Audit Report: Workflow e53e1c45434547afaedf7a0d079d5615

Overview

- Source JSON: reports\workflow-report-e53e1c45434547afaedf7a0d079d5615.json
- Run ID: e53e1c45434547afaedf7a0d079d5615
- Papers input: papers\alcott food deserts.pdf
- Started at: 2026-02-15T21:03:21.277446+00:00
- Finished at: 2026-02-15T21:10:07.499430+00:00
- Duration: 0:06:46.221984

Effective Configuration

- Chat model: gpt-5-nano
- 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: failed
- report_store: stored

Agentic Summary

- Status: completed
- Main question: What is the key contribution?
- Report question set: both
- Structured questions generated: 84
- Confidence mean/median: 0.3042301084348596 / 0.29929904275261066
- Confidence labels: low=65, medium=18, high=0

Final Answer

- Key contribution: A robust decomposition of the nutrition–income gap in healthy eating into supply-side (prices, availability) and demand-side (preferences) factors, showing demand dominates. Across specifications, supply explains ~7–12% and demand ~88–93% of the gap (roughly 90% to 10%). (Context: Main empirical goal; page 41, 275–411; page 27, 275–482)

- Data foundation: Uses a rich, nationally representative combination of Homescan panel data (demographics, calories, health measures) and Nielsen RMS price/sales data (2006–2016; ~35,000 stores per year; excludes liquor; channels: grocery, mass, drug, convenience). All results weighted; Nielsen data treated as researchers' data (Nielsen not involved in analysis). (Context: Data and samples used; page 1, 275–483; page 8, 0–349; page 8, 275–393)
- Price treatment: RMS prices deflated to 2010 USD; excludes bulk produce and other nonpackaged items. (Context: page 8, 0–349; page 8, 275–393)
- Identification strategy and endogeneity bounds: Combines a partial-equilibrium thought experiment with geographic variation to bound location effects; online Appendix maps retailer presence; produce prices predicted to be cheaper on the West Coast and more expensive on the East Coast, addressing endogeneity via geographic variation. (Context: page 27, 275–482; page 32, 275–431)
- Instrumental/identification details: Constructs a multi-faceted instrument from geographic price variation across the five largest RMS retailers; Appendix figures illustrate these variations and their identification implications. (Context: page 32, 275–431)
- Illustrative empirical result (Coke vs. Pepsi): County-level Coke market-share exercise yields^ 0.16; a 10 percentage-point higher Coke share is associated with a 1.6 percentage-point higher Coke share within Coke+Pepsi purchases. (Context: page 27, 275–482)
- Robustness and interpretation notes: Conclusions are from Nielsen-based analyses with multiple robustness checks (alternative specifications and fixed effects); Nielsen explicitly did not influence the analysis. (Context: page 1, 275–483)
- Illustrative nutrient-preference evidence: Table IV reports nutrient preferences by income quartile (including units and interpretation, e.g., WTP for a health index). (Context: page 36, 0–349; page 38, 275–377)

If you'd like, I can extract a concise formal summary with a minimal set of equations or tailor the bullets to a specific subsection (data, estimation, or robustness results).

Sub-Answers

Sub-answer 1

- Question: ResponseTextConfig(format=ResponseFormatText(type='text'), verbosity='medium')
- Question tokens estimate: `19`
- Answer:

- Data and samples used
  - Homescan panel data (demographics, calories, health measures) and Nielsen RMS price/sales data (2006–2016; about 35,000 stores observed per year; exclude liquor; channels: grocery, mass merchandiser, drug, convenience). All Homescan results are weighted for national representativeness. Sources and caveats noted (Nielsen data are the researchers' data and Nielsen had no role in analysis). (Context: page 1, words 275–483; page 8, words 0–349; page 8, 275–393)
- How prices are treated
  - Prices from RMS are deflated to 2010 USD; RMS excludes bulk produce and other nonpackaged items. (Context: page 8, 0–349; page 8, 275–393)
- Main empirical goal
  - Decompose differences in healthy eating across income groups into supply-side (prices, availability) and demand-side (preferences) factors. Robustness checks with census-region-by-product fixed effects and urban/rural interactions. (Context: page 41, 275–411; page 27, 275–482)
- Key result: share of nutrition–income gap explained by demand vs supply
  - Across specifications, supply explains roughly 7%–12% and demand explains roughly 88%–93% of the nutrition–income relationship; about 90% is due to demand, ~10% to supply. (Context: page 41, 275–411)
- Methodological approach to endogeneity and location
  - Use a partial equilibrium thought experiment and geographic variation to bound the role of location. Online Appendix figures map geographic presence of retailers; produce prices are predicted to be cheap on the West Coast and expensive on the East Coast (addressing endogeneity via geographic/instrumental variation). (Context: page 27, 275–482; page 32, 275–431)
- Illustrative empirical result (Coke vs. Pepsi)
  - An empirical exercise using county Coke market shares shows a statistically significant^ 0.16: moving to a county with a 10 percentage point higher Coke share is associated with a 1.6 percentage point increase in Coke's share within Coke+Pepsi purchases. (Context: page 27, 275–482)
- Instrumental/identification notes
  - The study constructs a multi-faceted instrument based on geographic price variation across major retailers (the five largest RMS retailers),

with online Appendix figures illustrating these variations and their implication for identification. (Context: page 32, 275–431)

- Notes on data interpretation and scope
  - The conclusions reflect the researchers' analyses of Nielsen data and are not Nielsen's statements; results are presented with several robustness checks, including alternative specifications and fixed effects. (Context: page 1, 275–483)
- Example of reported nutrient preferences by income (illustrative table)
  - Table IV shows estimated nutrient preferences by income quartile across several nutrients and health indices; units and interpretation are provided (e.g., WTP for health index). (Context: page 36, 0–349; page 38, 275–377)

If you'd like, I can extract a concise formal summary (with equations) or tailor the bullets to a specific subsection (e.g., data, estimation, or robustness results).

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 (0.33069776207848245)`
- Retrieval method: `local`
- Evidence type: `Intro/abstract statement from the paper`
- Data source: Nielsen Homescan data
- Answer:

To explain the nutrition–income gap: why the wealthy eat more healthfully than the poor in the United States and whether neighborhood environments (food deserts) account for it.

- 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=0-40, section=Introduction, note=States the main research question about nutritional inequality and why the wealthy eat more healthfully.

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

- Category: `A) Research question / contribution`

- Confidence: `low (0.3459747648478517)`
- Retrieval method: `local`
- Evidence type: `Empirical finding using Homescan/Nielsen data and a novel price-based instrument across markets to separate supply and demand effects on nutrition.`
- Data source: Homescan data; Nielsen Datasets at the Kilts Center (Nielsen Company data)
- Answer:

The paper's primary contribution is showing that supply-side factors (food deserts / local supermarket access) explain only a small share of nutritional inequality, while demand differences across income groups drive the bulk of the gap; using a novel market-level price instrument, the authors argue that policy aimed at increasing healthy-s grocery supply would have limited impact compared to demand-side measures.

- Quote snippet:

  > exposing low-income households to the same products and prices available to high-income households reduces nutritional inequality by only about 10%, while the remaining 90% is driven by differences in demand.

- Citation anchors:
  - page=1, words=0-80, section=I. INTRODUCTION, note=States the main contribution: questioning the role of neighborhood supply in nutritional inequality.
  - page=6, words=0-60, section=IV–VI (Supply vs. Demand discussion), note=Documents the 10% (supply) vs 90% (demand) decomposition of the nutrition–income relationship.

A03: What is the central hypothesis being tested?

- Category: `A) Research question / contribution`
- Confidence: `low (0.2907905449325526)`
- Retrieval method: `local`
- Evidence type: `Hypothesis / research question about place effects on the nutrition-income relationship`
- Data source: Homescan panel data (Nielsen) with Retail Market Store (RMS) data for geographic areas
- Assumption flag: `True`
- Assumption notes: Assumes that endogeneity in moving decisions biases estimates upward; uses household fixed effects and event-study design to isolate place effects; discusses instrument validity in related sections.
- Answer:

The central hypothesis is that geographic place effects (the impact of location/retail environment) help explain the nutrition–income relationship, but

these place effects are small, contributing only a tiny portion of the gap (bounded to a few percent).

- Quote snippet:

  > we can bound the medium-term, partial equilibrium effects of place as contributing no more than 2%–3% of the nutrition– income relationship.

- Citation anchors:
  - page=4, words=275-423, section=Results - Place effects and bounds, note=States that place effects contribute no more than 2%–3% of the nutrition–income relationship.
  - page=23, words=275-408, section=Methodology - Testing for place effects, note=Describes testing for place effects via within-household changes after moves.
  - page=27, words=275-482, section=Bounding the effect of location on the nutrition–income relationship, note=Explicitly notes that they can bound the extent to which location explains the nutrition–income relationship.

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

- Category: `A) Research question / contribution`
- Confidence: `low (0.30495106371040515)`
- Retrieval method: `local`
- Evidence type: `Table II results and surrounding text describing dependent variables`
- Data source: Nielsen Homescan data (household-by-quarter) with RMS store data
- Table/Figure: Table II
- Answer:

The main dependent variables are (1) expenditure shares — the share of total grocery expenditures recorded in Homescan, and (2) the Health Index — the overall measure of how healthy households' grocery purchases are.

- 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=page 21, words=0-60, section=Table II, note=Notes define dependent variables: expenditure shares and Health Index.
  - page=page 21, words=275-350, section=Table II notes, note=Panel A/B data definitions and Health Index description.

  - page=page 26, words=0-60, section=Event study/Place effects, note=Health Index as dependent variable in within-household and move analyses.

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

- Category: `A) Research question / contribution`
- Confidence: `low (0.24219802017593928)`
- Retrieval method: `local`
- Evidence type: `text`
- Data source: Homescan data; RMS store data
- Answer:

Ebcqt and Hm are the key treatment/exposure variables. Ebcqt is the supermarket-entry indicator within a distance band by q quarters after quarter t; Hm is the Health Index of groceries sold in RMS stores in geographic area m (ZIP code or county).

- 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 variable definitions (Ebcqt, Bbit), note=Definition of Ebcqt (supermarket-entry indicator) used in the event-study.
  - page=24, words=0-349, section=Health Index exposure (Equation 3), note=Definition of Hm as the Health Index of groceries purchased in RMS stores in geographic area m.

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

- Category: `A) Research question / contribution`
- Confidence: `medium (0.38667750315445915)`
- Retrieval method: `local`
- Evidence type: `textual`
- Data source: Homescan panel data and Nielsen RMS scanner data
- Table/Figure: Figure V
- Answer:

United States grocery retail market (US consumer food retail), using panel data from 2004–2016 (Homescan) and Nielsen RMS scanner data from 2006–2016.

- Quote snippet:

Event Study of Moves across Counties Using 2004–2016 Homescan data

- Citation anchors:
  - page=8, words=0-349, section=Data sources, note=RMS data period 2006–2016; mentions U.S. context via Dietary Guidelines.
  - page=26, words=0-282, section=Figure V/Methods, note=Header for Figure V: Event Study of Moves across Counties Using 2004–2016 Homescan data.
  - page=32, words=0-349, section=Geographic scope, note=Mentions West Coast / East Coast produce variation, implying US geographic scope.

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

- Category: `A) Research question / contribution`
- Confidence: `low (0.2802098579059066)`
- Retrieval method: `local`
- Evidence type: `empirical evidence from a structural demand model showing demand dominates supply in explaining the nutrition-income gap`
- Data source: Homescan panel data; RMS store data; Nielsen Datasets (Kilts Center)
- Answer:

Demand-side factors, not supply, explain the nutrition–income gap; differences in demand for healthy groceries by income account for most of nutritional inequality.

- Quote snippet:

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

- Citation anchors:
  - page=6, words=0-40, section=Empirical findings on supply vs. demand, note=Key finding that demand dominates the nutrition–income relationship

A08: What alternative mechanisms are discussed?

- Category: `A) Research question / contribution`
- Confidence: `low (0.21444864256378068)`
- Retrieval method: `local`
- Evidence type: `Discussion of alternative policy mechanisms and mediating channels affecting healthy eating outcomes`
- Answer:

They discuss an alternative policy mechanism (subsidies for healthy foods) and identify education and nutrition knowledge as mediating mechanisms that influence healthy grocery demand and nutritional inequality.

