Audit Report: Workflow d9a0a007cbf54c8881e328aab6d6cf0f

Overview

- Source JSON: reports\workflow-report-d9a0a007cbf54c8881e328aab6d6cf0f.json
- Run ID: d9a0a007cbf54c8881e328aab6d6cf0f
- Papers input: papers\Calorie Posting in Chain Restaurants - Bollinger et al. (2011).pdf
- Started at: 2026-02-16T04:22:51.161738+00:00
- Finished at: 2026-02-16T04:31:33.697368+00:00
- Duration: 0:08:42.535630

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: indexed
- report_store: pending

Agentic Summary

- Status: completed
- Main question: What are the paper's main contribution, identification strategy, key results, and limitations?
- Report question set: both
- Structured questions generated: 84
- Confidence mean/median: 0.27970699443533575 / 0.28419734314792333
- Confidence labels: low=74, medium=9, high=0

Final Answer

- Main contributions
  - Empirically assesses NYC's mandatory calorie posting on chain restaurants using two datasets (store-day transaction data and

cardholder data) in NYC vs control cities (Boston/Philadelphia); identification via a post-policy indicator and fixed effects/controls. (page 13, 275-528)
  - Decomposes effects into extensive vs. intensive margins, showing calories drop primarily through reduced food purchases (extensive margin) and, to a lesser extent, lower calories per item. (page 18, 0-349)
  - Provides distributional and heterogeneity analyses, showing larger reductions for higher-calorie consumers and higher quantiles; differential effects by income, education, etc. (pages 21-22, 0-349)
  - Includes external validation via a Seattle placebo, suggesting beverage-focused effects and limited noise from the policy elsewhere. (page 15, 0-254)

- Identification strategy
  - Difference-in-differences design comparing NYC to control cities with a POST indicator (1 after April 1, 2008), using fixed effects and additional controls; evidence drawn from store-day transaction data and cardholder data. (page 13, 275-528)

- Key results
  - Overall effect: calories per store-day fall by about 4.6; calories per transaction fall less, consistent with more total transactions after posting. (page 13, 275-528)
  - Item- and per-item effects (Table 4): beverages per transaction up modestly (0.005***); food items per transaction down (−0.029***); calories per beverage and per food item both decline (log beverage calories per beverage: −0.008***; log calories per food item: −0.039***). (page 18, 0-349)
  - Mechanism: the calorie reduction is driven by both extensive margin (fewer food items purchased) and intensive margin (lower calories per item), with food items contributing notably. (page 18, 0-349)
  - Main channel: extensive margin dominates—about three-quarters of the total calorie reduction comes from not purchasing food items. (page 18, 0-349)
  - Distributional and heterogeneity patterns: larger reductions at higher quantiles; percent change roughly 5–6% from 75th to 99th percentile; effects differ by income, education, age, gender, and customer type (Posting × interactions). (pages 21-22, 0-349)
  - Summary takeaway: mandatory posting reduces calories per transaction mainly via reduced food purchases (extensive margin) and, to a lesser extent, lower calories per item; effects stronger for higher-calorie consumers and at higher consumption quantiles. (page 18, 0-349; pages 21-22, 0-349)

- Limitations

  - Limitations are not explicitly enumerated in the provided excerpts; robustness checks include a Seattle placebo, which supports external validity to some extent but does not fully generalize beyond the studied contexts. (Seattle placebo: page 15, 0-254)

Sub-Answers

Sub-answer 1

- Question: ResponseTextConfig(format=ResponseFormatText(type='text'), verbosity='medium')
- Question tokens estimate: `19`
- Answer:
- Objective and data
  - Analyzes the impact of NYC's mandatory calorie posting on chain restaurants using two datasets (store-day transaction data and cardholder data) in NYC vs control cities (Boston/Philadelphia). POST = 1 after April 1, 2008; results rely on NYC vs controls and fixed effects/controls. (page 13, 275-528)
- Overall impact on calories
  - Calorie posting reduces average calories per store-day by about 4.6 calories, with the reduction in calories per transaction smaller, consistent with an increase in total transactions after posting. (page 13, 275-528)
- Seattle placebo comparison (external validation)
  - In Seattle, beverage calories per transaction fell by 4.6 calories, while food calories per transaction rose by 0.8 calories, suggesting beverage-focused effects and little noise from the posting policy elsewhere. (page 15, 0-254)
- Item- and per-item effects
  - Table 4 (transaction data) shows: number of beverages per transaction increases modestly (0.005***); number of food items per transaction falls (−0.029***); calories per beverage and per food item also decline (log beverage calories per beverage: −0.008***; log food calories per food item: −0.039***). (page 18, 0-349)
  - Interpretation: the reduction in calories is driven both by fewer items purchased (extensive margin with food items) and by lower calories per item (intensive margin). Food items contribute notably to the effect. (page 18, 0-349)
- Main channel: extensive margin dominates

  - About three quarters of the total calorie reduction is attributed to people opting not to buy food items (extensive margin of food demand). (page 18, 0-349)

- Distributional and heterogeneity patterns
  - Quantile evidence: calorie reductions are larger for higher quantiles; the percent change is roughly stable around 5–6 percent from 75th to 99th percentile, though absolute decreases are larger at higher quantiles. (page 22, 0-349)
  - Heterogeneity: effects differ by characteristics (e.g., income, education, age, gender, customer type); e.g., higher-income and more-educated groups show additional differential effects (Posting × median income, Posting × percent with college degree, etc.). (page 21, 0-349)

- Summary takeaway
  - Mandatory calorie posting reduces calories per transaction primarily through reduced food purchases (extensive margin) and, to a lesser extent, lower calories per item; effects are larger for higher-calorie consumers and more pronounced at higher consumption quantiles. (page 18, 0-349; page 21-22, 0-349)

Sources used:

- NYC calorie posting effects and overall magnitude: page 13, words 275-528
- Seattle placebo example: page 15, words 0-254
- Table 4 item-level and margin findings: page 18, words 0-349
- Extensive-margin interpretation: page 18, words 0-349
- Distributional/heterogeneity results: pages 21-22, words 0-349

Structured Q&A Appendix

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

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

- Category: `A) Research question / contribution`
- Confidence: `low (0.3160348019347735)`
- Retrieval method: `local`
- Evidence type: `explicit research question stated in the introduction`
- Data source: Transaction data and cardholder data from Starbucks in NYC, Boston, and Philadelphia
- Answer:

Whether mandatory calorie posting affects consumers' behavior, specifically calories per transaction, at Starbucks.

