exposure within 0–10/10–15 minutes and assume no congestion delay.
  - page=48, words=0-323, section=Conclusion/scope, note=Effects assessed over the several-year time horizon allowed by the data; partial-equilibrium attribution.

G04: How might the results change in different time periods or markets?

- Category: `G) External validity and generalization`
- Confidence: `medium (0.35904641697863765)`
- Retrieval method: `local`
- Evidence type: `direct text`
- Data source: Nielsen Homescan and RMS
- Table/Figure: Figure IV (Event Study of Supermarket Entry)
- Assumption flag: `False`
- Answer:

They are largely identified within market–year cells, so period-wide shocks are differenced out, while cross-market differences can matter. In particular, results can differ by geography because retailers' presence and relative pricing vary across markets (e.g., produce cheaper on the West Coast, costlier on the East Coast). Effects also differ by market type: expenditure share changes are larger for low-income and food-desert areas. For timing, the movers design captures medium-term changes (a few years after a move) and is likely an upper bound on true place effects.

- Quote snippet:

  identification comes only from variation in the relative prices across product groups within a market in a given year.

- Citation anchors:
  - page=32, words=0-349, section=Instrument and identification, note=Identification comes from relative prices across product groups within a market-year.
  - page=32, words=275-431, section=Instrument and identification, note=Produce predicted cheap on West Coast and expensive on East Coast, implying geographic differences.

- page=24, words=0-349, section=Movers event study, note=Estimates reflect medium-term effects after moves; authors view them as likely upper bounds due to possible endogeneity.
- page=21, words=275-503, section=Heterogeneity in effects, note=Expenditure share changes are larger for low-income households and in food deserts.
- page=19, words=0-310, section=Event Study of Supermarket Entry, note=2004–2016 event study; separate full-sample vs. food-desert subsample panels.

H01: Are key variables measured directly or via proxies?

- Category: H) Measurement validity
- Confidence: medium (0.3645774780007952)
- Retrieval method: local
- Evidence type: mixed (direct measures and constructed proxies/instruments)
- Data source: Nielsen Homescan and RMS
- Assumption flag: False
- Answer:

Mixed. Many variables are directly observed or self-reported (e.g., demographics, income bins, purchases/prices, survey measures), while key constructs are proxies/constructed (e.g., adult-equivalent household size and daily calorie need from guidelines; an IV based on out-of-market chain pricing to proxy cost advantages; residualized income quartiles; unobserved product characteristics).

- Quote snippet:

  Our household size variable measures the number of adult "equivalents" in the household, where children are scaled into adults by their calorie needs.

- Citation anchors:
  - page=8, words=0-349, section=Data description (Homescan), note=Direct self-reports of demographics; construction of daily calorie need and adult-equivalent household size; additional survey measures.
  - page=30, words=0-349, section=Demand model, note=Constructed variables (price per calorie, characteristics per calorie) and inclusion of an unobserved product characteristic .
  - page=30, words=275-393, section=Price endogeneity, note=Motivates need for an instrument (proxy) due to correlation between prices and preferences.
  - page=31, words=275-432, section=IV construction (part 1), note=Defines IV components using out-of-market chain prices to capture cost advantages, excluding local demand.
  - page=32, words=0-349, section=IV construction (part 2), note=Formal definition of the price instrument Pjmt as a weighted

average cost advantage across UPCs.

- – page=34, words=0-349, section=Instrument validity checks, note=Validates IV is unrelated to predictable tastes; describes use across applications.
- – page=34, words=275-385, section=Estimation setup, note=Constructs income quartiles after residualizing by household size/age/year—another constructed measure.
- – page=42, words=275-356, section=Controls and survey measures, note=Direct measures of demographics and additional survey variables used as controls.

H02: What measurement error risks are acknowledged or likely?