- Quote snippet:

  > Education explains the largest share of the relationship between demand for healthy groceries and income, and nutrition knowledge explains the second-largest share.

- Citation anchors:
  - page=43, words=0-349, section=VII.B., note=Education and nutrition knowledge as mediating mechanisms in healthy grocery demand.
  - page=43, words=275-387, section=VII.C. Using Subsidies to Reduce Nutritional Inequality, note=Introduces subsidies for healthy foods as an alternative policy mechanism.

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

- Category: `A) Research question / contribution`
- Confidence: `medium (0.3940432466415256)`
- Retrieval method: `local`
- Evidence type: `policy_implications`
- Data source: Nielsen data; Health Index measures; SNAP budget figures (2016); Healthy Food Financing Initiative subsidy data.
- Table/Figure: Table II, Panel B; Online Appendix Table A4; Figure VII
- Assumption flag: `True`
- Assumption notes: Conclusions rely on a structural demand model with assumptions about supply–demand interactions, endogeneity, and the exogeneity of certain shocks; results discussed as suggestive and contingent on model validity.
- Answer:

The main policy implications are that supply-side policies (like encouraging supermarket entry) have limited effects on healthy eating, while means-tested subsidies for healthy groceries could be much more cost-effective; thus policy should shift from promoting access to healthy groceries toward targeted subsidies (e.g., subsidy to bottom-quartile households) to reduce nutritional inequality.

- Quote snippet:

  > Our analyses suggest that supply-side policies, such as encouraging supermarket entry, will have limited effects on healthy eating.

- Citation anchors:
  - page=46, words=Our-eating., section=VII.C. Using Subsidies to Reduce Nutritional Inequality, note=Key claim that supply-side policies have limited effects; subsidies may be more cost-effective.

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

- Category: `A) Research question / contribution`
- Confidence: `medium (0.3897719677164252)`
- Retrieval method: `local`
- Evidence type: `policy-welfare interpretation based on model results`
- Data source: Homescan add-on survey (Nielsen) and SNAP budget data (Center for Budget and Policy Priorities).
- Table/Figure: Table V and Figure VII
- Assumption flag: `True`
- Assumption notes: Subsidy modeled as an ad valorem subsidy proportional to Health Index within a stylized framework; assumes fixed supply conditions per counterfactual analysis and that endogeneity does not overturn results.
- Answer:

Welfare interpretation: supply-side policies (e.g., eliminating food deserts) have limited welfare benefits for reducing nutritional inequality, while demand-side means-tested subsidies for healthy groceries could substantially raise low-income households' healthy eating to the level of high-income households at a modest cost relative to the SNAP budget. This implies welfare gains would be driven mainly by targeted subsidies rather than supply expansion, though the authors note that a principled welfare-maximization analysis is needed for definitive policy guidance.

- Quote snippet:

  > a means-tested subsidy for healthy groceries could increase low-income households' healthy eating to the level of high-income households at an additional cost of only about 15% of the current SNAP budget.

- Citation anchors:
  - page=46, words=our-budget., section=VII.C. Using Subsidies to Reduce Nutritional Inequality, note=Key welfare finding: means-tested subsidy could raise low-income health index to high-income level at cost ~15% of SNAP budget.
  - page=46, words=Thus-inequality., section=VII.C. Using Subsidies to Reduce Nutritional Inequality, note=Policy implication: food deserts policy likely yields little welfare progress.

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

- Category: `A) Research question / contribution`
- Confidence: `low (0.27977498161653724)`
- Retrieval method: `local`

- Evidence type: `summary of limitations drawn from the article's discussion of caveats and interpretation`
- Data source: Nielsen data and Nielsen Datasets at the Kilts Center; other appended data sources noted in the article
- Assumption flag: `True`
- Assumption notes: Explicit acknowledgement that the model rests on multiple assumptions; discussion of endogeneity and the exclusion restriction; results are not claimed to be causal without further validation.
- Answer:

The authors acknowledge several key limitations: (1) the analysis relies on many strong assumptions and should be viewed as suggestive rather than definitive, requiring demonstration projects or randomized trials to confirm effects; (2) potential endogeneity concerns and an instrument-based identification, with explicit notes that endogeneity could bias place effects upward and that the exclusion restriction is an important assumption; (3) the analysis covers only a several-year time horizon, limiting conclusions about long-run effects on eating behavior; (4) the estimated effects of supply-side interventions are small, implying limited policy impact from supermarket entry alone; (5) the study relies on observational data and thus cannot fully establish causality or generalize to other settings without further validation; (6) policy prescriptions (e.g., subsidies) require careful welfare optimization and measurement of market failures before firm recommendations.

- Quote snippet:

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

- Citation anchors:
  - page=47, words=Our-policies., section=VII.C., note=Acknowledges that results rely on many assumptions and are suggestive; calls for demonstration projects.
  - page=34, words=affect-allow., section=Conclusion, note=Notes limited horizon: effects observed may be limited to several-year time horizon the data allow.
  - page=22, words=The-households., section=Results, note=Illustrates that local supermarket access explains only about 1.5% of the health index difference, implying small policy impact from entry.

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

- Category: `A) Research question / contribution`
- Confidence: `low (0.31228149123870697)`
- Retrieval method: `local`
- Evidence type: `novel data and identification claim`

- Data source: Nielsen Homescan panel data and RMS retail scanner data (and related data like TDLinx and ZBP for entry)
- Answer:

The paper claims that its data/identification novelty lies in (1) a novel price-based identification strategy/instrument that exploits cross-market variation in retailer presence and relative prices, enabling identification where other instruments may fail; and (2) the use of a new Homescan add-on survey that links large-sample scanner data to measures of health preferences and nutrition knowledge.

- Quote snippet:

  > our structural demand analysis builds on the framework introduced by Dubois, Griffith, and Nevo (2014) and adds a novel identification strategy and price instrument.

- Citation anchors:
  - page=7, words=0-349, section=II.A, note=Statement that identification strategy and price instrument are novel.
  - page=7, words=0-349, section=II.A, note=Notably, the new Homescan add-on survey provides opportunity to connect scanner data to health measures.
  - page=34, words=0-349, section=VI.C, note=Instrument is novel in the literature and usable where others fail exclusion restriction.

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.3494502439440793)`
- Retrieval method: `local`
- Evidence type: `instrumental variable approach; identification via price instrument Pjmt`
- Data source: Homescan grocery purchases data cited in the QJE article
- Assumption flag: `True`
- Assumption notes: Exogeneity of the price instrument: E[ijt · Pjmt] = 0; assumes retailer pricing advantages vary by market independently of households' tastes.
- Answer:

An instrumental-variables identification strategy using a novel price instrument that captures retailers' cross-market cost advantages and geographic presence to identify demand across product groups, assuming the instrument is uncorrelated with households' idiosyncratic tastes.

- Quote snippet:

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

Pjmt for group j in household i's market.

- Citation anchors:
  - page=33, words=275-397, section=Identifying assumption, note=Paragraph stating the key identifying assumption: E i jt Pjmt = 0
  - page=31, words=275-432, section=Instrument construction, note=Definition and construction of the price instrument Pjmt from retailer pricing advantages and market presence

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.27880782530873593)`
- Retrieval method: `local`
- Evidence type: `textual description of study design`
- Data source: Nielsen Homescan panel data
- Assumption notes: The authors acknowledge that random assignment (experimental design) is ideal but not feasible; they rely on observational data with fixed effects and event-study methods to address endogeneity, characteristic of a quasi-experimental approach.
- Answer:

Quasi-experimental

- Quote snippet:

  While the ideal experiment to measure place effects would randomly assign households to different neighborhoods, households in our data move for reasons that may create endogeneity concerns.

- Citation anchors:
  - page=23, words=275-408, section=Place effects / identification, note=States that the ideal experiment would randomly assign households to neighborhoods and that moves raise endogeneity concerns, implying a quasi-experimental, observational approach with controls.

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

- Category: `B) Identification strategy / causal design`
- Confidence: `medium (0.4358316001092912)`
- Retrieval method: `local`
- Evidence type: `explanation of instrument-based identification using supply-side cost advantages and retailer presence across markets (VI.B).`
- Answer:

The exogenous variation comes from a price instrument built on retailers' cost advantages in supplying UPCs and retailers' differing presence across markets, which induces relative price variation across product groups within a market in a given year.

- Quote snippet:

  > We construct our instrument as follows. For retail chain r's cost advantage in supplying UPC k relative to the national average is thus ln (pkrt,−m) = ln (pkrt,−m) − ln (pkt,−m).

- Citation anchors:
  - page=page 32, words=We construct our instrument as follows.-markets., section=VI.B. Price Endogeneity, note=Instrument uses retailer cost advantages and cross-market retailer presence to create exogenous price variation.

B04: What is the treatment definition and timing?

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.12924186504223473)`
- Retrieval method: `local`
- Evidence type: `Definition of treatment and timing in the study's event-study specification.`
- Data source: Nielsen Homescan data, 2004–2016
- Table/Figure: Figure IV (Event Study) and Table II (regression results)
- Assumption flag: `True`
- Assumption notes: Driving times computed with Google Maps API assuming no congestion delay; treatment defined by supermarket entry within 0–10 or 10–15 minute drive of tract centroid.
- Answer:

Treatment is the entry of a supermarket within a 0–10 minute or 10–15 minute driving distance from a census tract centroid, with timing defined in quarters after the entry quarter t (via the event-study indicator Ebcqt).

- 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=0-349, section=Event-study treatment definition, note=Defines Sbct and Ebcqt; treatment is supermarket entry within driving-time bands and timing is in quarters after quarter t.
  - page=17, words=275-411, section=Event-study treatment definition, note=Further description of timing and indicators for entry within bands.

B05: What is the control/comparison group definition?

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.19746373462545141)`
- Retrieval method: `local`
- Evidence type: `textual description of the baseline/control group in Figure VI counterfactuals`
- Data source: QJE article 'Food Deserts and Nutritional Inequality', Figure VI description (page 40).
- Table/Figure: Figure VI
- Answer:

The control/comparison group is the base category, defined as each income group retaining its own preferences and facing its own local supply conditions.

- Quote snippet:

  The base category measures the Health Index for each income group when each group retains their own preferences and faces their own local supply conditions.

- Citation anchors:
  - page=page 40, words=0-285, section=FIGURE VI, note=Baseline (base category) description for health index counterfactuals

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

- Category: `B) Identification strategy / causal design`
- Confidence: `medium (0.39140356643933755)`
- Retrieval method: `local`
- Evidence type: `Equation (7) baseline regression specification from VI.A Empirical Model`
- Data source: Homescan data (Nielsen Homescan) used for calorie demand model
- Answer:

ln Yi jt = − ln p̃i jt − ˜c ãi jct − + j + m + t + i jt , c=2

- Quote snippet:

  ln Yi jt = − ln p̃i jt − ˜c ãi jct − + j + m + t + i jt , c=2

- Citation anchors:
  - page=page 30, words=0-349, section=VI.A. Empirical Model, note=Equation (7) baseline regression for annual calorie demand; fixed effects included.

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.33492295257670746)`
- Retrieval method: `local`
- Evidence type: `textual evidence from methods sections describing FE specifications`
- Data source: Homescan data (QJE article analysis)
- Answer:

Fixed effects included: unit fixed effects (household fixed effects phi_i); time fixed effects (year indicators mu_t); two-way fixed effects combining household and year fixed effects; higher-dimensional fixed effects including household-by-census-tract fixed effects and census-division-by-quarter fixed effects.

- Quote snippet:

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

- Citation anchors:
  - page=page 21, words=275-503, section=Estimation, note=mentions census-division-by-quarter and household-by-census-tract fixed effects
  - page=page 24, words=0-349, section=Equation (3) and fixed effects, note=defines mu_t as year indicators and phi_i as household fixed effect
  - page=page 24, words=275-452, section=Event study discussion, note=discusses FE structure and robust SE (two-way clustering) around moves

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.33360752589448195)`
- Retrieval method: `local`
- Evidence type: `textual`
- Answer:

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

- Quote snippet:

  Robust standard errors, clustered by household and census tract, are in parentheses.

- Citation anchors:

  - page=page 17, words=275-411, section=Estimation details, note=Robust standard errors with two-way clustering by household and census tract.
  - page=page 21, words=275-503, section=Results/Robust SEs, note=Robust standard errors, clustered by household and census tract, are in parentheses.