- Quote snippet:

  Ultimately, whether calorie posting affects people's behavior is an empirical question.

- Citation anchors:
  - page=3, words=275-296, section=Introduction, note=Location of the explicit empirical question about calorie posting effects.

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

- Category: `A) Research question / contribution`
- Confidence: `low (0.29760137826352995)`
- Retrieval method: `local`
- Evidence type: `textual_evidence`
- Data source: Starbucks transaction data; cardholder data; in-store consumer surveys
- Answer:

The paper's primary contribution is providing a large-scale, long-horizon analysis of calorie posting, using a dataset far larger and broader than prior work (over 100 million Starbucks transactions across NYC, Boston, and Philadelphia, plus 1.5 million cardholder observations and surveys), enabling examination of the time path, product substitution, heterogeneity in responses, and profits, and clarifying learning and salience as mechanisms.

- Quote snippet:

  the dataset we study is much larger and broader—the universe of over 100 million transactions at Starbucks in Boston, NYC, and Philadelphia over a 14-month period. We also analyze individual-level data (1.5 million transactions of anonymous customers over time), as well as a survey that focuses on testing consumers' knowledge).

- Citation anchors:
  - page=8, words=0-349, section=Comparison with prior research and data contributions, note=States dataset is much larger/broader and includes multiple data sources, enabling long-run analysis and examination of substitution, heterogeneity, and profits.

A03: What is the central hypothesis being tested?

- Category: `A) Research question / contribution`
- Confidence: `low (0.2901430171019807)`
- Retrieval method: `local`
- Evidence type: `empirical analysis of transaction/cardholder data`

- Data source: Starbucks transaction data and cardholder data
- Answer:

The central hypothesis is that mandatory calorie posting at chain restaurants reduces calories per transaction (i.e., lowers calories purchased), and that this effect may operate through learning about calories or through increased salience of calorie information.

- Quote snippet:

  > Ultimately, whether calorie posting affects people's behavior is an empirical question.

- Citation anchors:
  - page=3, words=275-564, section=Introduction – empirical question and mechanisms, note=States that the effect of calorie posting is an empirical question and discusses learning vs salience.
  - page=29, words=0-349, section=Results – commuters and NYC effects, note=Presents results showing changes in calories per transaction for commuters and non-commuters.
  - page=30, words=0-349, section=Results – learning vs salience conclusions, note=Discusses interpretation: both learning and salience contribute.

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

- Category: `A) Research question / contribution`
- Confidence: `low (0.3047339608395138)`
- Retrieval method: `local`
- Evidence type: `definition of dependent variables from results section`
- Data source: transaction data and cardholder data
- Answer:

The main outcomes of interest are calories per transaction measures: beverage calories per transaction, food calories per transaction, and total calories per transaction (log forms).

- Quote snippet:

  > the dependent variable is log(beverage calories per transaction). Based on the transaction data we estimate that calorie posting caused a trivial decrease in beverage calories per transaction.

- Citation anchors:
  - page=12, words=0-349, section=Table 3 – Regression results, note=Defines dependent variables: log(beverage calories per transaction), log(food calories per transaction), and log(beverage + food calories) per transaction.

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

- Category: `A) Research question / contribution`
- Confidence: `low (0.24409435786648528)`
- Retrieval method: `local`
- Evidence type: `regression specification / treatment variable definition`
- Data source: Transaction data and cardholder data (Starbucks)
- Answer:

POST, the calorie-posting dummy variable equal to one for NYC stores after April 1, 2008, is the key treatment/exposure variable.

- Quote snippet:

  y_sct = x_sct + POS Tc t + _sct, where y_sct is a measure of calories per transaction at store s in city c on day t, POS Tc t is a dummy equal to one if calories were posted (i.e., NYC stores after April 1, 2008)

- Citation anchors:
  - page=11, words=0-349, section=Regression specification / Methods, note=Defines the POST variable as a dummy equal to one if calories were posted (NYC stores after April 1, 2008).

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

- Category: `A) Research question / contribution`
- Confidence: `medium (0.3533244294259882)`
- Retrieval method: `local`
- Evidence type: `textual`
- Data source: Starbucks transaction data and cardholder data from NYC, Boston, and Philadelphia (USA).
- Table/Figure: Table 1—Summary Statistics for Transaction Data and Cardholder Data
- Answer:

United States (New York City as the primary setting) in the chain-restaurant sector (Starbucks), examining the NYC calorie-posting law during 2008–2009 (with NYC stores compared to Boston and Philadelphia).

- Quote snippet:

  Our transaction data cover all 222 Starbucks locations in NYC, and all 94 Starbucks locations in Boston and Philadelphia.

- Citation anchors:

- page=5, words=0-40, section=A. Data Summary, note=Describes NYC/Boston/Philadelphia coverage and the 2008–2009 period

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

- Category: `A) Research question / contribution`
- Confidence: `low (0.28587559983424177)`
- Retrieval method: `local`
- Evidence type: `textual claim from article`
- Data source: Starbucks transaction and cardholder data
- Answer:

A combination of learning and salience effects.

- Quote snippet:

  a combination of learning and salience effects.

- Citation anchors:
  - page=page 25, words=0-349, section=A. Why is There an Effect?, note=Authors state that behavior may be driven by a combination of learning and salience effects.

A08: What alternative mechanisms are discussed?