- Category: `H) Measurement validity`
- Confidence: `low (0.2590133017456877)`
- Retrieval method: `local`
- Evidence type: `textual evidence with minor inference`
- Data source: Nielsen Homescan and RMS; ZBP
- Assumption flag: `True`
- Assumption notes: Self-reported BMI/diabetes/attitudes/knowledge likely contain reporting error; income in 16 bins implies coarseness. These risks are not explicitly labeled as 'measurement error' but follow from the described data features.
- Answer:

Acknowledged/likely measurement error risks include: (1) Undercoverage of non-packaged/bulk items, since both Homescan and RMS exclude bulk produce and other nonpackaged goods (mitigated partly by evidence that ~60% of produce calories among magnet households are packaged). (2) RMS is a sample, not a census; the store-type mix in RMS may not match a county's true distribution, requiring regression adjustment with ZBP and leaving scope for residual area-level measurement error in the Health Index. (3) Household residence is observed only at year-end, so exposure around moves is measured annually; within-year moves and local shopping patterns can cause misclassification, mitigated via a 50% local-shopping restriction. (4) Self-reported health-related variables (e.g., BMI, diabetes, health importance, nutrition knowledge) are likely subject to reporting/recall error. (5) Income is recorded in 16 bins, implying coarse measurement of income.

- Quote snippet:

  RMS does not include sales of bulk produce and other nonpackaged items.

- Citation anchors:
  - – page=8, words=0-349, section=Data coverage and demographics, note=Homescan excludes nonpackaged items; magnet sample shows 60% of produce calories packaged; income recorded in 16 bins;

presence of self-reports.

- – page=8, words=275-393, section=RMS coverage, note=RMS excludes bulk produce and other nonpackaged items.
- – page=24, words=275-452, section=RMS store-type composition vs true distribution, note=RMS is not a complete census; adjust area Health Index for store-type distribution using ZBP.
- – page=24, words=0-349, section=Location timing, note=Household location observed only at end of each year; event study uses annual data around moves.
- – page=25, words=275-468, section=Local shopping restriction, note=50% local shopping restriction to mitigate misclassification around moves.

H03: Are there validation checks for key measures?

- Category: `H) Measurement validity`
- Confidence: `low (0.26944750917915244)`
- Retrieval method: `local`
- Evidence type: `reported statistical validation and falsification tests`
- Data source: Nielsen Homescan
- Assumption flag: `False`
- Answer:

Yes. They validate the key dietary-quality measure by showing the linearized Health Index closely matches the standard HEI (r=0.91) and that both correlate with BMI and diabetes. They also conduct falsification tests for their instrument, finding no association with predictable tastes or differential patterns across healthy product groups by income.

- Quote snippet:

  the correlation between "true" HEI and linearized HEI is 0.91 in our household-by-year Homescan data.

- Citation anchors:
  - – page=11, words=0-349, section=Methods for Health Index (HEI linearization), note=Reports r=0.91 between true HEI and linearized HEI; correlations with BMI and diabetes
  - – page=34, words=0-349, section=Instrument validity checks, note=Instrument not associated with predicted purchases; no systematic differences across healthy categories by income

H04: Do the authors discuss construct validity for core outcomes?

- Category: `H) Measurement validity`
- Confidence: `low (0.2371564135213259)`
- Retrieval method: `local`

54

- Evidence type: `Within-study correlational (convergent) validity evidence`
- Data source: Nielsen Homescan household-by-year data
- Table/Figure: Online Appendix Table A2
- Assumption flag: `False`
- Answer:

Yes. They provide construct validity evidence for their core outcome (the Health Index) by showing it strongly correlates with the original HEI (r = 0.91) and that both the "true" and linearized HEI are highly correlated with BMI and diabetes status.

- Quote snippet:

  the correlation between "true" HEI and linearized HEI is 0.91… both the "true" HEI and our linearized HEI are highly correlated with Homescan panelists' BMI and diabetes status.

- Citation anchors:
  - page=11, words=0-349, section=Description/validation of Health Index (HEI), note=Reports 0.91 correlation with true HEI; notes correlations with BMI and diabetes.
  - page=11, words=275-399, section=III.A. Purchase Disparities: The Nutrition–Income Relationship, note=States both true and linearized HEI are highly correlated with BMI and diabetes status.