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

- Category: `B) Identification strategy / causal design`
- Confidence: `medium (0.373655378166589)`
- Retrieval method: `local`
- Evidence type: `verbatim quote from context`
- Assumption flag: `True`
- Assumption notes: Exogeneity/exclusion restriction: the price instrument is assumed uncorrelated with unobserved tastes (E[i jt Pjmt] = 0). The key economic content is that chains do not have comparative pricing advantages where customers have unobserved stronger tastes.
- Answer:

Exclusion restriction (instrument exogeneity): the instrument must be uncorrelated with households' idiosyncratic preferences, i.e., E[i jt Pjmt] = 0.

- Quote snippet:

  Our identifying assumption is that household i's idiosyncratic preferences for product group j are uncorrelated with the price instrument Pjmt for group j in household i's market: E i jt Pjmt = 0.

- Citation anchors:
  - page=33, words=275-397, section=VI.C. Method of Moments Estimation, note=Identifying assumption: instrument exogeneity (no correlation between idiosyncratic preferences and price instrument).

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.26627537021464204)`
- Retrieval method: `local`
- Evidence type: `Two tests of instrument validity using observable characteristics to infer potential correlations with tastes, i.e., no association with predicted purchases and no systematic differences across income counties for health-related product groups.`
- Data source: Online Appendix Table A9

- Assumption notes: Identifying assumption: the price instrument Pjmt is uncorrelated with households' idiosyncratic preferences for product group j in a market. The authors provide two tests using observable characteristics to assess correlation with tastes; these tests cannot rule out correlation with unobserved tastes.
- Answer:

They perform two observable-characteristics-based tests to assess instrument validity: (1) predicting purchases of product groups using household demographics and county-average demographics shows the price instrument is not associated with those predicted purchases after fixed effects, and (2) the instrument is not systematically different in high–Health Index product groups between low- and high-income counties. See Online Appendix Table A9 for details.

- 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=33, words=275-397, section=VI.B Instrument validity tests, note=Describes the identifying assumption and the two tests linking the instrument to observable tastes.
  - page=34, words=0-349, section=VI.B Instrument validity tests, note=Continuation of tests; states the first test and introduces the second test comparing health-index groups across counties.
  - page=34, words=275-385, section=VI.B Instrument validity tests, note=Mentions Online Appendix Table A9 for details.

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.3221077412965917)`
- Retrieval method: `local`
- Evidence type: `Event-study analyses around supermarket entry and moves (Figure IV and Figure V); includes balanced-panel design and post-entry results; no explicit pre-trend discussion provided.`
- Data source: Homescan data (2004–2016), household-by-quarter
- Table/Figure: Figure IV and Figure V
- Answer:

Yes. The paper reports event-study analyses around supermarket entry (Figure IV) and around moves across counties (Figure V). Post-entry results show the entrant expenditure share rising by about 2 percentage points one year after entry (larger effects in food deserts), while post-move analyses find no statistically significant change in the Health Index after moves. The text describes the

event-study setup and uses a balanced panel for identification, but it does not explicitly report pre-trend (pre-entry) coefficient results.

- 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=17, words=275-411, section=Event-study introduction, note=Mentions using the Homescan sample weights and presenting event-study results before estimating equations.
  - page=18, words=0-349, section=Supermarket entry event-study design, note=Describes Figure IV, [0, 10)q parameters, balanced panel, and omitted category q = -1.
  - page=19, words=0-310, section=Supermarket entry event-study results, note=States the post-entry result: entrant expenditure share up ~2 percentage points after one year; larger in food deserts.
  - page=24, words=0-349, section=Moves event-study setup, note=Notes balanced panel around a move and graphical event-study before estimating equation (3).
  - page=26, words=0-282, section=Moves event-study results, note=Figure V presents moves across counties; postmove Health Index results are not statistically significant.

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.18509711708403398)`
- Retrieval method: `local`
- Evidence type: `textual`
- Data source: QJE article 'Food Deserts and Nutritional Inequality'
- Assumption flag: `True`
- Assumption notes: E ijt Pjmt = 0; instrument uncorrelated with unobserved tastes; cross-market variation identifies the effect.
- Answer:

The study uses Pjmt as an instrumental variable for prices. Pjmt is the weighted average cost advantage that retail chains in market m have for UPCs in product group j, constructed from each chain's presence Nrmt and average sales per store Njrt across markets, together with the national-level calories Nkt and the within-market price gaps ln(pkrt,−m). The instrument variation comes from the interaction of ln(pkrt,−m) and chain presence across markets. The first stage regresses log price pkrt,−m on Pjmt (with product group, market, and year fixed effects). First-stage F-statistics range from 243 to 260 across income quartiles, supporting relevance. The identifying assumption is that ijt is uncorrelated with Pjmt (E ijt Pjmt = 0); they also test whether the instrument correlates

with predicted tastes and find no such relation.

- 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, section=IV instrument construction, note=Definition of instrument construction and rationale (we construct our instrument as follows).
  - page=page 32, words=0-349, section=Instrument definition, note=Definition of Pjmt and how it uses Nrmt, Njrt, Nkt, and ln(pkrt,−m).
  - page=page 32, words=275-431, section=Variation and first-stage, note=Variation arises from ln(pkrt,−m) and retailer presence across markets.
  - page=page 33, words=0-349, section=Assumptions and tests, note=Identifying assumption E ijt Pjmt = 0; tests for correlation with tastes.
  - page=page 33, words=275-397, section=First-stage statistics, note=First-stage F-statistics 243-260; instrument relevance.
  - page=page 34, words=0-349, section=Estimation and tests, note=Method of moments estimation; description of tests of instrument validity.

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.30465824904407074)`
- Retrieval method: `local`
- Evidence type: `not_reported_in_context`
- Data source: QJE article 'Food Deserts and Nutritional Inequality', 134(4), 1793 (online version: https://academic.oup.com/qje/article/134/4/1793/5492274)
- Answer:

The provided excerpt does not report the first-stage strength metrics (F-statistic, partial R-squared, or explicit relevance evidence) for the instrument; it describes the instrument construction and validity tests but does not give numeric first-stage statistics.

- 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=32, words=0-349, section=Instrument construction, note=Definition of the price instrument Pjmt (equation (9)) and its pricing-based variation.
  - page=33, words=275-397, section=Instrument validity, note=Identifying assumption E[ijt Pjmt] = 0 and discussion of tests relating the instrument to predictable tastes.
  - page=34, words=0-349, section=Validity tests, note=Discussion of tests for no systematic relation between the instrument and predictable tastes; no first-stage statistics reported.

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.26723474727038254)`
- Retrieval method: `local`
- Evidence type: `not_mentioned_in_context`
- Answer:

The provided context does not specify a regression discontinuity design, running variable, cutoff, or bandwidth.

- Quote snippet:

  > We compute driving times between each census tract centroid and the address of each entering supermarket using the Google Maps application program interface (API) and assuming no congestion delay.

- Citation anchors:
  - page=17, words=0-349, section=FOOD DESERTS AND NUTRITIONAL INEQUALITY, note=No RDD running variable, cutoff, or bandwidth discussed; analysis uses fixed effects and event-study with entry timing.
  - page=24, words=0-349, section=THE QUARTERLY JOURNAL OF ECONOMICS, note=No RDD discussion; endogeneity/place effects discussed instead.

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.32135184799968625)`
- Retrieval method: `local`
- Evidence type: `methodology description from the article describing the event-study DiD design and the regression equations.`

- Data source: Homescan data (with fixed effects; supermarket-entry analysis), using robust standard errors with two-way clustering.
- Answer:

They use a staggered difference-in-differences (DiD) implemented as an event-study style two-way fixed-effects regression. They define an indicator Ebcqt for whether a supermarket entered within a distance band by q quarters after quarter t and estimate dynamic treatment effects bq from equation (2) using a balanced panel around one entry, with household and census-tract fixed effects and two-way clustering.

- 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=0-349, section=Event study and DiD framework, note=Definition of the event-study indicator Ebcqt and the regression framework for estimating bq in a DiD setting.
  - page=18, words=0-349, section=Balanced panel around one entry and equation (2), note=Description of the balanced-panel identification strategy and the regression specification (2) used to identify bq and bq.

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

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.3232268221350246)`
- Retrieval method: `local`
- Evidence type: `textual dataset description`
- Data source: Homescan, Nielsen RMS, TDLinx, ZIP Code Business Patterns (ZBP), USDA nutrition databases (Food and Nutrient Database for Dietary Studies; National Nutrient Database for Standard Reference), Nielsen Datasets
- Answer:

Homescan, Nielsen RMS, Nielsen TDLinx, ZIP Code Business Patterns (ZBP), and USDA nutrition databases (Food and Nutrient Database for Dietary Studies; National Nutrient Database for Standard Reference).

- Quote snippet:

  > We measure supermarket entry using Nielsen's TDLinx data set, a geocoded census of all food retailers in the United States, including the month each store opened.

- Citation anchors:

  - page=1, words=0-349, section=Introduction, note=Nielsen Datasets at the Kilts Center; data sources referenced.
  - page=8, words=0-349, section=II.A, note=Homescan household data used; demographic and nutrition variables described.
  - page=8, words=275-393, section=II.A, note=Nielsen RMS data described (prices, sales, channels).
  - page=9, words=275-389, section=II.B, note=ZIP Code Business Patterns data for ZIP code establishments.
  - page=10, words=0-349, section=II.B, note=TDLinx data set for supermarket entry dates.
  - page=34, words=0-349, section=II.C, note=USDA nutrition databases for nutrition facts.

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

- Category: `C) Data, sample, and measurement`
- Confidence: `medium (0.4075221686940247)`
- Retrieval method: `local`
- Evidence type: `textual evidence from the data/methods sections`
- Data source: Homescan households and Nielsen RMS data
- Answer:

household

- Quote snippet:

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

- Citation anchors:
  - page=page 17, words=275-411, section=IV, note=mentions using the Homescan data and clustering by household; describes data at household level
  - page=page 9, words=275-389, section=II.B. Grocery Retail Establishments, note=notes collapsing data to the household (or store)-by-year level

C03: What is the sample period and geographic coverage?

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.29929904275261066)`
- Retrieval method: `local`
- Evidence type: `data_period_and_geography`
- Data source: Homescan data (Nielsen) and Nielsen RMS data
- Answer:

Homescan data: 2004–2016; RMS data: 2006–2016; geographic coverage: United States.

- Quote snippet:

  There were 6,721 entries of supermarkets in the United States between 2004 and 2016.

- Citation anchors:
  - page=9, words=0-349, section=Data description, note=Homescan data cover 2004–2016 (731,994 household-by-year observations).
  - page=8, words=0-349, section=Data description, note=Nielsen RMS data cover 2006–2016.
  - page=10, words=0-349, section=Geographic coverage, note=There were 6,721 entries of supermarkets in the United States between 2004 and 2016.

C04: What are the sample restrictions / inclusion criteria?

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.27031212147650685)`
- Retrieval method: `local`
- Evidence type: `text`
- Data source: Homescan and Nielsen RMS
- Answer:

Sample restrictions / inclusion criteria include: (1) using Homescan households (with an 8,000-household magnet subsample in 2004–2006) and Nielsen RMS data (2006–2016); (2) RMS data exclude liquor stores and do not include bulk produce or other nonpackaged items; (3) results are weighted for national representativeness; (4) for move-related analyses, observations are restricted to a balanced panel around a move with a 50% local shopping restriction to curb multiple moves and exposure variation; (5) analyses use four household income quartiles, with the bottom quartile defined as the lowest quartile of residuals from regressing household income on household size, age, and year indicators; (6) data are organized at the household-by–product group–by-year level, with J = 45 product groups in the Method of Moments estimation.

- Quote snippet:

  For 2004–2006, however, the data include an 8,000-household "magnet" subsample that also recorded prices and weights of nonpackaged groceries.

- Citation anchors:
  - page=page 8, words=For-subsample., section=Data sources, note=Magnet subsample included for 2004–2006.
  - page=page 8, words=We exclude liquor stores.-stores., section=Nielsen RMS data, note=RMS excludes liquor stores; excludes bulk/ nonpackaged items.
  - page=page 24, words=Before estimating equation-move., section=Move-panel design, note=Balanced panel around a move.