- Category: `A) Research question / contribution`
- Confidence: `low (0.22576387223949634)`
- Retrieval method: `local`
- Evidence type: `textual`
- Data source: Bollinger et al., Calorie Posting in Chain Restaurants, American Economic Journal: Economic Policy
- Answer:

The document discusses two alternative mechanisms: a learning effect and a salience effect (often in combination).

- Quote snippet:

  One reason why calorie posting may affect consumer choice is a learning effect. Another possible explanation for the observed reduction in calories per transaction is a salience effect.

- Citation anchors:
  - page=25, words=One-effects., section=A. Why is There an Effect?, note=Mentions learning and salience as explanations; notes behavior may be driven by a combination.
  - page=25, words=Another-salience effects., section=A. Why is There an Effect?, note=Explicit mention of salience effect as an alternative mechanism.

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

- Category: `A) Research question / contribution`
- Confidence: `medium (0.3767458184348189)`
- Retrieval method: `local`
- Evidence type: `Discussion and Conclusion (policy implications) from the paper`
- Data source: Starbucks NYC transaction data, cardholder data, and in-store surveys
- Answer:

Mandatory calorie posting across chain restaurants reduces calories per transaction (mainly by lowering food calories) with little effect on beverage calories or visit frequency; while the direct obesity benefits may be small, the costs are very low and the policy could yield long-run benefits such as menu innovation toward lower-calorie options and public education, making it potentially worthwhile, especially since both learning and salience appear to drive effects.

- Quote snippet:

  > mandatory calorie posting caused food calories per transaction to fall by 14 percent, on average, but had a negligible impact on beverage calories per transaction.

- Citation anchors:
  - page=24, words=275-564, section=IV. Discussion, note=Contains the main policy implications and discussion of learning/salience, costs, and potential benefits.
  - page=34, words=0-349, section=IV. Discussion, note=Concludes with the summary of findings and policy relevance, including low costs and potential long-run effects.

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

- Category: `A) Research question / contribution`
- Confidence: `low (0.31027249337452484)`
- Retrieval method: `local`
- Evidence type: `policy implications / welfare interpretation based on results`
- Data source: Starbucks calorie posting study; transaction and cardholder data in NYC vs control cities
- Table/Figure: Table 9
- Assumption flag: `True`
- Assumption notes: Assumes reductions in calories translate into health/obesity benefits; obesity outcomes not measured in the study
- Answer:

The welfare interpretation is that calorie posting may yield a welfare gain through informing consumers and reducing calorie intake (especially from high-

calorie foods) with little impact on visit frequency or profits, suggesting health-oriented benefits with minimal consumer or producer distortion. However, the paper emphasizes that obesity outcomes are not measured, so the welfare impact remains uncertain and contingent on whether reduced calories translate into meaningful health gains.

- Quote snippet:

  > IV. Discussion To summarize briefly, the analyses above show that mandatory calorie posting caused food calories per transaction to fall by 14 percent, on average, but had a negligible impact on beverage calories per transaction.

- Citation anchors:
  - page=24, words=275-564, section=IV. Discussion, note=Summary of welfare-relevant results: calories per transaction fell; beverage calories largely unaffected; discussion of potential welfare implications and need for obesity data

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

- Category: `A) Research question / contribution`
- Confidence: `low (0.27592370260423454)`
- Retrieval method: `local`
- Evidence type: `textual limitation`
- Answer:

The main limitations are (1) obesity effects are not directly measured because BMI data are not yet available, so they cannot link calorie posting to obesity outcomes, and (2) the analysis uses data from only one chain (Starbucks), making it unclear whether the findings generalize to other chains; plus the possibility that some effects could be offset by changes in eating at home.

- Quote snippet:

  > There are two important limitations to this research. First, we do not directly measure the effect of calorie posting on obesity itself. A second limitation is that we have data for only one chain (Starbucks).

- Citation anchors:
  - page=3, words=275-564, section=limitations, note=Two main limitations: no direct obesity measurement and data for only one chain; potential offset by home eating.

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

- Category: `A) Research question / contribution`
- Confidence: `low (0.3097688817473264)`
- Retrieval method: `local`

- Evidence type: `data/identification novelty claim`
- Data source: Starbucks transaction data and anonymous cardholder data
- Answer:

The paper claims its novelty lies in using cardholder-level data to identify how calorie information affects individual behavior and to separate intensive and extensive margins from changes in consumer composition, i.e., novel data and identification.

- Quote snippet:

  > The advantage of the cardholder data is that we can assess how the calorie information causes particular individuals to change behavior, isolating intensive and extensive margins from changes in consumer composition.

- Citation anchors:
  - page=5, words=0-349, section=Data and identification novelty, note=States that the advantage of the cardholder data allows isolating effects on intensive and extensive margins from consumer composition.
  - page=3, words=275-564, section=Data suitability, note=Notes that the detailed transaction data are uniquely well-suited to answering the question.

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.27842383323913433)`
- Retrieval method: `local`
- Evidence type: `regression-based identification using within-city variation (difference-in-differences)`
- Data source: Transaction data and cardholder data from Starbucks in NYC, Boston, and Philadelphia.
- Assumption flag: `True`
- Assumption notes: Parallel trends assumption: pre-treatment trends in NYC and control cities are similar; identification relies on NYC changes following posting being exogenous relative to controls.
- Answer:

The identification uses a within-city variation over time (difference-in-differences) by comparing NYC before and after calorie posting to Boston/Philadelphia controls, attributing changes to posting.

- Quote snippet:

  > identification of the effect of calorie posting stems from within-city variation over time.

- Citation anchors:

  - page=11, words=275-522, section=Estimation/Identification, note=Statement that identification stems from within-city variation over time.

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.3052100415762042)`
- Retrieval method: `local`
- Evidence type: `Describes treatment/control groups and post-treatment indicator in the methods/regression specification.`
- Data source: Transaction data and cardholder data from Starbucks across NYC, Boston, Philadelphia, and Seattle
- Table/Figure: Table 2
- Assumption flag: `False`
- Answer:

Quasi-experimental design (a natural experiment using treatment and control cities with before/after comparison, estimated via regression with a post-treatment indicator).