I01: What policy counterfactuals are considered or implied?

- Category: `I) Policy counterfactuals and welfare`
- Confidence: `low (0.3292589015742387)`
- Retrieval method: `local`
- Evidence type: `text`
- Table/Figure: Figure VI
- Assumption flag: `False`
- Answer:

The paper considers or implies several policy counterfactuals: (1) a partial-equilibrium 'place effect' move—shifting a household from a low- to a high-income retail environment; (2) supply-side equalization—setting all groups' prices to those of high-income areas and also equalizing product nutrient characteristics (eliminating food-desert supply differences); (3) demand-side preference counterfactuals—imposing high-income nutrient and product-group preferences on other groups; (4) encouraging supermarket entry (supply-side policy); and (5) subsidies for healthy foods—implemented as a simple subsidy that scales with product healthfulness, including a means-tested healthy-grocery subsidy (estimated to close the gap at about 15% of the SNAP budget).

- Quote snippet:

55

We focus on a simple subsidy that scales in a product's healthfulness.

- Citation anchors:
  - page=27, words=275-482, section=Partial-equilibrium place effect thought experiment, note=Move from low- to high-income retail environment
  - page=40, words=0-285, section=Figure VI counterfactuals, note=Equalizing prices and product nutrient characteristics; setting preferences to high-income group
  - page=43, words=275-387, section=VII.C Using Subsidies, note=Supply-side policy of supermarket entry; healthy-food subsidies scaling with healthfulness
  - page=48, words=0-323, section=Conclusion, note=Means-tested subsidy could close gap at ~15% of SNAP; equalizing supply has limited effect

I02: What are the main welfare tradeoffs or distributional impacts discussed?

- Category: I) Policy counterfactuals and welfare
- Confidence: medium (0.3951936988208884)
- Retrieval method: local
- Evidence type: Empirical event-study plus structural model counterfactuals and cost calculations
- Data source: Allcott et al. (2019), The Quarterly Journal of Economics
- Table/Figure: Table V: Impacts of Means-Tested Healthy Grocery Subsidies
- Answer:

The paper highlights a key welfare tradeoff between supply-side store-entry policies and targeted consumer subsidies. Equalizing supply (e.g., eliminating food deserts) reduces nutritional inequality by at most about 10%, implying limited welfare gains from costly store subsidies. In contrast, a means-tested subsidy tied to product healthfulness and targeted to bottom-quartile households can achieve the same or greater improvements at transparent costs: about $84 million to match the effect of one nearby supermarket entry, $830 million to match full supply equalization, and $11 billion (15% of the SNAP budget) to close the entire gap. The authors caution that healthy eating is not the sole social objective and government outlays are not a complete measure of social costs, so a full welfare framework is needed, but distributionally, targeted subsidies are more cost-effective at improving low-income households' healthy grocery consumption than supply-side interventions.

- Quote snippet:

  an annual subsidy of $84 million would increase healthy eating by the same amount as one additional supermarket entry

- Citation anchors:
  - page=48, words=0-323, section=Conclusion, note=Equalizing supply reduces gap by no more than ~10%; means-tested subsidy could close gap at ~15% of SNAP
  - page=46, words=0-349, section=VII.C, note=$84M matches supermarket entry; $830M matches full supply equalization; $11B closes gap
  - page=45, words=0-349, section=VII.C, Table V, note=Table V quantifies subsidy parameters and total payments to bottom-quartile households
  - page=43, words=275-387, section=VII.C, note=Introduces using subsidies for healthy foods as alternative to supply-side policies
  - page=2, words=0-349, section=Introduction, note=Counterfactual: equalizing products/prices cuts inequality ~10%; 90% due to demand, challenging supply-side focus

I03: Are cost-benefit or incidence analyses provided?