  - page=page 25, words=The 50%-living., section=Move-panel design, note=50% local shopping restriction.
  - page=page 20, words=To match-indicators., section=Bottom income quartile, note=Definition of bottom income quartile as residuals from regressions on size, age, year.
  - page=page 34, words=For estimation-quartiles,, section=Method of Moments, note=Four income quartiles used in MoM estimation.
  - page=page 34, words=Data are at the household-by-level., section=Data structure, note=Data are household-by–product group–by-year level.
  - page=page 34, words=We define J = 45-minimize, section=Product groups, note=J = 45 product groups used in defining product groups.
  - page=page 8, words=The Nielsen RMS data consist-stores., section=Data scope, note=RMS data span 2006–2016 across stores.

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

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.2834155924949919)`
- Retrieval method: `local`
- Evidence type: `descriptive statistics / data description`
- Data source: Homescan data; Nielsen RMS data
- Answer:

731,994

- Quote snippet:

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

- Citation anchors:
  - page=page 9, words=0-349, section=Notes, note=N Homescan data include 731,994 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 (0.3486200980697796)`
- Retrieval method: `local`
- Evidence type: `Definition and construction of Health Index; normalization`
- Data source: Nielsen Homescan data (2004–2016)
- Answer:

The key outcome is the Health Index, a composite measure of the healthfulness of grocery purchases. It is built as a weighted average of UPC-level Health Index values within product groups (calorie-weighted when computing area-level

Health Index, then aggregated to households), and is normalized to have a mean of 0 and a standard deviation of 1 across households (i.e., a z-score).

- 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=21, words=275-503, section=Health Index construction, note=Definition and normalization of Health Index; Health Index normalized to mean 0 and sd 1 across households.
  - page=24, words=275-452, section=Area average Health Index, note=Calorie-weighted averaging to compute area-level Health Index; description of RMS-based weighting for groceries.

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

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.2928188734652247)`
- Retrieval method: `local`
- Evidence type: `constructed exposure indicators and event-study design`
- Data source: Homescan data; RMS store data (for Health Index construction); RMS counts used to describe local store availability.
- Answer:

Treatment/exposure is measured with constructed indicators: (1) Ebcqt, an indicator for whether a supermarket entered within a distance band b of census tract c by q quarters after quarter t; (2) Bit, an indicator for being in a balanced panel around a household move; and (3) Hm, the Health Index of packaged groceries purchased in RMS stores in geographic area m (ZIP code or county), used to capture area-level exposure to the local food environment. The Health Index is built as a calorie-weighted average of groceries sold in RMS stores and regressed against area distribution of store channels to align with ZBP data. These constructions underpin the event-study design assessing exposure to entry effects.

- 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 / Measures of exposure, note=Definition of Ebcqt: an indicator for whether a supermar-

ket entered within distance band b by q quarters after t.
  - page=24, words=0-349, section=Balanced panel around a move, note=Definition of Bit: an indicator for whether observation is part of a balanced panel around a move.
  - page=24, words=0-349, section=Health-Index exposure, note=Definition of Hm: Health Index of groceries purchased in RMS stores in geographic area m (ZIP code or county).

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

- Category: `C) Data, sample, and measurement`
- Confidence: `medium (0.35566120363772863)`
- Retrieval method: `local`
- Evidence type: `textual`
- Answer:

Yes. The key covariates are the time-varying household characteristics included in Xit: the natural log of income, the natural log of years of education, indicators for each integer age from 23–90, household size, race indicators, an indicator for whether the household head is married, employment status, and weekly work hours. These covariates are the baseline controls used in the regressions (e.g., Equation 3) and are included together with household fixed effects and year indicators. The analysis also notes robustness checks where they add observed-changes controls (income, job responsibilities, household composition, 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, an indicator for whether the household heads are married, employment status, and weekly work hours.

- Citation anchors:
  - page=17, words=275-411, section=Xit covariates and model specification, note=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, an indicator for whether the household heads are married, employment status, and weekly work hours.
  - page=21, words=275-503, section=Xit covariates and fixed effects, note=mentions employment status and weekly work hours as part of covariates; notes fixed effects context.
  - page=24, words=0-349, section=Equation (3) covariates, note=Xit described as a vector of potentially time-varying household covariates (Table I).
  - page=27, words=275-482, section=Controls/demographics, note=Discusses including household demographics Xit and the

minimal impact on results.
  - page=38, words=275-377, section=Method of Moments Estimation, note=Xit denotes demographics including log years of education, race, married, employment status, and weekly work hours.
  - page=42, words=275-356, section=Demographics and remaining covariates, note=Xit denotes the remaining demographics from Section IV plus two additional variables.

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

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.30557918260254957)`
- Retrieval method: `local`
- Evidence type: `data handling (missing data and weighting)`
- Data source: Homescan (Nielsen) data and RMS price data
- Answer:

Missing data are handled by dropping observations with zero purchases (about 10.6% of observations at the household-by–product group–by-year level); there is no imputation described. Homescan results are weighted for national representativeness, though some analyses explicitly state that observations are not weighted.

- 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=VI.C. Method of Moments Estimation, note=Dropping observations with zero purchases (10.6%).
  - page=8, words=0-349, section=Homescan data weighting, note=Homescan results weighted for national representativeness.
  - page=21, words=275-503, section=Estimation rules / Data weighting, note=Some analyses are not weighted for national representativeness.

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

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.3414683090295303)`
- Retrieval method: `local`
- Evidence type: `direct quote`
- Data source: Homescan panel data and Nielsen RMS price data
- Answer:

Yes. The study uses survey weights (Homescan sample weights) to weight results; Homescan results are weighted for national representativeness, and the

estimation uses the Homescan sample weights.

- Quote snippet:

  use the Homescan sample weights for this analysis.

- Citation anchors:
  - page=8, words=0-349, section=Data description / Weighting, note=Homescan results weighted for national representativeness.
  - page=17, words=275-411, section=Estimation, note=Use the Homescan sample weights for this analysis.

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

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.2554771359486147)`
- Retrieval method: `local`
- Evidence type: `Cross-source data linkage using UPC crosswalks, retailer-code mappings, and geographic harmonization`
- Data source: Homescan panel data; Nielsen RMS data; TDLinx store data; ZIP Code Business Patterns data; USDA nutrition facts databases via crosswalks
- Answer:

Yes. The study links data across Homescan, RMS, ZBP, TDLinx, and USDA nutrition data. Linkage is done by (1) matching UPC-level nutrition facts to UPCs using USDA crosswalks; (2) aligning RMS price/sales data with Homescan UPCs via retailer codes and identifying entrants by linking TDLinx stores to Homescan retailer codes; (3) combining RMS store data with ZIP Code Business Patterns data and adjusting for differences in store-channel distributions across RMS and ZBP; (4) using ZIP-code/county geography to compute area-level Health Index for place-effect analysis.

- 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=Nutrition facts and data linkage, note=UPC-level USDA nutrition facts matched to UPCs using crosswalks developed by the USDA
  - page=10, words=275-424, section=Store-entry linkage, note=TDLinx stores linked to Homescan retailer codes to identify entrants; spurious entrants dropped
  - page=8, words=0-349, section=Data sources overview, note=Description of Homescan, RMS, and data sources used for the analysis

- page=24, words=0-349, section=Geographic linkage and area health index, note=RMS data linked with ZIP/county geography; regression adjustment for RMS vs ZBP distribution differences

C12: What summary statistics are reported for main variables?

- Category: `C) Data, sample, and measurement`
- Confidence: `medium (0.3556705368873661)`
- Retrieval method: `local`
- Evidence type: `Table I descriptive statistics (Panel A and Panel B)`
- Data source: Homescan households data; ZIP Code Business Patterns establishment counts
- Table/Figure: Table I
- Answer:

Table I reports descriptive statistics (means and standard deviations) for main Homescan variables in Panel A and establishment counts for ZIP codes in Panel B.

- Quote snippet:

  > Table I, Panel A presents descriptive statistics for Homescan households. Unless otherwise stated, all Homescan results are weighted for national representativeness.

- Citation anchors:
  - page=8, words=0-349, section=Table I Panel A descriptive statistics, note=Describes Table I Panel A descriptive statistics for Homescan households and weighting.
  - page=9, words=0-349, section=Table I Panel A and Panel B descriptive statistics, note=Gives Panel A descriptive statistics (means and standard deviations) and Panel B ZIP code establishment counts.

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

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.3250038147860591)`
- Retrieval method: `local`
- Evidence type: `descriptive statistics table; descriptive figure; geographic map`
- Data source: QJE article: Food Deserts and Nutritional Inequality
- Table/Figure: Table I; Figure II; Online Appendix Figure A11
- Answer:

Yes. There are descriptive baseline visuals: Table I provides baseline descriptive statistics; Figure II shows store healthfulness and size by ZIP code median

income; and Online Appendix Figure A11 maps the geographic presence of the five largest RMS retailers, establishing baseline spatial patterns.

- Quote snippet:

  TABLE I DESCRIPTIVE STATISTICS

- Citation anchors:
  - page=9, words=0-20, section=TABLE I DESCRIPTIVE STATISTICS, note=baseline descriptive statistics
  - page=14, words=0-60, section=FIGURE II Store Average Healthfulness and Size by ZIP Code Median Income, note=baseline pattern by income
  - page=32, words=275-431, section=Online Appendix Figure A11, note=geographic retailer presence map

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

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `medium (0.3935327573307571)`
- Retrieval method: `local`
- Evidence type: `Direct quote from article's results section.`
- Data source: Homescan data (Nielsen Homescan; used to construct Health Index by expenditure shares).
- Answer:

Negative; about a 0.9% reduction in the top–bottom Health Index gap (i.e., the gap is reduced by roughly 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, section=Table II, Panel B results, note=Headline main effect: effect of one supermarket entry within a 10-minute drive on Health Index gap between top and bottom income quartiles.

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

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `low (0.22635778599379758)`
- Retrieval method: `local`
- Evidence type: `Instrumental variables identification; chain-based price instrument with fixed effects; exogeneity assumption`
- Assumption flag: `True`

- Assumption notes: Exogeneity of the chain-based price instrument: price variation across markets due to chain presence is independent of unobserved consumer tastes.
- Answer:

The preferred specification is an instrumental variables (IV) specification that uses price variation instrumented by grocery chains' comparative advantages across product groups and markets, with product-group and market fixed effects, relying on an exogeneity assumption that chain-driven price variation is independent of unobserved preferences. This specification has a very strong first stage and identifies how prices and availability affect demand across income groups.

- Quote snippet:

  > The key identifying assumption is that geographic variation in prices due to the presence of specific chains is independent of geographic variation in unobserved preferences.

- Citation anchors:
  - page=5, words=0-349, section=Identification via chain-based price instrument, note=Introduces the instrument using chain-based price variation and the key identifying assumption.
  - page=33, words=275-397, section=Identifying assumption and instrument validity, note=States the exogeneity requirement: chains do not have comparative pricing advantages in groups with unobserved tastes.
  - page=35, words=0-349, section=Moment conditions and estimation, note=Describes the identifying moments and how characteristic preferences are estimated (̃, , ,).

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

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `medium (0.44036819972758057)`
- Retrieval method: `local`
- Evidence type: `Regression results from Figure IV panel using Homescan data showing expenditure-share changes and health index effects by supermarket entry.`
- Data source: Homescan panel data
- Table/Figure: Figure IV
- Answer:

The effect is economically meaningful but modest. Supermarket entry raises the entrant's expenditure share by about 2 percentage points one year after entry, which translates to roughly $2 more in grocery expenditures per $100 spent by nearby residents (about 3% in food-desert areas). Health-Index impacts are

very small (around 0.02 standard deviations for the bottom income quartile), and overall the nutrition–income gap is driven largely by demand rather than supply.

- Quote snippet:

  > the entrant expenditure share increases by about 2 percentage points one year after entry. In other words, a retail chain adding a new store earns an additional $2 for every $100 in grocery expenditures by nearby residents.

- Citation anchors:
  - page=21, words=275-503, section=Figure IV results for entrant expenditure share and Health Index, note=Contains the 2 percentage-point increase in entrant expenditure share and the $2 per $100 interpretation, plus the 3% figure for food deserts.