- Quote snippet:

  POS_Tc t is a dummy equal to one if calories were posted (i.e., NYC stores after April 1, 2008)

- Citation anchors:
  - page=11, words=0-349, section=II. Effect of Mandatory Calorie Posting on Calorie Consumption, A. Calories Per Transaction, note=Shows the regression specification including POS_Tc t and the treatment/control framework.
  - page=7, words=275-566, section=Methods/Design, note=Describes Seattle vs NYC setup with control cities and the differential posting law.

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.2985466131536938)`
- Retrieval method: `local`
- Evidence type: `Identification strategy description (policy-induced exogenous shock with within-city, city-week variation; difference-in-differences using NYC as treated and Boston/Philadelphia as controls).`
- Data source: Starbucks transaction data and anonymized cardholder data for NYC, Boston, and Philadelphia (with Seattle as robustness check).

- Answer:

The source of exogenous variation is the timing of calorie posting in New York City (NYC) starting April 1, 2008, which creates a treatment (NYC) versus control (Boston and Philadelphia) setup. Identification relies on within-city variation over time (city-week level variation) to estimate the effect of posting.

- Quote snippet:

  > In both cases (transaction data and cardholder data), identification of the effect of calorie posting stems from within-city variation over time.

- Citation anchors:
  - page=11, words=275-522, section=Data/Identification, note=NYC calorie posting commenced on April 1, 2008; treated NYC vs controls for identification.
  - page=12, words=0-349, section=Identification, note=Identification stems from within-city variation over time; policy variation at the city-week level.

B04: What is the treatment definition and timing?

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.13190735780502477)`
- Retrieval method: `local`
- Evidence type: `textual evidence (methods: treatment definition and timing)`
- Data source: Starbucks transaction and cardholder data (NYC, Boston, Philadelphia)
- Answer:

Treatment is the NYC calorie-posting intervention, defined by a post-period indicator that equals one for NYC stores after the posting date (April 1, 2008).

- Quote snippet:

  > POS Tc t is a dummy equal to one if calories were posted (i.e., NYC stores after April 1, 2008)

- Citation anchors:
  - page=11, words=0-349, section=II.A - Calories Per Transaction, note=Definition of treatment: POS is a dummy equal to one if calories were posted (i.e., NYC stores after April 1, 2008).

B05: What is the control/comparison group definition?

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.1800548992326989)`
- Retrieval method: `local`

- Evidence type: `Textual evidence from the study describing the control group.`
- Data source: NYC vs Boston/Philadelphia transaction and cardholder data; Seattle robustness check with Portland/San Francisco as controls.
- Answer:

The control/comparison group consists of Boston and Philadelphia, used as controls for NYC.

- Quote snippet:

  Qualitatively, however, it appears that Boston and Philadelphia are reasonable controls for NYC.

- Citation anchors:
  - page=6, words=0-349, section=Table 1—Summary Statistics; Control Group, note=States that Boston and Philadelphia are used as controls for NYC.
  - page=14, words=275-590, section=Robustness checks with Seattle; control cities, note=Mentions Seattle's control cities Portland and San Francisco used in robustness checks.

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.30863926179799794)`
- Retrieval method: `local`
- Evidence type: `Baseline regression specification (Equation 1) from the paper`
- Data source: Bollinger et al. 2011, American Economic Journal: Economic Policy
- Answer:

The baseline estimating equation is y_sct = x_sct + POS_Tc t + _sct, where y_sct is a measure of calories per transaction at store s in city c on day t, POS_Tc t is a dummy equal to one if calories were posted (i.e., NYC stores after April 1, 2008), and x_sct includes week fixed effects (to control for seasonality), day-of-week dummies, holiday dummies, temperature and temperature squared, and precipitation and precipitation squared. In the transaction data, store fixed effects are included; in the cardholder data, individual fixed effects are included.

- Quote snippet:

  y_sct = x_sct + POS Tc t + _sct, where y_sct is a measure of calories per transaction at store s in city c on day t

- Citation anchors:
  - page=11, words=275-522, section=Estimation framework / Regression specification, note=Equation (1): y_sct = x_sct + POS Tc

t + _sct; y_sct is a measure of calories per transaction; POS Tc t is a dummy equal to one if calories were posted (NYC stores after April 1, 2008); x_sct includes week fixed effects (to control for seasonality), day-of-week dummies, holiday dummies, temperature and temperature^2, precipitation and precipitation^2; store fixed effects used with transaction data; individual fixed effects used with cardholder data.

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.25447351634256876)`
- Retrieval method: `local`
- Evidence type: `Description of fixed-effects structure from regression specifications across the cited pages.`
- Data source: American Economic Journal: Economic Policy, February 2011
- Answer:

Unit fixed effects: store fixed effects (transaction data) and individual fixed effects (cardholder data). Time fixed effects: week fixed effects, day-of-week dummies, and holiday dummies; robustness check adds date fixed effects. Two-way fixed effects: store fixed effects with week fixed effects. Higher-dimensional fixed effects: date fixed effects in robustness checks (drop weekly/daily effects).

- Quote snippet:

  Store fixed effects to control for all time-invariant, store-specific heterogeneity. Store fixed effects also control for time-invariant city characteristics. In the cardholder data we include individual consumer fixed effects (and drop the store fixed effects).

- Citation anchors:
  - page=11, words=275-330, section=Regression specification and time controls, note=Mentions week fixed effects to control for seasonality, day-of-week dummies, holiday dummies, and weather controls.
  - page=12, words=0-60, section=Fixed effects in store vs cardholder specifications, note=Notes store fixed effects for time-invariant, store-specific heterogeneity; cardholder data use individual fixed effects (and drop store fixed effects).
  - page=12, words=60-120, section=Robustness checks and additional fixed effects, note=Robustness check includes date fixed effects (instead of day-of-week/week fixed effects).