- Category: I) `Policy counterfactuals and welfare`
- Confidence: `low (0.29904970035413087)`
- Retrieval method: `local`
- Evidence type: `text`
- Data source: Allcott et al. (2020), The Quarterly Journal of Economics
- Assumption flag: `False`
- Answer:

No. The excerpt does not provide a full cost-benefit or incidence analysis; it offers cost-effectiveness comparisons (e.g., required subsidy amounts) and explicitly notes that a welfare/optimal policy analysis would be needed before advocating policies.

- Quote snippet:

  Before advocating for or against such a subsidy, one would need to measure the relevant market failures and study optimal policy in a principled welfare maximization framework.

- Citation anchors:
  - page=48, words=0-323, section=Main text, note=States that optimal policy/welfare analysis would be needed before advocating a subsidy, implying no full cost-benefit/incidence analysis is provided.
  - page=46, words=0-349, section=Main text, note=Provides cost-effectiveness comparisons (e.g., $84M, $830M, $11B; statements about being cost-effective), not a full cost-benefit or incidence analysis.

I04: What policy recommendations are stated or implied?

- Category: I) Policy counterfactuals and welfare
- Confidence: low (0.3321727108993598)
- Retrieval method: local
- Evidence type: direct textual evidence
- Assumption flag: True
- Assumption notes: Model-based subsidy results are suggestive and depend on assumptions; authors recommend randomized demonstration projects to assess effects and note economic/political considerations around SNAP modifications.
- Answer:

Shift policy away from supply-side efforts (e.g., subsidizing supermarket entry or focusing on local access) toward demand-side, means-tested subsidies for healthy foods—potentially scaled by product healthfulness—and test such approaches (including SNAP-related designs) via randomized demonstration projects. Additionally, improving education and nutrition information is suggested as a way to reduce nutritional inequality.

- Quote snippet:

  policy makers interested in reducing nutritional inequality might redirect efforts away from promoting access to healthy groceries and toward means-tested subsidies.

- Citation anchors:
    - page=47, words=0-349, section=Main text, note=Recommends redirecting efforts toward means-tested subsidies; calls for demonstration projects and notes SNAP considerations.
    - page=43, words=0-349, section=Main text, note=Education and nutrition knowledge explain part of inequality; suggests roles for improving education and nutrition information.
    - page=43, words=275-387, section=Main text, note=Introduces alternative policy: subsidies for healthy foods, scaled by product healthfulness.
    - page=46, words=0-349, section=Main text, note=Indicates means-tested subsidies are more cost-effective than supermarket entry subsidies at increasing healthy purchases.

J01: What missingness or attrition patterns are reported?

- Category: J) Data quality and integrity
- Confidence: low (0.3176327986938804)
- Retrieval method: local
- Evidence type: text
- Data source: Nielsen Homescan Panel (2004–2016)
- Assumption flag: False

- Answer:

Two patterns are reported: (1) 10.6% of household–product group–year observations are dropped due to zero purchases; "Baby food" has the most missing, and the missingness shares across income groups are not correlated with product-group characteristics. (2) The Homescan data omit away-from-home food and mostly exclude nonpackaged/bulk groceries, so those items are missing from coverage.

- Quote snippet:

  We drop 10.6% of observations at the household-by–product group–by-year level because they have zero purchases. "Baby food" is the product group with the most missing observations.

- Citation anchors:
  - page=35, words=275-428, section=VI.C. Method of Moments Estimation, note=Reports 10.6% dropped observations, baby food most missing, and lack of correlation of missingness with characteristics across income groups.
  - page=7, words=0-349, section=II.A. Nielsen Homescan and Retail Scanner Data, note=Describes missing coverage for away-from-home consumption and mostly nonpackaged/bulk groceries.