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

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `low (0.33060827674792864)`
- Retrieval method: `local`
- Evidence type: `robustness_checks`
- Data source: QJE article; Online Appendix; Homescan dataset; RMS retailer price data
- Assumption flag: `True`
- Assumption notes: Robustness checks rely on the IV exogeneity/exclusion restriction. They conduct tests to show the instrument is uncorrelated with predicted tastes, though unobserved tastes cannot be tested.
- Answer:

The authors run several robustness checks and find that the main conclusions remain robust. Key checks include: (1) adding county-by-quarter fixed effects and a control for convenience-store entry; results do not qualitatively change. (2) including demographic covariates; these only modestly raise R2 and the interpretation that place effects are likely upper bounds remains. (3) within-household fixed effects event-study design to bound medium-term place effects. (4) an instrumental-variables approach with a price-based instrument; two validity tests show the instrument is not correlated with predicted tastes and does not differ systematically across high-/low-income counties for high–Health Index product groups, supporting exogeneity. (5) online appendix evidence showing the instrument's variation arises from within-market price differences, not demand shocks. Overall, the robustness checks support the conclusion that observed place effects are small in magnitude and likely upper bounds; the main qualitative results 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=27, words=275-482, section=Robustness checks, note=Adding county-by-quarter fixed effects or a control for convenience-store entry does not qualitatively change the results.
  - page=24, words=0-349, section=Place effects bounds interpretation, note=We interpret the results as likely upper bounds on true place effects.
  - page=34, words=0-349, section=Instrument validity tests, note=Two tests show the instrument is not associated with predicted tastes and is not systematically different across income counties for high–Health Index product groups.
  - page=32, words=0-349, section=Instrument variation and identification, note=Online Appendix shows that most variation in the instrument comes from retailer presence and within-market price variation, not cross-market demand shocks.

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

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `low (0.20131888759938782)`
- Retrieval method: `local`
- Evidence type: `instrument validity placebo tests (two robustness checks)`
- Data source: Homescan data (and county demographics as described in the text)
- Answer:

They run two placebo-like tests to assess instrument validity: (1) predict purchases of product groups using household demographics and county-average demographics and check that the instrument is not associated with these predicted purchases (controlling for fixed effects); (2) assess whether the instrument is systematically different in high–Health Index product groups between low- and high-income counties, finding no systematic difference. Overall, the instrument passes these placebo tests, supporting its validity.

- Quote snippet:

  Two tests: (1) predict purchases using household and county demographics; (2) test instrument differences in high–Health Index products across low- vs high-income counties; instrument not associated with predicted purchases and not systematically different.

- Citation anchors:
  - page=33, words=275-397, section=VI.C. Method of Moments Estimation, note=Two tests of instrument validity: (1) predict purchases

using household and county demographics; (2) test instrument differences across high–Health Index product groups in low- vs high-income counties; results show no association with predicted purchases and no systematic difference.

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

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `low (0.25122207858629836)`
- Retrieval method: `local`
- Evidence type: `textual`
- Data source: Online Appendix Table A9
- Table/Figure: Online Appendix Table A9
- Answer:

Two falsification (unaffected-outcome) tests are conducted: (1) predict purchases of each product group j using household demographics and then predict county-level purchases from county-average demographics, finding the instrument is not associated with these predicted purchases; (2) test whether the instrument is systematically different between high-Health Index product groups in low- vs high-income counties, finding no systematic difference.

- Quote snippet:

  First, we predict purchases of product group j using household demographics and then predict county-level purchases; we find the instrument is not associated with predicted purchases.

- Citation anchors:
  - page=33, words=275-397, section=Falsification tests of instrument validity, note=Describes two tests to check for unaffected outcomes: predicted purchases using demographics, and differences across high-Health Index groups by county income.

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

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `medium (0.4341618378637985)`
- Retrieval method: `local`
- Evidence type: `textual`
- Data source: Quarterly Journal of Economics article on nutrition and inequality
- Table/Figure: Figure VII
- Answer:

Heterogeneity results reported include: by income, analyses are conducted on four household income quartiles; higher-income households have stronger preferences for convenience and fresher foods with shorter shelf lives, and there

are differences in willingness-to-pay across quartiles; by region, there is a region/census division heterogeneity illustrated in Figure VII showing the share of the income–healthy grocery demand correlation explained by regional indicators. Size and baseline exposure heterogeneity are not explicitly reported in the provided excerpts.

- Quote snippet:

  For estimation, we construct separate data sets for four household income quartiles

- Citation anchors:
  - page=38, words=0-349, section=VI.C; VII, note=Income heterogeneity: separate data sets for four household income quartiles; higher-income households have stronger preferences; WTP differences across quartiles.
  - page=44, words=0-238, section=VII, note=Region heterogeneity: Figure VII shows census division contributions to explaining the income–healthy demand relationship.
  - page=20, words=275-432, section=Results discussion, note=Bottom income quartile analyses referenced for heterogeneity by income.

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

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `low (0.16600285978385657)`
- Retrieval method: `local`
- Evidence type: `instrument validity mechanism tests (exclusion restriction)`
- Data source: Homescan data
- Assumption notes: Tests are designed to support the exclusion restriction by showing the instrument is not related to predictable tastes.
- Answer:

Two mechanism tests are performed to assess the instrument's validity: (1) test whether purchases of product group j, predicted from household demographics, align with county-level purchases predicted from county demographics, and whether the instrument is not associated with these predicted purchases after standard fixed effects; (2) test whether the instrument is not systematically different in high–Health Index product groups in low- versus high-income counties, relying on the nutrition–income relationship. The results show no relationship to predicted tastes, supporting the instrument's exclusion restriction and its validity for identifying cross-sectional variation.

- Quote snippet:

  We implement two tests, each of which predicts tastes in a different way.

- Citation anchors:
  - page=34, words=0-349, section=VI.B Instrument validity tests, note=Two tests of instrument validity are described: predicted purchases not associated with the instrument after fixed effects; no systematic difference in high–Health Index product groups across income counties.

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

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `medium (0.3718231411053623)`
- Retrieval method: `local`
- Evidence type: `Robustness checks and sampling sensitivity`
- Data source: Homescan (RMS) panel data; analyses consider RMS store distributions
- Answer:

The results are robust to alternative samples/bandwidths/controls: analyses that avoid using the Homescan sample weights (and thus rely on a non-representative sample) still yield similar qualitative conclusions, and adding extra controls such as county-by-quarter fixed effects or convenience-store-entry controls does not qualitatively change the results; Observations are not weighted for national representativeness, and the two-way robust standard errors are used to account for clustering; thus sensitivity to sampling/controls appears limited in the reported robustness checks.

- Quote snippet:

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

- Citation anchors:
  - page=24, words=0-349, section=Place effects & sampling sensitivity, note=Within-household design; not using Homescan sample weights.
  - page=23, words=275-408, section=Robustness checks, note=Adding county-by-quarter fixed effects or a control for convenience store entry does not qualitatively change the results.
  - page=21, words=275-503, section=Sample/weights considerations, note=Observations are not weighted for national representativeness.
  - page=24, words=275-452, section=Estimation robustness, note=Robust standard errors with two-way clustering by household and geographic area.

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

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `low (0.25299725681325247)`

- Retrieval method: `local`
- Evidence type: `conclusion_summary`
- Assumption flag: `True`
- Assumption notes: Results are described as suggestive due to many model assumptions; authors call for demonstration projects/randomized experiments to confirm effects.
- Answer:
- Supply-side (new supermarket) interventions have small effects on healthy eating; they explain at most about 1.5% of the nutrition gap between high- and low-income households.
- People travel long distances for shopping, and entry of a new supermarket mostly diverts purchases from other supermarkets rather than improving overall healthy eating.
- Place effects explaining nutritional inequality are limited: partial equilibrium place effects are no more than ~3%, and fully equalizing supply would reduce the gap by at most ~10%.
- Means-tested subsidies for healthy groceries could raise low-income households' healthy eating to the level of high-income households at an additional cost of about 15% of the current SNAP budget.
- Because the model relies on many assumptions, the authors view these results as suggestive and advocate demonstration projects or randomized experiments to confirm policy effects.
- A reframing of food deserts: eliminating them is unlikely to yield large gains in reducing nutritional inequality; policy emphasis may shift from access to subsidies.
- Quote snippet:

  Entry of a new supermarket has economically small effects on healthy grocery purchases, and we can conclude that local access to supermarkets explains no more than about 1.5% of the Health Index difference.

- Citation anchors:
  - page=48, words=0-40, section=Conclusion, note=Policy implications: small supply-side effects, potential of subsidies; quantified effects (1.5%, 3%, 10%, 15%).
  - page=47, words=0-20, section=Conclusion, note=Caveat about model assumptions and need for demonstration projects.

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

- Category: `E) Citations and related literature`
- Confidence: `low (0.29822085113381763)`
- Retrieval method: `local`
- Evidence type: `literature review / framing of prior work`
- Answer:

The most important prior papers cited are Bitler & Haider (2011), which is central because it challenges the assumption that supply-side factors explain food deserts and questions whether food deserts cause unhealthy eating; Powell et al. (2007) and Larson, Story, and Nelson (2009) provide foundational empirical links between store availability, neighborhood characteristics, and healthy eating; Horowitz et al. (2004) and Jetter & Cassady (2005) document barriers and costs to accessing healthy foods in low-income areas; Allcott et al. (2017) and related replication work (2019) extend the geographic analysis of poverty, nutrition, and food choices across markets, reinforcing the role of geography and market structure in nutritional inequality. Together, these works establish the key debate the paper engages with: whether access constraints drive nutritional inequality and to what extent supply conditions matter relative to other factors.

- Quote snippet:

  "Bitler and Haider (2011) observe that "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=275-415, section=Literature discussion, note=Mentions Bitler & Haider (2011) as addressing whether food deserts are caused by supply-side factors.
  - page=49, words=0-349, section=References, note=Cites Bitler & Haider (2011) and Allcott et al. (2019) replication data; Powell et al. (2007) appears here.
  - page=50, words=0-349, section=References, note=Cites Larson, Story, and Nelson (2009) on neighborhood environments and access to healthy foods.
  - page=1, words=0-349, section=References, note=Lists foundational early work such as Alwitt & Donley (1997); Horowitz et al. (2004); Jetter & Cassady (2005).
  - page=2, words=275-415, section=References, note=Allcott, Diamond, Dubé, Handbury, Rahkovsky, Schnell (2017) on geography of poverty and nutrition; related replication data (2019).
  - page=49, words=0-349, section=References, note=Replication data for Allcott et al. (2019) and additional prior work including Bitler & Haider (2011).

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

- Category: `E) Citations and related literature`
- Confidence: `low (0.26223774805232425)`
- Retrieval method: `local`
- Evidence type: `explicit citation and subsumption of prior related works`
- Answer:

This work directly builds on and subsumes the prior work by Handbury, Rahkovsky, and Schnell (2015) and by Allcott, Diamond, and Dubé (2017).

- Quote snippet:

  > This article subsumes and replaces our previous work, Handbury, Rahkovsky, and Schnell (2015) and Allcott, Diamond, and Dubé (2017).

- Citation anchors:
  - page=49, words=275-543, section=page 49 words 275-543, note=Contains the sentence stating the article subsumes and replaces the authors' previous work (Handbury et al. 2015; Allcott et al. 2017).
  - page=49, words=0-349, section=page 49 words 0-349, note=References list includes Handbury, Rahkovsky, and Schnell (2015) and Allcott, Diamond, and Dubé (2017).

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

- Category: `E) Citations and related literature`
- Confidence: `low (0.2676619056751341)`
- Retrieval method: `local`
- Evidence type: `text excerpt citing benchmark/comparison papers`
- Answer:

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

- Quote snippet:

  > DellaVigna and Gentzkow (2019) 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=VI.C. Method of Moments Estimation, note=Mentions benchmark/comparison papers 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 (0.3182354015902439)`
- Retrieval method: `local`
- Evidence type: `dataset_usage`
- Data source: Nielsen Homescan consumer panel data; Nielsen RMS retail scanner data; Nielsen Datasets (Kilts Center); ZIP Code Business Patterns data
- Answer:

The study cites Nielsen data sources (Homescan consumer panel data and RMS retail scanner data) and related Nielsen Datasets via the Kilts Center, plus ZIP Code Business Patterns data for ZIP-code establishment counts. Homescan provides household demographics, calorie-needs calculations, and purchase data (with a magnet subsample and PanelViews health measures), while RMS contributes weekly UPC-level prices and sales volumes across stores; ZIP Code Business Patterns supplies geographic establishment counts for regional 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=1, words=275-483, section=Data sources / Acknowledgments, note=Mentions Nielsen data and Datasets at Kilts Center (Nielsen data sources).
  - page=8, words=0-349, section=Data sources - Homescan and RMS, note=Describes Nielsen RMS data (prices and sales) and Homescan magnet subsample; PanelViews reference appears here.
  - page=9, words=275-389, section=Data sources - ZIP Code patterns, note=States ZIP Code Business Patterns data for ZIP code establishment counts.
  - page=8, words=0-349, section=Data sources - PanelViews, note=Notes Homescan PanelViews add-on survey for health measures (BMI, diabetes) in 2017.