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.2807844345302521)`
- Retrieval method: `local`
- Evidence type: `textual`
- Data source: cardholder data (Table 9)
- Table/Figure: Table 9
- Answer:

They used robust (heteroskedasticity-robust) standard errors, reported in parentheses; the regression is at the transaction (observation) level (each observation is a transaction). The notes do not indicate any clustering.

- Quote snippet:

  Robust standard errors in parentheses.

- Citation anchors:
  - page=29, words=0-349, section=Table 9 notes, note=Robust standard errors in parentheses; observation is a transaction

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.2723451096565549)`
- Retrieval method: `local`
- Evidence type: `parallel_trends`
- Data source: Starbucks transaction data and cardholder data
- Table/Figure: Figure 2
- Assumption flag: `True`
- Assumption notes: The validity relies on parallel trends between NYC and the control cities; pre-trend similarity is reported, and results are corroborated across data sources and with within-individual identification.
- Answer:

Parallel trends assumption (differences-in-differences): the treated NYC stores would have followed the same time trend in calories per transaction as the control cities (Boston/Philadelphia) in the absence of calorie posting.

- Quote snippet:

  First, with both datasets we see no evidence of pre-trend differences between NYC and Boston/Philadelphia.

- Citation anchors:
  - page=14, words=0-60, section=Pre-trend checks, note=No evidence of pre-trend differences between NYC and control cities

(Boston/Philadelphia).
  - page=14, words=275-590, section=Identification & robustness checks, note=Within-individual variation; robustness checks across datasets; pre-trends discussed.

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.24037184619592758)`
- Retrieval method: `local`
- Evidence type: `parallel-trends check; within-individual variation design`
- Data source: Starbucks transaction data and cardholder data
- Table/Figure: Figure 2
- Assumption flag: `True`
- Assumption notes: Pre-trend equivalence across NYC and control cities; identification supported by within-individual variation in cardholder data to isolate intensive-margin effects.
- Answer:

Evidence supports the identifying assumption by showing no pre-trend differences between NYC and control cities (parallel trends) in both transaction and cardholder analyses, and by using cardholder data that isolates intensive-margin changes via within-individual variation over time.

- Quote snippet:

  no evidence of pre-trend differences between NYC and Boston/Philadelphia

- Citation anchors:
  - page=14, words=0-349, section=Figure 2 results, note=No pre-trend differences between NYC and controls; supports parallel trends.
  - page=12, words=275-624, section=Cardholder vs Transaction analysis, note=Cardholder data isolate intensive-margin changes; tracks same individuals; identification via within-individual variation.
  - page=14, words=275-590, section=Cardholder data identification, note=Within-individual variation over time identifies effects.

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.3177309306712598)`
- Retrieval method: `local`
- Evidence type: `Event-study/pre-trend test comparing NYC vs control cities; Figure 2 shows no pre-trend differences; post-change effects observed starting Apr 1, 2008.`

- Data source: Transaction and cardholder data from Starbucks in NYC, Boston, and Philadelphia (pre/post calorie posting).
- Table/Figure: Figure 2
- Assumption flag: `True`
- Assumption notes: Support for parallel trends assumption is evidenced by no pre-trend differences between NYC and control cities prior to calorie posting.
- Answer:

Yes. The study conducts event-study style pre-trend checks by comparing NYC stores to control cities (Boston/Philadelphia). It finds no evidence of pre-trend differences before calorie posting, and the post-policy drop in calories per transaction begins around April 1, 2008. Cardholder data show no pre-trend differences; any Christmas/New Year deviations appear only in transaction data.

- Quote snippet:

  First, with both datasets we see no evidence of pre-trend differences between NYC and Boston/Philadelphia.

- Citation anchors:
  - page=9, words=0-349, section=Results: Pre-trend checks and Figure 2, note=No evidence of pre-trend differences; drop begins around Apr 1, 2008; transaction vs cardholder differences noted.
  - page=14, words=0-349, section=Figure 2 and pre-trend discussion, note=Explicit statement of no pre-trend differences; post-period pattern follows posting date.

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.09999694849378597)`
- Retrieval method: `local`
- Evidence type: `text`
- Data source: Transaction data; Cardholder data
- Answer:

No instrumental variables are used. The study relies on a post-treatment indicator for NYC after April 1, 2008 and city controls with fixed effects and other controls; there is no IV and no first-stage regression reported.

- Quote snippet:

  The results shown in Table 3 are based on specifications in which calorie posting is binary—i.e., the POST variable is simply a dummy equal to one at NYC stores on every day after April 1, 2008.

- Citation anchors:

- page=13, words=275-528, section=Methods (Table 3 specification), note=POST variable is simply a dummy equal to one at NYC stores on every day after April 1, 2008.
- page=15, words=0-254, section=Figure 2 regression specification, note=Regression of log(calories per transaction) on separate week effects for NYC and control regions with day-of-week effects and weather controls.

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.24514799960328948)`
- Retrieval method: `local`
- Evidence type: `not_reported`
- Data source: Starbucks transaction data and cardholder data
- Answer:

The provided text does not report any first-stage statistics (F-stat, partial R^2) or explicit evidence of first-stage strength.

- Quote snippet:

  we estimate regressions of the following form: (1) ysct = xsct + POSTct + sct, where ysctis a measure of calories per transaction

- Citation anchors:
  - page=11, words=275-522, section=Methods/Empirical specification, note=Describes regression form with POST indicator; no instrument or first-stage statistics mentioned.
  - page=3, words=275-564, section=Limitations/Context, note=Discusses study limitations and data, with no mention of first-stage strength.
  - page=30, words=0-349, section=Discussion/Conclusion, note=Notes learning and salience; no first-stage statistics are discussed.

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.22339790055511907)`
- Retrieval method: `local`
- Evidence type: `design description; not an RDD`
- Answer:

The study does not implement an RDD. It uses a post-treatment indicator for NYC stores after April 1, 2008 within a regression framework with fixed effects, which is a difference-in-differences style design rather than an RDD. Therefore there is no running variable, cutoff, or bandwidth specification for an RDD.