J02: How are outliers handled (winsorization, trimming, exclusions)?
- Category: `J) Data quality and integrity`
- Confidence: `low (0.33104689525725783)`
- Retrieval method: `local`
- Evidence type: `textual`
- Data source: Nielsen Homescan and RMS, as described in the provided context
- Assumption flag: `False`
- Answer:

Outlier handling via winsorization or trimming is not described in the provided context. The only stated handling is sample exclusions/restrictions: they exclude liquor stores and nonpackaged/bulk items from RMS/Homescan coverage, and for mover analyses they use balanced panels and a 50% local shopping restriction. No explicit winsorizing or trimming of variables is mentioned.

- Quote snippet:

  We exclude liquor stores... RMS does not include sales of bulk produce and other nonpackaged items.

- Citation anchors:
  - page=8, words=275-393, section=Data description, note=Excludes liquor stores; RMS does not include bulk/nonpackaged items

- page=24, words=275-452, section=Mover design setup, note=Defines balanced panel around a move
- page=25, words=275-468, section=Mover design restriction, note=50% local shopping restriction to exclude atypical movers

J03: Are there data audits or validation steps described?

- Category: J) Data quality and integrity
- Confidence: low (0.23155706377695187)
- Retrieval method: local
- Evidence type: textual evidence
- Data source: Nielsen TDLinx validated against retailer administrative records; Nielsen Homescan magnet subsample
- Table/Figure: Online Appendix Figure A1 (mentioned)
- Assumption flag: False
- Answer:

Yes. They validate TDLinx store openings against retailer administrative records, exclude spurious 'entrants' using M&A transfer lists and same-block checks, assess packaged-goods representativeness with a magnet subsample, and run instrument validity tests.

- Quote snippet:

  In validation checks, we found that TDLinx data closely match entry dates and loca- tions for 1,914 supermarkets ... from four retailers' administrative records.

- Citation anchors:
  - page=10, words=0-349, section=II.B. Grocery Retail Establishments, note=Validation of TDLinx opening dates/locations against retailer administrative records
  - page=10, words=275-424, section=II.B. Grocery Retail Establishments, note=Procedures to exclude spurious entrants (M&A transfers; same-block prior-year stores)
  - page=8, words=0-349, section=Data and representativeness, note=Magnet subsample used to check representativeness of focusing on packaged groceries
  - page=34, words=0-349, section=Instrument validity tests, note=Two tests showing instrument unrelated to predictable tastes

J04: Is there evidence of reporting bias or selective sample inclusion?

- Category: J) Data quality and integrity
- Confidence: low (0.295750837001939)
- Retrieval method: local

- Evidence type: `Explicit methodological limitations and sample representativeness statements in the text`
- Data source: Allcott et al. (2019), The Quarterly Journal of Economics: Food Deserts and Nutritional Inequality
- Table/Figure: Figure V
- Assumption flag: `False`
- Answer:

Yes. The authors explicitly note reporting and selection issues: Homescan largely captures only packaged UPC items and excludes away-from-home food; some analyses use movers who are not nationally representative and are not weighted; and remaining endogeneity may bias results upward.

- Quote snippet:

  Because the set of movers is not nationally representative, we again do not use the Homescan sample weights for this analysis.

- Citation anchors:
  - page=7, words=120-260, section=II.A. Nielsen Homescan and Retail Scanner Data, note=Homescan excludes away-from-home purchases and mostly nonpackaged groceries; potential reporting limitation.
  - page=13, words=275-405, note=Results are similar in a 'magnet' sub-sample that includes bulk purchases, addressing concerns about missing nonpackaged items.
  - page=21, words=275-320, note=Observations are not weighted for national representativeness.
  - page=24, words=0-200, note=Movers are not nationally representative; any remaining endogeneity likely biases estimates toward finding place effects.

K01: What goodness-of-fit or diagnostic metrics are reported?

- Category: `K) Model fit and diagnostics`
- Confidence: `low (0.2633273068183583)`
- Retrieval method: `local`
- Evidence type: `Reported statistics in tables and text (correlations, SEs, CIs, significance, R2 comment)`
- Data source: Nielsen Homescan 2004–2016 (as analyzed in the article's tables/notes)
- Table/Figure: Table III; Table IV
- Answer:

They report correlation coefficients ( ) for HEI components, robust clustered standard errors, significance stars (10%, 5%, 1%), 95% confidence interval upper bounds, and note that regression R2 changes only slightly with added controls.