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

- Category: `E) Citations and related literature`
- Confidence: `medium (0.43651591775325593)`
- Retrieval method: `local`
- Evidence type: `Textual discussion of IV/diD-like event study methods and a set of econometric references embedded in the`

`article's references.`

- Data source: Nielsen RMS data (Retail Measurement Services) and Homescan panel data; Nielsen datasets at Kilts Center
- Assumption flag: `True`
- Assumption notes: Assumes demand shocks are uncorrelated across markets.
- Answer:

The article cites instrumental variables (IV) as its main econometric approach, uses an event-study/within-household fixed-effects design (a DiD-style framework around moves), and references key econometric papers such as Kling, Liebman, and Katz (2007) on experimental neighborhood effects and Hausman (1996) on endogeneity; it also cites DellaVigna and Gentzkow (2019) as an example of pricing-based instruments.

- Quote snippet:

  > This instrument is novel in the literature, and it can be used in situations in which other instruments do not generate identification or fail the exclusion restriction.

- Citation anchors:
  - page=page 32, words=0-349, section=VI.B. Instrumental Variables, note=Describes the price-based instrument and cross-market variation for identification.
  - page=page 24, words=0-349, section=Event study / within-household fixed effects, note=Notes use of an event study around moves and household fixed effects.
  - page=page 34, words=0-349, section=IV discussion / instrument mention, note=Mentions DellaVigna and Gentzkow (2019) and describes instrument novelty.
  - page=page 50, words=0-349, section=References (econometric methods), note=Cites Kling, Liebman, and Katz (2007) Experimental Analysis of Neighborhood Effects; Hausman (1996); and other econometric methods.

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

- Category: `E) Citations and related literature`
- Confidence: `low (0.2710586230403233)`
- Retrieval method: `local`
- Evidence type: `literature_position_against_claim`
- Answer:

Yes. The paper positions itself against the view that food deserts (neighborhood supply-side factors) drive nutritional inequality, citing Bitler and Haider (2011) as arguing there is not sufficient evidence that food deserts are the systematic

cause of unhealthy eating.

- 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=0-349, section=Introduction, note=Mentions Bitler & Haider (2011) arguing against the assumed supply-side cause of food deserts.
  - page=2, words=275-415, section=Introduction, note=Continues discussion referencing Bitler & Haider (2011) and their conclusion about evidence.

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

- Category: `E) Citations and related literature`
- Confidence: `low (0.32828281577445884)`
- Retrieval method: `local`
- Evidence type: `Citation to online appendix and data/code repository`
- Data source: Harvard Dataverse (doi:10.7910/DVN/MSOBYI)
- Answer:

Yes. The article references an Online Appendix and a data/code repository (Harvard Dataverse) for replicating tables and figures.

- Quote snippet:

  Supplementary MATERIAL An Online Appendix for this article can be found at The Quarterly Journal of Economics online. Code replicating tables and figures can be found in Harvard Dataverse, doi: 10.7910/DVN/MSOBYI.

- Citation anchors:
  - page=48, words=0-323, section=Supplementary MATERIAL, note=Online Appendix and code repository (Harvard Dataverse doi: 10.7910/DVN/MSOBYI)

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

- Category: `E) Citations and related literature`
- Confidence: `low (0.3164201485851723)`
- Retrieval method: `local`
- Evidence type: `literature gap / synthesis`
- Answer:

The citations leave open whether supply-side factors (food deserts / store availability) causally explain unhealthy eating among low-income groups; specifically, there is not enough evidence to determine if food deserts systematically cause poor dietary quality, and some literature notes that existing work implicitly assumes supply-side causes without solid causal evidence.

- 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, section=Literature gap discussion, note=Bitler & Haider (2011) claim there is not sufficient evidence to determine whether food deserts are the systematic cause of unhealthy eating by low-income people; they also note that much existing research implicitly assumes supply-side causes.

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

- Category: `F) Replication and transparency`
- Confidence: `low (0.19556110191614948)`
- Retrieval method: `local`
- Evidence type: `text`
- Data source: Harvard Dataverse (Allcott et al. 2019) doi: 10.7910/DVN/MSOBYI
- Answer:

Yes. Replication code and data files are provided in Allcott et al. (2019) on the Harvard Dataverse (doi: 10.7910/DVN/MSOBYI); the article also notes an Online Appendix available through the QJE website.

- 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=Mentions online appendix and replication code in Harvard Dataverse.

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

- Category: `F) Replication and transparency`
- Confidence: `low (0.24512781532025157)`
- Retrieval method: `local`
- Evidence type: `not mentioned`

- Answer:

No. The provided context does not mention a pre-analysis plan or a registered trial.

- Quote snippet:

  Supplementary Material An Online Appendix for this article can be found at The Quar- terly Journal of Economics online.

- Citation anchors:
  - page=page 34 words 0-349, words=0-349, section=Supplementary Material, note=Mentions supplementary material/online appendix but no reference to pre-analysis plan or registered trial

F03: Are data access constraints disclosed (restricted access, proprietary data, NDAs)?

- Category: `F) Replication and transparency`
- Confidence: `low (0.28718933021838233)`
- Retrieval method: `local`
- Evidence type: `Proprietary data disclosure; data availability statements`
- Data source: Nielsen RMS data; Nielsen Datasets at Kilts Center; TDLinx; ZIP Code Business Patterns; replication data via Harvard Dataverse
- Answer:

Yes. The study discloses that part of its data come from proprietary Nielsen RMS data accessed via Nielsen Datasets, with explicit notes that Nielsen has no role in the analysis and that the conclusions are those of the researchers. It also states that replication data are available (via Harvard Dataverse).

- Quote snippet:

  The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen.

- Citation anchors:
  - page=page 1, words=0-349, section=Data sources and access restrictions, note=Nielsen RMS data are proprietary; conclusions are researchers' and not Nielsen's; Nielsen not involved in analysis.
  - page=page 1, words=275-483, section=Data sources and access restrictions, note=Continuation of Nielsen data disclosure and data source to Kilts Center.
  - page=page 49, words=0-349, section=References/Data availability, note=Replication data for this article are available via Harvard Dataverse.

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

- Category: `F) Replication and transparency`
- Confidence: `low (0.22463398165468668)`
- Retrieval method: `local`
- Evidence type: `Methods / Data construction`
- Data source: Homescan data; Nielsen RMS data; TDLinx; ZIP Code Business Patterns
- Answer:

Yes. The paper documents key data cleaning and construction steps, including exclusion of liquor stores and nonpackaged items, restricting to supermarket-type entries with retailer-code matching, deflating prices, using TDLinx to identify true entrants, dropping spurious entrants, using ZIP Code Business Patterns data, and collapsing data to household/store-year level.

- 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=II.B Grocery Retail Establishments, note=Excludes liquor stores; nonpackaged items not included; channel composition and deflation mentioned.
  - page=10, words=0-349, section=II.B Grocery Retail Establishments, note=Includes only club stores, supercenters, and grocery stores; retailer-code matching; excludes small mom-and-pop groceries.
  - page=10, words=275-424, section=II.B Grocery Retail Establishments, note=6,721 supermarket entries; use of ZIP Code Business Patterns data; TDLinx-based entrant screening.
  - page=10, words=275-424, section=II.B Grocery Retail Establishments, note=Use list of TDLinx stores transferred through mergers to exclude spurious entrants; drop identical-subtype entries in same census block year-over-year.
  - page=9, words=275-389, section=II.B Grocery Retail Establishments, note=Empirical analyses note: collapse data to household (or store)-by-year level for analyses.

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

- Category: `F) Replication and transparency`
- Confidence: `low (0.28259493276264314)`
- Retrieval method: `local`
- Evidence type: `supplementary_materials`
- Data source: Online Appendix
- Answer:

Robustness and sensitivity analyses are partially reported; the main text refers to robustness checks and validity tests, but full robustness/sensitivity results are reported in the Online Appendix rather than the main article.

- Quote snippet:

  See Online Appendix Table A9 for details.

- Citation anchors:
  - page=7, words=0-349, section=Data and Appendix materials, note=All appendix material is in the Online Appendix.
  - page=21, words=275-503, section=Robustness/Sensitivity via Appendix Tables, note=Online Appendix Table A4 shows robustness of results.
  - page=34, words=0-349, section=Instrument robustness and Appendix A9, note=See Online Appendix Table A9 for details.
  - page=38, words=275-377, section=Sensitivity and robustness checks, note=As shown, Online Appendix Table A11 provides similar estimates.
  - page=41, words=275-411, section=Specification robustness, note=Online Appendix Table A11 repeats this decomposition across alternatives.
  - page=48, words=0-323, section=Endogeneity robustness, note=Robustness check; endogeneity discussion; place effects estimated as no more than 3% of differences.

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

- Category: `G) External validity and generalization`
- Confidence: `low (0.2928615356894132)`
- Retrieval method: `local`
- Evidence type: `Data scope and generalizability discussion from the study methods and limitations.`
- Data source: Nielsen Homescan Panel (2004–2016) measuring household grocery purchases
- Assumption flag: `False`
- Assumption notes: Generalization is limited to U.S. take-home grocery purchases among Homescan-type households; not meals away from home; Homescan sample is not nationally representative; movers are not nationally representative.
- Answer:

Generalizable populations are U.S. households whose grocery purchases are captured in take-home retail data (e.g., the Nielsen Homescan panel) and who experience changes in local retail environments, such as moving across counties or living in areas with varying food retail access. The study's external validity is limited to such take-home grocery contexts in the United States and may not

extend to meals eaten away from home or non-U.S. settings.

- Quote snippet:

  We use the Nielsen Homescan Panel for 2004–2016 to measure household grocery purchases, and we focus on explaining income-related differences in the take-home market (i.e., groceries) rather than meals away from home.

- Citation anchors:
  - page=7, words=0-349, section=II.A. Nielsen Homescan and Retail Scanner Data, note=Describes use of Nielsen Homescan to measure household grocery purchases and focus on take-home market.
  - page=24, words=0-349, section=II.B. Endogeneity and Place Effects, note=Notes that movers are not nationally representative and weights are not used, limiting generalizability.
  - page=26, words=0-349, section=FIGURE V and related results, note=Illustrates within-county place effects and the scope of medium-term generalization to location changes.

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

- Category: `G) External validity and generalization`
- Confidence: `low (0.24063857935569227)`
- Retrieval method: `local`
- Evidence type: `data limitations and generalizability note`
- Data source: Nielsen Homescan Panel (2004–2016)
- Answer:

The study notes that its findings are least generalizable to the national population and to broader diets beyond take-home, packaged grocery purchases. Specifically, the data are not nationally representative and focus on home grocery purchases (packaged items), excluding food bought away from home.

- Quote snippet:

  Observations are not weighted for national representativeness.

- Citation anchors:
  - page=7, words=0-9, section=II.A. Nielsen Homescan and Retail Scanner Data, note=Observations are not weighted for national representativeness.
  - page=7, words=60-120, section=II.A. Nielsen Homescan and Retail Scanner Data, note=Homescan does not include data on food purchased away from home; focuses on take-home grocery purchases.

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

- Category: `G) External validity and generalization`
- Confidence: `low (0.24266437175378278)`

- Retrieval method: `local`
- Evidence type: `limitations and boundary conditions`
- Data source: Homescan panel data; RMS store data; ZBP data
- Assumption flag: `True`
- Assumption notes: Explicit assumption that endogeneity would generate upward bias in estimated place effects, enabling interpretation of results as upper bounds; acknowledges medium-term horizon as the scope of analysis.
- Answer:

Yes. The authors discuss boundary conditions and scope limits, acknowledging endogeneity that may bias estimates, a panel that is only long enough to study medium-term (a few years) effects, and that place effects are interpreted as upper bounds rather than definitive long-run conclusions.