- Quote snippet:

  POS Tc_t is a dummy equal to one if calories were posted (i.e., NYC stores after April 1, 2008)

- Citation anchors:
  - page=11, words=0-349, section=Methods/Identification, note=Definition of treatment: POS_t is a dummy equal to one if calories were posted (NYC stores after April 1, 2008) and the regression specification.

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

- Category: `B) Identification strategy / causal design`
- Confidence: `low (0.2889126741456382)`
- Retrieval method: `local`
- Evidence type: `textual description of the estimation strategy with post dummy and event-study variant`
- Data source: Transaction and cardholder data from Starbucks stores in NYC, Boston, and Philadelphia; Seattle data used for robustness
- Table/Figure: Figure 2 and Table 3 (DiD with post indicator and timing analysis)
- Assumption flag: `True`
- Assumption notes: Parallel trends between NYC and control cities prior to posting; pre-trend differences examined and found not evidence of divergence (Figure 2).
- Answer:

The paper uses a standard difference-in-differences setup with a post indicator for NYC after the calorie-posting date (April 1, 2008). They also implement an event-study style variation by using separate week dummies for NYC and the control cities to examine timing/dynamics of the effect. Additionally, they reference a staggered adoption across cities by noting Seattle's post-law date (January 1, 2009) as a robustness check.

- Quote snippet:

  y_sct = x_sct + _POST Tc_t + _sct, where y_sct is a measure of calories per transaction at store s in city c on day t, POST Tc_t is a dummy equal to one if calories were posted (i.e., NYC stores after April 1, 2008)

- Citation anchors:
  - page=11, words=275-522, section=Estimation, note=Defines the regression y_sct = x_sct + _POST Tc_t + _sct; POST is a dummy for NYC stores after April 1, 2008; includes week fixed effects and other controls.
  - page=14, words=0-349, section=Results/Timing, note=Figure 2 and

description of plotting the difference between NYC weekly fixed effects and control weekly effects (event-study style timing evidence).
  - page=16, words=275-558, section=Robustness / Additional Adoption, note=Notes Seattle's post-law date (Jan 1, 2009) as an additional adoption date and robustness check.

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

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.2803849335009675)`
- Retrieval method: `local`
- Evidence type: `dataset_description`
- Data source: Starbucks transaction data; Starbucks cardholder data
- Answer:

Two datasets: the Starbucks transaction data and the Starbucks cardholder data.

- Quote snippet:

  Our transaction data cover all 222 Starbucks locations in NYC, and all 94 Starbucks locations in Boston and Philadelphia.

- Citation anchors:
  - page=5, words=0-349, section=Data Summary, note=Describes the transaction data: 222 NYC locations and 94 Boston/Philadelphia locations.
  - page=5, words=275-591, section=Data Summary, note=Explicitly states dataset naming: 'first dataset as the transaction data and the second dataset as the cardholder data'.

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

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.32252574139462475)`
- Retrieval method: `local`
- Evidence type: `text`
- Data source: Starbucks transaction data and cardholder data
- Answer:

transaction

- Quote snippet:

  For each transaction we observe the time and date, store location, items purchased, and price of each item.

- Citation anchors:

  - page=5, words=0-349, section=Data Summary, note=Unit of observation is a transaction; data include time/date, store, items, price per transaction.

C03: What is the sample period and geographic coverage?

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.2306843139246988)`
- Retrieval method: `local`
- Evidence type: `text`
- Data source: Starbucks transaction data and cardholder data (NYC, Boston, Philadelphia).
- Answer:

Sample period: January 1, 2008 to February 28, 2009; geographic coverage: Starbucks locations in New York City, Boston, and Philadelphia.

- Quote snippet:

  Our transaction data cover all 222 Starbucks locations in NYC, and all 94 Starbucks locations in Boston and Philadelphia.

- Citation anchors:
  - page=5, words=0-349, section=Data Summary, note=Geographic coverage (NYC; Boston & Philadelphia) and time window around calorie posting.

C04: What are the sample restrictions / inclusion criteria?

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.24538555087080954)`
- Retrieval method: `local`
- Evidence type: `direct textual evidence from the study's data and methods sections`
- Data source: Starbucks transaction and cardholder data used in Bollinger et al., American Economic Journal: Economic Policy, 2011
- Answer:

Sample restrictions include: (1) Cardholder analysis uses a subsample of anonymous cardholders who averaged at least one transaction per week in one of NYC, Boston, or Philadelphia during the pre-period before calorie posting in NYC; this yields 7,520 NYC and 3,772 Boston/Philadelphia individuals in the cardholder sample. (2) Transaction data include all transactions from 222 NYC stores and 94 stores in Boston and Philadelphia, but exclude stores not open for the entire period and exclude any transactions with more than four units of a single item (balanced panel).

- Quote snippet:

We define a subsample containing any individual that averaged at least one transaction per week in one of NYC, or Boston and Philadelphia, in the period before calorie posting in NYC.

- Citation anchors:
  - page=5, words=0-349, section=Data and sample description (cardholder subsample), note=Definition of cardholder subsample: averaged at least one transaction per week in NYC, Boston, or Philadelphia during the pre-period before calorie posting in NYC.
  - page=6, words=275-517, section=Sample restrictions (transaction data), note=Exclude stores not open for the entire period (balanced panel) and exclude transactions with more than four units of any one item.

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

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.29791024856995396)`
- Retrieval method: `local`
- Evidence type: `Data summary from the Cardholder dataset describing sample size.`
- Data source: Starbucks transaction data (universe of transactions) and cardholder data (subsample with 1 transaction per week in NYC/Boston/Philadelphia).
- Answer:

1.51 million transactions (cardholder data).

- Quote snippet:

  There are 7,520 such individuals in NYC and 3,772 such individuals in Boston and Philadelphia, generating a combined 1.51 million transactions for us to study.

- Citation anchors:
  - page=5, words=0-349, section=Data Summary – Cardholder sample size, note=Reports 7,520 NYC cardholders and 3,772 in Boston/Philadelphia, generating a combined 1.51 million transactions.
  - page=5, words=275-591, section=Data Summary – Cardholder sample size, note=Contains the same statement about the cardholder subsample size and total transactions (1.51 million).