- Quote snippet:

95% confidence interval upper bound 0.100 0.097 0.217 0.220 ... Robust standard errors, clustered by household and local area (ZIP code or county), are in parentheses.

- Citation anchors:
  - page=11, words=275-399, section=Section III.A text, note=Reports correlation coefficients ( ) for dietary components and HEI validity.
  - page=27, words=0-349, section=Table III notes and surrounding text, note=Shows 95% CI upper bounds, clustered robust SEs, significance stars; mentions R2 changes slightly.
  - page=21, words=275-503, section=Table II notes, note=States robust clustered SEs and significance star levels.

K02: Are functional form choices tested (logs, levels, nonlinearities)?

- Category: K) Model fit and diagnostics
- Confidence: low (0.2986677962411813)
- Retrieval method: local
- Evidence type: textual (model description)
- Data source: Article text (pp. 30, Sections VI.A–VI.B)
- Answer:

Not in the provided context. The paper specifies a log-linear demand equation derived from Cobb–Douglas/CES preferences but does not report tests comparing logs vs levels or other nonlinear forms.

- Quote snippet:

  ln Yijt = − ln p̃ijt − ˜c ãijct − + j + m + t + ijt

- Citation anchors:
  - page=30, words=0-349, section=VI.A Empirical Model, note=Presents log-linear demand equation (7); no mention of functional-form tests.
  - page=30, words=275-393, section=VI.A/VI.B, note=Discusses implications of Cobb–Douglas form; does not indicate testing alternative forms.

K03: Are residual checks or specification tests reported?

- Category: K) Model fit and diagnostics
- Confidence: low (0.2418967000364113)
- Retrieval method: local
- Evidence type: textual evidence
- Table/Figure: Online Appendix Table A9; Online Appendix Table A11
- Assumption flag: False
- Answer:

Yes. The paper reports multiple specification and robustness checks, including tests of instrument validity, checks for pre-trends around moves, and robustness

to adding alternative controls (e.g., county-by-quarter fixed effects, convenience-store entry).

- Quote snippet:

  Adding county-by-quarter fixed effects or a control for convenience store entry does not qualitatively change the results.

- Citation anchors:
  - page=21, words=275-503, section=Entry effects robustness, note=Adding county-by-quarter fixed effects or a control for convenience store entry does not change results.
  - page=25, words=275-468, section=Move analysis pre-trend check, note=No statistically significant evidence of premove trends in Health Index after moves.
  - page=34, words=0-349, section=Instrument validity tests, note=Two tests show the instrument is not associated with predicted purchases and not systematically different across income-related health groups.
  - page=38, words=275-377, section=Additional robustness, note=Estimates are very similar with alternative controls; supports exclusion restriction robustness.

K04: How sensitive are results to alternative specifications or estimators?

- Category: K) Model fit and diagnostics
- Confidence: medium (0.4193826056196361)
- Retrieval method: local
- Evidence type: text
- Data source: Provided context excerpts (pages 21, 27, 38)
- Assumption flag: False
- Answer:

Results appear robust. Including demographic controls has very little impact; adding county-by-quarter fixed effects or a control for convenience-store entry does not qualitatively change results; re-estimating with extensive census-region-by–product-group (and urban/rural, high-income) fixed effects yields very similar estimates. Restricting to supercenter entry increases expenditure-share effects, but Health Index effects remain economically small and mostly insignificant. Adding shelf-life and convenience characteristics lowers WTP magnitudes but preserves the qualitative preference patterns.

- Quote snippet:

  Adding county-by-quarter fixed effects or a control for convenience store entry does not qualitatively change the results.