- Quote snippet:

  > Any endogeneity in moving decisions likely biases these estimates upward.

- Citation anchors:
  - page=4, words=275-423, section=Limitations / Endogeneity, note=Endogeneity may bias estimates upward; panel length limited to a few years.
  - page=24, words=0-349, section=Scope of within-household design, note=Within-household design estimates place effects in the medium term; results are upper bounds on true effects.
  - page=48, words=0-323, section=Time horizon and bounds, note=Several-year time horizon; endogeneity discussed as a boundary condition; place effects bounded.

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

- Category: `G) External validity and generalization`
- Confidence: `medium (0.3592734269131566)`
- Retrieval method: `local`
- Evidence type: `study-design and results interpretation`
- Data source: Homescan data; RMS store data; Nielsen data (ZBP context mentioned); author analyses
- Assumption flag: `True`
- Assumption notes: Assumes within-market price variation identifies effects in a given year and that demand shocks are uncorrelated across markets (instrument validity).
- Answer:

Across different time periods or markets, the estimated effects may differ due to variations in market structure (how many and what types of retailers

are present), the degree of within-market price variation, and changes in where households move or shop. The study's results come from a 2004–2016 event-study using within-household fixed effects and market-level controls; thus, longer horizons or markets with different retailer mixes or price dynamics could yield larger or smaller expenditure-share effects and different Health Index responses. Heterogeneity by income and food-desert status also matters, with effects typically larger for low-income households in food deserts and most entrant gains reflecting diverted sales rather than new purchases. The identification rests on within-market price variation in a given year, so markets with little within-market variation or different competitive dynamics may produce different estimates, and the validity of the instrument relies on limited cross-market demand shocks being uncorrelated.

- Quote snippet:

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

- Citation anchors:
  - page=19, words=0-310, section=Figure IV Event Study; time period and controls, note=Time frame 2004–2016; controls and robustness notes
  - page=21, words=275-503, section=Results and heterogeneity, note=Expenditure-share changes larger for low-income and food deserts; most gains are diverted
  - page=24, words=0-349, section=Place effects; medium-term within-household design, note=Medium-term place effects; annual data; moves; fixed effects
  - page=32, words=0-349, section=Identification via within-market variation, note=Identification comes from within-market price variation across product groups in a market/year
  - page=34, words=0-385, section=Instrument validity and cross-market identification, note=Tests of instrument validity; cross-market identification; assumption about demand shocks

H01: Are key variables measured directly or via proxies?

- Category: `H) Measurement validity`
- Confidence: `medium (0.36465428277471224)`
- Retrieval method: `local`
- Evidence type: `Methods/Data description; direct measurements for some items; derived constructs for calorie needs; purchases as proxy for consumption.`
- Data source: Homescan household data and Nielsen RMS price data
- Table/Figure: Table I, Panel A
- Answer:

Key variables are measured both directly and via proxies: demographic and

health-related variables are collected directly from respondents (self-reports, BMI, diabetes status, nutrition knowledge quiz, income, age, race, employment, etc.), while calorie-related measures are derived or proxied (daily calorie needs computed from Dietary Guidelines and household composition; total calories consumed (Yijt) inferred from calories purchased).

- Quote snippet:

  self-reports of the importance of staying healthy, a detailed nutrition knowledge quiz, body mass index (weight/height2), and diabetes status

- Citation anchors:
  - page=8, words=0-349, section=Data sources / Measurements, note=Direct self-reports and health measures (BMI, diabetes status) and demographic variables.
  - page=30, words=0-349, section=Variables / Yijt, note=Total calories consumed (Yijt) defined from purchases – a proxy for actual consumption.

H02: What measurement error risks are acknowledged or likely?

- Category: `H) Measurement validity`
- Confidence: `low (0.25921969594910366)`
- Retrieval method: `local`
- Evidence type: `Data limitations and measurement error risks in store sampling and self-reported measures.`
- Data source: Homescan and Nielsen RMS data
- Answer:

Measurement error risks acknowledged or likely include: (a) RMS data do not form a complete census of stores, so the distribution of store types in RMS may not match the true county distribution, risking mismeasurement of the area Health Index; the authors address this by using a calorie-weighted Health Index and regression-adjusting for differences in store-channel type distributions; (b) Homescan data include self-reports of health-related variables (e.g., importance of staying healthy, BMI, diabetes status), which are subject to measurement error; (c) most Homescan households are observed for only a few years, limiting within-household estimates of place effects and potentially biasing medium-term assessments; (d) the area-average Health Index is adjusted for differences between RMS and true store distributions observed in ZBP data to mitigate sampling-based measurement bias.

- Quote snippet:

  the distribution of store types in the RMS sample may not match a county's true distribution.

- Citation anchors:

  - page=24, words=0-349, section=Store-level data limitations and measurement risk, note=RMS data do not contain the complete census of stores; distribution of store types may not match true county distribution.
  - page=24, words=275-452, section=Health Index construction, note=Calorie-weighted Health Index of RMS stores with regression-adjustment for distribution differences.
  - page=8, words=0-349, section=Data sources and self-reported measures, note=Includes self-reports of health-related variables (staying healthy importance, BMI, diabetes status).
  - page=24, words=0-349, section=Panel design limitations, note=Most Homescan households are observed for only a few years; within-household design estimates medium-term place effects.

H03: Are there validation checks for key measures?

- Category: `H) Measurement validity`
- Confidence: `low (0.2696332542228456)`
- Retrieval method: `local`
- Evidence type: `Validation of health/dietary measures (HEI, Health Index)`
- Data source: Nielsen Homescan data (household-by-year data, 2004–2016)
- Table/Figure: Online Appendix Table A1 and A2 referenced for validation of HEI and Health Index
- Answer:

Yes. They validate key measures by showing the linearized HEI is highly consistent with the true HEI (correlation 0.91) and that the Health Index correlates with BMI and diabetes; the Health Index is normalized to mean 0 and SD 1, with Online Appendix Tables A1–A2 providing further validation.

- Quote snippet:

  > We find that the linearization makes little difference: 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=HEI validation, note=Description of linearized HEI validation against true HEI (correlation 0.91).
  - page=11, words=275-399, section=HEI validation, note=Online Appendix Table A2 showing correlations of HEI/Health Index with BMI and diabetes.

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

- Category: `H) Measurement validity`
- Confidence: `low (0.23725491403165053)`

- Retrieval method: `local`
- Evidence type: `construct validity assessment via correlations with health indicators`
- Data source: Homescan (Nielsen) household data; RMS store purchase data; BMI/diabetes status used for validation
- Table/Figure: Online Appendix Tables A1 and A2
- Answer:

Yes. The authors assess construct validity of core outcomes by showing that the Health Index (their linearized HEI) correlates strongly with established health indicators (BMI and diabetes status), indicating the measure captures healthfulness of diets.

- Quote snippet:

  "Online Appendix Table A2 shows that 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=III. STYLIZED FACTS: PURCHASES AND SUPPLY OF HEALTHFUL FOODS III.A. Purchase Disparities: The Nutrition–Income Relationship, note=Discussion of linearized HEI/Health Index construction and validation
  - page=11, words=275-399, section=III. STYLIZED FACTS: PURCHASES AND SUPPLY OF HEALTHFUL FOODS III.A. Purchase Disparities: The Nutrition–Income Relationship, note=Evidence that both true HEI 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 (0.3293398427266181)`
- Retrieval method: `local`
- Evidence type: `Policy counterfactuals described in the article (Figure VI counterfactuals and Section VII.C).`
- Data source: Nielsen data and RMS store data (Nielsen Datasets at the Kilts Center; data center at UChicago Booth).
- Answer:

The article considers several policy counterfactuals: (i) equalizing prices across income groups for each product group; (ii) setting all groups' prices to the high-income group's prices; (iii) setting prices and product nutrient characteristics to those of the high-income group; (iv) setting nutrient preferences to those of the high-income group; (v) setting product group preferences to those of the high-income group; (vi) implementing a means-tested subsidy for healthy groceries (costing about 15% of the current SNAP budget) to raise low-income

households' healthy eating to the level of high-income households; (vii) considering supply-side approaches like encouraging supermarket entry (found to have limited effects); and (viii) a partial-equilibrium place-based counterfactual where a household moves from a low- to a high-income retail environment to bound the importance of location effects.

- Quote snippet:

  > Our first counterfactual measures the effect of prices by equalizing the prices that all income groups pay for each product group.

- Citation anchors:
  - page=40, words=We-categories., section=VI.C. Method of Moments Estimation, note=First counterfactual: price equalization across income groups for each product group.
  - page=40, words=The second column of points in-categories., section=VI.C. Method of Moments Estimation, note=Describes the five counterfactual categories (price-only; prices+nutrients; nutrient preferences; product group preferences).
  - page=48, words=In combination with the assumption-budget., section=VII.C. Using Subsidies to Reduce Nutritional Inequality, note=Means-tested subsidy for healthy groceries and forecasted cost.
  - page=27, words=We consider a partial equilibrium thought experiment-environment., section=Place effects, note=Implied policy-relevant counterfactual about mobility to higher-income retail environments.

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

- Category: `I) Policy counterfactuals and welfare`
- Confidence: `medium (0.3951286157637345)`
- Retrieval method: `local`
- Evidence type: `Policy and welfare analysis`
- Data source: The Quarterly Journal of Economics article 'Food Deserts and Nutritional Inequality'.
- Table/Figure: Table V
- Assumption flag: `True`
- Assumption notes: Endogeneity could generate upward bias in place-effects; results may understate the impact of supply/demand factors if assumptions do not hold.
- Answer:

The article finds that welfare gains from supply-side reforms (e.g., supermarket entry) are small; most nutritional inequality stems from demand differences and preferences, as well as nutrition knowledge. A means-tested subsidy for healthy groceries could substantially reduce the health index gap between high- and low-

income households, at a fraction of the SNAP budget; fully equalizing supply would reduce the gap by about 10%, while subsidies could achieve parity at roughly 15% of SNAP expenditures. Eliminating food deserts would generate little progress.

- Quote snippet:

  Our analyses suggest that supply-side policies, such as encouraging supermarket entry, will have limited effects on healthy eating.

- Citation anchors:
  - page=2, words=0-349, section=Introduction, note=Counterfactual: exposure to same products/prices reduces inequality by ~10%; demand differences drive ~90%.
  - page=43, words=0-349, section=VII.A/VII.C Intro, note=Introduces subsidies as alternative policy; supply-side effects found to be limited.
  - page=45, words=0-349, section=VII.C Subsidies, note=Table V presents subsidies to close health-index gaps; means-tested subsidies discussed.
  - page=46, words=0-349, section=VII.C Subsidies, note=Cost comparisons: subsidies vs supply-equalization; relates to SNAP budget.
  - page=37, words=0-349, section=WTP/Preferences by income, note=Higher-income households show different preferences; education and nutrition knowledge matter.

I03: Are cost-benefit or incidence analyses provided?

- Category: `I) Policy counterfactuals and welfare`
- Confidence: `low (0.2991275275287277)`
- Retrieval method: `local`
- Evidence type: `cost-effectiveness analysis`
- Data source: SNAP budget data (2016 SNAP budget of $71 billion) and subsidy cost figures discussed in the article
- Table/Figure: Table V
- Assumption flag: `True`
- Assumption notes: Assumes the demand estimates take the existing SNAP program as given and uses a partial-equilibrium framework; endogeneity concerns addressed by fixed effects and the claim that the subsidy analysis leaves SNAP unchanged otherwise.
- Answer:

Yes. The article provides a cost-effectiveness (policy cost-benefit) analysis of a means-tested healthy food subsidy, including explicit subsidy cost figures and comparisons to the SNAP budget.

- Quote snippet:

  Thus, our model suggests that adding a healthy food subsidy to SNAP could eliminate this measure of nutritional inequality at an

additional cost of only about 15% of the current SNAP budget.