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

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.3267221472357983)`
- Retrieval method: `local`

- Evidence type: `textual`
- Data source: Transaction data; Cardholder data
- Answer:

The key outcome is calories per transaction, measured as calories per transaction at a store, city, and day (with breakdowns for beverage calories per transaction, food calories per transaction, and total calories per transaction). Transformations include log specifications: for example, log(beverage + food calories) and log(calories per transaction) are used in regressions, with coefficients interpreted as percent changes (e.g., a −0.119 coefficient implying an 11.2% decrease). The analysis also examines quantiles of calories per transaction and uses both level and log forms to decompose intensive vs. extensive-margin effects.

- Quote snippet:

  y_sct is a measure of calories per transaction at store s in city c on day t

- Citation anchors:
  - page=11, words=275-522, section=Methods/Empirical specification, note=Defines y_sct as calories per transaction and presents the regression form including the POST dummy.
  - page=12, words=275-624, section=Results/Specifications, note=Mentions log(beverage + food calories) and presents a −0.063 coefficient; discusses cardholder data.
  - page=12, words=275-624, section=Results/Interpretation, note=Reports a 11.2 percent decrease in food calories per transaction (coefficient −0.119) and a 5.0 percent total calories change using cardholder data.
  - page=16, words=275-558, section=Tables/Decomposition, note=Discusses margins and decomposition into number of beverages/food items per transaction.
  - page=22, words=0-349, section=Tables/Quantiles, note=Describes quantile-based analysis of calories per transaction (Table 6) and percent changes across quantiles.

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

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.2924329921257721)`
- Retrieval method: `local`
- Evidence type: `Describes treatment construction (POST dummy) and alternative specifications; lists constructed outcomes and fixed-effects approach.`
- Data source: Transaction dataset and cardholder dataset
- Assumption flag: `True`

- Assumption notes: Assumes identification comes from within-city variation over time (city-week level) with NYC post-implementation as the treatment; assumes no other concurrent shocks differentially affecting NYC stores during the post period.
- Answer:

Treatment/exposure is measured with a binary post variable (POST) that equals 1 for NYC stores on every day after April 1, 2008, indicating calorie posting. Identification relies on within-city variation over time (city-week level). A robustness alternative uses separate week dummies for NYC and control cities, excluding the POST indicator. Constructed variables include the treatment indicator itself and various calorie-related outcome metrics (e.g., log calories per transaction for beverages/food, total calories per transaction) as well as other derived measures (number of items per transaction, calories per item, etc.).

- Quote snippet:

  POS Tc is a dummy equal to one if calories were posted (i.e., NYC stores after April 1, 2008)

- Citation anchors:
  - page=11, words=275-522, section=Estimation specification, note=POS Tc is a dummy equal to one if calories were posted (i.e., NYC stores after April 1, 2008).
  - page=11, words=0-349, section=Identification strategy, note=Identification stems from within-city variation over time; policy variation is at the city-week level.
  - page=12, words=0-349, section=Methods, note=Store fixed effects for transaction data and individual fixed effects for cardholder data; identification from within-city variation over time.
  - page=12, words=275-624, section=Robustness / Alternative specification, note=An alternative approach uses separate week dummies for NYC and control cities, excluding the POST variable.

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

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.28724934374498645)`
- Retrieval method: `local`
- Evidence type: `text`
- Data source: Transaction data and cardholder data from the Starbucks calorie posting study
- Answer:

Yes. The key covariates/controls are week fixed effects, day-of-week dummies, holiday dummies, and weather controls (temperature and precipitation and their squares). These are included in the baseline regression specification for both

the transaction data and the cardholder data; in the cardholder analysis, there are also individual fixed effects. Thus, the controls that are always included across specifications are week fixed effects, day-of-week dummies, holiday dummies, and weather controls. Individual fixed effects are added in the cardholder dataset but not necessarily in the transaction dataset.

- Quote snippet:

  x_sct includes week fixed effects (to control for seasonality), day-of-week dummies, holiday dummies, temperature and temperature squared, and precipitation and precipitation squared.

- Citation anchors:
  - page=11, words=0-349, section=Methods/Regression controls, note=Baseline controls include week fixed effects, day-of-week dummies, holiday dummies, temperature and temperature squared, and precipitation and precipitation squared.
  - page=29, words=0-349, section=Cardholder data regression, note=Cardholder regressions include individual, week, and day-of-week fixed effects, and weather controls.

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

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.2782624872026636)`
- Retrieval method: `local`
- Evidence type: `data cleaning / exclusion criteria`
- Data source: Transaction data and cardholder data from Starbucks locations in NYC, Boston, and Philadelphia (2008–2009).
- Answer:

Missing data are handled by dropping incomplete observations: the analysis uses a balanced panel by excluding transactions that were not open for the full data period and excluding transactions with more than four units of any item. The paper notes that the excluded transactions represent 2.2 percent of all transactions; there is no indication of imputation or weighting.

- Quote snippet:

  The excluded transactions represent only 2.2 percent of all transactions.

- Citation anchors:
  - page=6, words=0-349, section=Data cleaning / exclusions, note=Excludes transactions not observed during the entire data period (balanced panel) and those with more than four units; 2.2% of transactions are excluded.

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

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.2957772185028331)`
- Retrieval method: `local`
- Evidence type: `textual`
- Assumption flag: `False`
- Answer:

No. The provided context does not indicate the use of survey weights or propensity weights. The analysis relies on regression specifications with fixed effects and other controls, but weights are not described.

- Quote snippet:

  to examine the effects of calorie posting while controlling for seasonality and other influences, we estimate regressions of the following form: (1) y_sct = x_sct + POST Tc t + _sct

- Citation anchors:
  - page=5, words=0-349, section=Data Summary, note=Data description; no weighting discussed.
  - page=11, words=0-349, section=Methods/Empirical strategy, note=Regression specification cited; includes fixed effects and weather controls; no weighting described.
  - page=11, words=275-522, section=Methods/Empirical strategy, note=Continuation of regression approach; no weighting described.
  - page=35, words=0-161, section=Appendix A, note=Survey appendix; no weighting described.