- Citation anchors:

- page=27, words=275-482, section=Robustness (demographic controls), note=Including household demographics has very little impact on results.
- page=21, words=275-503, section=Robustness (fixed effects, alternative entry definition), note=County-by-quarter FE and convenience-store entry control do not change results; supercenter-only effects larger for shares but Health Index effects small/insignificant.
- page=38, words=275-377, section=Robustness (additional fixed effects), note=Re-estimating with added region-by–product-group and interactions yields very similar estimates.
- page=38, words=0-349, section=Robustness (adding characteristics), note=Including shelf life and convenience lowers WTP levels but leaves preference patterns similar.

P01: ResponseTextConfig(format=ResponseFormatText(type='text'), verbosity='medium')

- Category: P) Previous questions

- Confidence: n/a

- Retrieval method: n/a

- Evidence type: n/a

- Answer:

- Data sources and coverage

  - Homescan records packaged groceries; in 2004–2006, an 8,000-household "magnet" subsample also recorded nonpackaged items. About 60% of magnet households' produce calories are from packaged goods, with no statistical income differences, suggesting packaged focus is reasonably representative for produce and that produce is a small share of groceries overall (page 8, words 0–349).
  - Households report income (16 bins), composition, demographics; researchers compute daily calorie needs using Dietary Guidelines and scale children to adult equivalents. PanelViews add-on (2017) provides health importance, nutrition quiz, BMI, diabetes; results are nationally weighted (page 8, words 0–349).
  - RMS: weekly UPC-level prices/volumes at ~42,000 stores from 160 chains (35,000/year observed) for 2006–2016; excludes liquor and bulk produce. Channel coverage: 53% grocery, 32% mass merchandiser, 55% drug, 2% convenience. Prices/incomes deflated to 2010 dollars using CPI-U (page 8, words 275–393).

- Identification and market structure facts

  - Substantial geographic variation in the presence of the five largest retailers; these chains set different relative prices across product groups,

yielding geographic price instruments. Example: produce predicted cheap on West Coast and expensive on East Coast, likely due to California production; market fixed effects address endogeneity from demand elasticity differences across markets (page 32, words 275–431).

- Within-household regressions: adding demographic controls changes results little and increases R^2 only slightly, implying possible unobserved within-household changes (Oster 2019). A migration-style thought experiment considers moving a household from low- to high-income retail environments (partial equilibrium) (page 27, words 275–482).
- Validation: replicates immediate brand choice migration effect for Coke vs. Pepsi withˆ   0.16; a 10 pp higher county Coke share raises a household's Coke share by 1.6 pp (page 27, words 275–482).

- Preference estimates (Table IV) and WTP for health

  - GMM estimates of nutrient preference parameters by income quartile show:
    * Higher-income quartiles have more negative WTP for sodium and added sugar, and more positive WTP for whole fruit and whole grains; patterns are monotonic across quartiles (page 36, words 0–349).
    * "WTP for Health Index" rises with income: Q1   0.429, Q2   0.631, Q3   0.820, Q4   1.141 (SEs shown in table) (page 36, words 0–349).
  - Units: sodium (g), whole/other fruit and dairy (cups), whole/refined grains and proteins (oz), added sugar (tsp), solid fats (calories). WTP for Health Index defined as $\sum_c \hat{\beta}_c s_c r_c$ (page 36, words 0–349).

- Decomposition: supply vs. demand in nutrition–income gaps

  - Across alternative specifications (including census-region-by–product group fixed effects and interactions with urban/rural and above-median income), supply explains 7%–12% and demand explains 88%–93% of the nutrition–income relationship; summary: about 90% demand, 10% supply (page 41, words 275–411; page 38, words 275–377).
  - The unobserved characteristic $\hat{\xi}$ is attributed to demand because it primarily determines demand elasticity with respect to price (page 41, words 275–411).

- Demographic controls used in outcomes regressions

  - Controls include household size, age indicators, log years of education, race, marital status, employment, weekly work hours, plus add-on variables such as the self-reported importance of staying healthy (page 42, words 275–356; page 8, words 0–349).