- Citation anchors:
  - page=45, words=275-394, section=VII.A. Decomposing Consumption Differences into Supply versus Demand Factors, note=Subsidy impact on Health Index and related cost-effectiveness discussion.
  - page=46, words=0-349, section=VII. EXPLAINING AND REDUCING NUTRITIONAL INEQUALITY, note=Annual subsidy amounts ($84M, $830M, $11B) and comparison to SNAP budget.
  - page=46, words=275-416, section=VII. EXPLAINING AND REDUCING NUTRITIONAL INEQUALITY, note=15% of SNAP budget stated as cost to achieve supply equalization.

I04: What policy recommendations are stated or implied?

- Category: `I) Policy counterfactuals and welfare`
- Confidence: `low (0.332063795932424)`
- Retrieval method: `local`
- Evidence type: `Policy recommendation and analysis from VII.C section arguing subsidies for healthy foods are more cost-effective than supply-side supermarket entry policies.`
- Data source: The Quarterly Journal of Economics article 'Food Deserts and Nutritional Inequality' (Handbury, Rahkovsky, Schnell, 2019/2020).
- Assumption flag: `True`
- Assumption notes: Model-based subsidies analysis relies on several assumptions; results are presented as suggestive and motivation for demonstration projects.
- Answer:

The article recommends focusing on means-tested subsidies for healthy foods rather than pursuing supply-side policies like promoting supermarket entry; i.e., redirect policy toward subsidies to reduce nutritional inequality.

- Quote snippet:

  "Our analyses suggest that supply-side policies, such as encouraging supermarket entry, will have limited effects on healthy eating. In this section, we study an alternative policy: subsidies for healthy foods."

- Citation anchors:
  - page=43, words=275-387, section=VII.C. Using Subsidies to Reduce Nutritional Inequality, note=Policy recommendation: redirect from supply-side policies to means-tested subsidies for healthy foods.

J01: What missingness or attrition patterns are reported?

- Category: `J) Data quality and integrity`
- Confidence: `low (0.3176962075412535)`

- Retrieval method: `local`
- Evidence type: `text`
- Data source: Nielsen Homescan panel
- Answer:

The study reports missingness in the data by dropping observations with zero purchases. Specifically, 10.6% of observations (at the household-by–product group–by-year level) are dropped because they have zero purchases.

- 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=VI.C. Method of Moments Estimation, note=Statement that 10.6% of observations are dropped due to zero purchases.

J02: How are outliers handled (winsorization, trimming, exclusions)?

- Category: `J) Data quality and integrity`
- Confidence: `low (0.33098625056552644)`
- Retrieval method: `local`
- Evidence type: `Not explicitly described; explicit exclusions are mentioned but not outlier-specific methods.`
- Data source: Homescan and Nielsen RMS data (excluding liquor stores and nonpackaged items)
- Answer:

Outliers are not described in the provided context; the text mentions exclusions such as liquor stores and nonpackaged items but does not specify any winsorization, trimming, or other outlier handling.

- Quote snippet:

  We exclude liquor stores.

- Citation anchors:
  - page=8, words=0-349, section=Data sources and sample exclusions, note=Excludes liquor stores; RMS excludes bulk/nonpackaged items.
  - page=8, words=275-393, section=Nielsen RMS data and exclusions, note=RMS does not include sales of bulk produce and other nonpackaged items.

J03: Are there data audits or validation steps described?

- Category: `J) Data quality and integrity`
- Confidence: `low (0.23150430471718103)`
- Retrieval method: `local`

- Evidence type: `data validation / audit`
- Data source: TDLinx store-entry data; validated against four retailers' administrative records
- Assumption flag: `False`
- Answer:

Yes. The article describes data validation checks comparing retailer-entry data (TDLinx) to independent records, validating store openings (1,914 supermarkets) and using cleaning steps to drop spurious entrants; this constitutes a data audit/validation.

- Quote snippet:

  In validation checks, we found that TDLinx data closely match entry dates and locations for 1,914 supermarkets

- Citation anchors:
  - page=page 10 words 0-349, words=0-349, section=II.B. Grocery Retail Establishments Studying, note=Validation checks comparing TDLinx store-entry data with four retailers' records; 1,914 supermarkets confirmed; spurious entrants excluded.
  - page=page 10 words 275-424, words=275-424, section=II.B. Grocery Retail Establishments Studying, note=Further data cleaning to drop potentially spurious entries where TDLinx shows a store of the same subtype in the same census block in the previous year.

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

- Category: `J) Data quality and integrity`
- Confidence: `low (0.2959580243215666)`
- Retrieval method: `local`
- Evidence type: `Data limitations and sample representativeness; selective sample inclusion and measurement bias.`
- Data source: Nielsen Homescan panel data (2004–2016)
- Answer:

Yes. The study notes that the Homescan panel is not nationally representative, that movers are not a nationally representative group, and that weights are not used; plus data limitations restrict to packaged foods and exclude away-from-home purchases, indicating selective sample inclusion and measurement limitations.

- Quote snippet:

  Observations are not weighted for national representativeness.

- Citation anchors:
  - page=7, words=0-349, section=Data limitations, note=Homescan excludes away-from-home purchases and relies on UPC-packaged items.

  - page=21, words=275-503, section=Representativeness / Sampling, note=Observations are not weighted for national representativeness; fixed effects used.
  - page=24, words=0-349, section=Sample representativeness of movers, note=Discussion of movers and use (or non-use) of weights; sampling biases due to non-representative movers.

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

- Category: `K) Model fit and diagnostics`
- Confidence: `low (0.26338528421278345)`
- Retrieval method: `local`
- Evidence type: `Table III diagnostics in the QJE article (association of Health Index with local area Health Index using movers).`
- Data source: Nielsen Homescan data (2004–2016)
- Table/Figure: Table III
- Answer:

Table III reports regression fit and diagnostic metrics: regression R-squared values, robust standard errors clustered by household and local area, and 95% confidence-interval upper bounds; significance levels are shown with stars for 10%, 5%, and 1%.

- Quote snippet:

  Robust standard errors, clustered by household and census tract, are in parentheses. $*$, $**$, $***$: statistically significant with 10%, 5%, and 1% confidence, respectively.

- Citation anchors:
  - page=27, words=0-349, section=TABLE III: ASSOCIATION OF HEALTH INDEX WITH LOCAL AREA HEALTH INDEX USING MOVERS, note=Mentions regression R2, robust standard errors, and 95% CI upper bounds; includes 0.100/0.097/0.217/0.220 as CIs.

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

- Category: `K) Model fit and diagnostics`
- Confidence: `low (0.29902457783586855)`
- Retrieval method: `local`
- Evidence type: `textual evidence from model description and estimation sections`
- Data source: Nielsen Homescan data (Homescan add-on survey used by Allcott, Lockwood, and Taubinsky 2019).
- Assumption flag: `True`
- Assumption notes: Identification assumption: E[ijt Pjmt] = 0; the text does not report tests of alternative functional forms for the model.

- Answer:

Functional form choices are not reported as being tested; the paper adopts a Cobb-Douglas structure with a log-linear calorie demand specification (ln Yi,t), and there is no description of alternative functional forms being empirically tested.

- Quote snippet:

  ln Yi jt = − ln p̃i jt − ˜c ãi jct − + j + m + t + i jt , c=2

- Citation anchors:
  - page=28, words=275-452, section=Model specification, note=States Cobb-Douglas preferences over groups; log-linear structure implied by utility over calories.
  - page=30, words=0-349, section=VI.A Empirical Model, note=Presents the log-linear calorie demand equation ln Yi jt = - ln p̃i jt - ˜c ãi jct - + j + m + t + i jt.
  - page=29, words=0-349, section=Model description, note=Describes Cobb-Douglas preferences over product groups and the implicit functional form.
  - page=33, words=275-397, section=Identification, note=Outlines the identifying assumption E i jt Pjmt = 0 and instrument validity discussion.
  - page=34, words=0-349, section=Estimation methods, note=Method of Moments/Data construction; no reference to testing alternative functional forms.

K03: Are residual checks or specification tests reported?

- Category: `K) Model fit and diagnostics`
- Confidence: `low (0.2418042923910048)`
- Retrieval method: `local`
- Evidence type: `instrument validity/specification tests`
- Data source: Homescan and Nielsen RMS data
- Table/Figure: Online Appendix Table A9
- Answer:

Yes. The authors report two specification tests of the instrument's validity (predicting tastes using alternative demographic-based predictions).

- Quote snippet:

  We implement two tests, each of which predicts tastes in a different way. First, we predict purchases of product group j using household demographics and then predict county-level purchases on the basis of county average demographics.

- Citation anchors:

  - page=page 33, words=275-397, section=VI.C. Method of Moments Estimation, note=Tests of instrument validity; two tests predicting tastes in different ways.

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

- Category: `K) Model fit and diagnostics`
- Confidence: `medium (0.4197420472585346)`
- Retrieval method: `local`
- Evidence type: `text`
- Data source: Online Appendix Tables A9 and A11; main text
- Assumption flag: `True`
- Assumption notes: Exogeneity assumption: E[ijt Pjmt] = 0; robustness tests show the instrument is not systematically related to predictable tastes, supporting the validity of the exclusion restriction.
- Answer:

Results are robust to a range of alternative specifications and estimators; adding demographic controls, county-by-quarter fixed effects, and different fixed effects specifications yields estimates that are very similar or unchanged, and tests of the price instrument's validity do not indicate violation of the exclusion restriction.

- 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=Robustness checks, note=Adding county-by-quarter fixed effects or a control for convenience store entry does not qualitatively change the results.
  - page=27, words=275-482, section=Robustness checks, note=Including household demographics has very little impact on results.
  - page=33, words=275-397, section=Instrument validity tests, note=Two tests predict no relation between instrument and predictable tastes; identifying assumption discussed.
  - page=34, words=0-349, section=Instrument validity and identification, note=Discussion of novel instrument and exogeneity tests.
  - page=38, words=275-377, section=Robustness of estimates, note=Online Appendix Table A11: estimates are all very 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 and samples used
  - Homescan panel data (demographics, calories, health measures) and Nielsen RMS price/sales data (2006–2016; about 35,000 stores observed per year; exclude liquor; channels: grocery, mass merchandiser, drug, convenience). All Homescan results are weighted for national representativeness. Sources and caveats noted (Nielsen data are the researchers' data and Nielsen had no role in analysis). (Context: page 1, words 275–483; page 8, words 0–349; page 8, 275–393)
- How prices are treated
  - Prices from RMS are deflated to 2010 USD; RMS excludes bulk produce and other nonpackaged items. (Context: page 8, 0–349; page 8, 275–393)
- Main empirical goal
  - Decompose differences in healthy eating across income groups into supply-side (prices, availability) and demand-side (preferences) factors. Robustness checks with census-region-by-product fixed effects and urban/rural interactions. (Context: page 41, 275–411; page 27, 275–482)
- Key result: share of nutrition–income gap explained by demand vs supply
  - Across specifications, supply explains roughly 7%–12% and demand explains roughly 88%–93% of the nutrition–income relationship; about 90% is due to demand, ~10% to supply. (Context: page 41, 275–411)
- Methodological approach to endogeneity and location
  - Use a partial equilibrium thought experiment and geographic variation to bound the role of location. Online Appendix figures map geographic presence of retailers; produce prices are predicted to be cheap on the West Coast and expensive on the East Coast (addressing endogeneity via geographic/instrumental variation). (Context: page 27, 275–482; page 32, 275–431)
- Illustrative empirical result (Coke vs. Pepsi)
  - An empirical exercise using county Coke market shares shows a statistically significant^ 0.16: moving to a county with a 10 percentage

point higher Coke share is associated with a 1.6 percentage point increase in Coke's share within Coke+Pepsi purchases. (Context: page 27, 275–482)

- Instrumental/identification notes
  - The study constructs a multi-faceted instrument based on geographic price variation across major retailers (the five largest RMS retailers), with online Appendix figures illustrating these variations and their implication for identification. (Context: page 32, 275–431)
- Notes on data interpretation and scope
  - The conclusions reflect the researchers' analyses of Nielsen data and are not Nielsen's statements; results are presented with several robustness checks, including alternative specifications and fixed effects. (Context: page 1, 275–483)
- Example of reported nutrient preferences by income (illustrative table)
  - Table IV shows estimated nutrient preferences by income quartile across several nutrients and health indices; units and interpretation are provided (e.g., WTP for health index). (Context: page 36, 0–349; page 38, 275–377)

If you'd like, I can extract a concise formal summary (with equations) or tailor the bullets to a specific subsection (e.g., data, estimation, or robustness results).