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

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.22255016300986244)`
- Retrieval method: `local`
- Evidence type: `text`
- Data source: Starbucks calorie posting study data (transaction data and anonymous cardholder data; NYC, Boston, Philadelphia; Seattle robustness check)
- Answer:

No—data are not merged across sources. The study uses two separate data sources: the transaction data and the anonymous cardholder data; the analysis then compares the separately estimated effects from each dataset rather than merging them into a single linked file.

- Quote snippet:

We refer to the first dataset as the transaction data and the second dataset as the cardholder data. In the analysis we compare the separately estimated effects of calorie posting on the cardholder data with transaction data.

- Citation anchors:
  - page=5, words=0-349, section=Data sources and linkage, note=Defines two datasets (transaction data and cardholder data) and states the analysis compares them rather than merging.
  - page=5, words=275-591, section=Data sources and linkage, note=Further description of cardholder vs. transaction data and their separate analyses compared in the study.

C12: What summary statistics are reported for main variables?

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.3270386890203643)`
- Retrieval method: `local`
- Evidence type: `Table 1`
- Data source: Table 1: Summary Statistics for Transaction Data and Cardholder Data (prior to calorie posting in NYC)
- Table/Figure: Table 1
- Answer:

The main variables' summary statistics are reported in Table 1 for two datasets (Transaction data and Cardholder data). The table lists: average weekly transactions per store; average weekly revenue per store; percent of transactions with brewed coffee; percent with beverage; percent with food; average number of items per transaction; average number of drink items per transaction; average number of food items per transaction; food attach rate; average dollars per transaction; average calories per transaction; average drink calories per transaction; and average food calories per transaction. The statistics are shown separately for Boston, New York City, and Philadelphia within each dataset, and are normalized to 1.00 to preserve confidentiality. All statistics are based on data prior to calorie posting in NYC (April 1, 2008).

- Quote snippet:

  Table 1—Summary Statistics for Transaction Data and Cardholder Data (Prior to policy change)

- Citation anchors:
  - page=6, words=0-349, section=Table 1, note=Summary statistics for transaction data and cardholder data; normalization to 1.00; data prior to calorie posting in NYC.

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

- Category: `C) Data, sample, and measurement`
- Confidence: `low (0.2833682878412908)`
- Retrieval method: `local`
- Evidence type: `descriptive_figures`
- Data source: Transaction data and cardholder data from Starbucks locations in NYC, Boston, and Philadelphia
- Table/Figure: Figure 1 and Figure 2
- Answer:

Yes. The paper includes descriptive figures that establish baseline patterns, notably Figure 1 showing average calories per transaction by city, and Figure 2 showing pre-trend checks (no pre-trend differences between NYC and control cities).

- Quote snippet:

  Figure 2 depicts the results for each dataset. There are no pre-trend differences between NYC and Boston/Philadelphia.

- Citation anchors:
  - page=page 9, words=0-349, section=Calories Per Transaction (Figure 1), note=Figure 1 shows average calories per transaction per week for NYC vs controls.
  - page=page 14, words=0-349, section=Baseline trend check (Figure 2), note=Figure 2 depicts results and shows no pre-trend differences between NYC and controls.

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

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `medium (0.4277382164232411)`
- Retrieval method: `local`
- Evidence type: `regression result reporting effect on calories per transaction (Table 3)`
- Data source: American Economic Journal: Economic Policy, calorie posting study using Starbucks data
- Table/Figure: Table 3
- Answer:

-5.8% decrease in average calories per transaction.

- Quote snippet:

  finding a 5.8 percent decrease in average calories per transaction, equivalent to 14.4 calories.

- Citation anchors:

- page=12, words=0-349, section=Table 3 results, note=Calorie posting reduces average calories per transaction by 5.8% (14.4 calories).

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

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `low (0.18139544853577588)`
- Retrieval method: `local`
- Evidence type: `Regression specification with fixed effects and weather controls used to identify the effect of calorie posting.`
- Data source: Transaction data and cardholder data for Starbucks (NYC calorie posting study)
- Answer:

The preferred specification is the regression in (1): y_sct = x_sct + _POST Tc t + _sct, where y_sct is calories per transaction, POST Tc t is a dummy for NYC stores after April 1, 2008, and x_sct includes week fixed effects, day-of-week dummies, holiday dummies, temperature and temperature squared, and precipitation and precipitation squared. This specification controls for seasonality and other influences (via fixed effects and weather controls) to isolate the effect of calorie posting, and is estimated separately for transaction data and cardholder data.

- Quote snippet:

  To examine the effects of calorie posting while controlling for seasonality and other influences, we estimate regressions of the following form: (1) y_sct = x_sct + _POST Tc t + _sct, where y_sct is a measure of calories per transaction...

- Citation anchors:
  - page=11, words=0-349, section=Regression specification, note=Describes the preferred specification and variables in the regression.
  - page=11, words=275-522, section=Regression specification, note=Continues with weather controls and sample usage.

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

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `medium (0.45859206251221274)`
- Retrieval method: `local`
- Evidence type: `Empirical findings from the Starbucks calorie-posting study (transaction and cardholder data), including robustness checks and heterogeneity analyses.`
- Data source: Starbucks transaction data (NYC) and anonymous cardholder data with fixed effects; control cities for seasonal variation.

- Answer:

The effects are economically meaningful but modest. Food calories per transaction fell about 14% on average after calorie posting (with roughly 75% of the reduction due to fewer food items purchased and about 25% from substitution to lower-calorie items). Total calories per transaction declined around 5–6% (about 5.8% in transaction data and 5.0% in cardholder data), translating to roughly 14.4 fewer calories per transaction. Beverage calories were largely unchanged. The impact was larger for high-calorie purchasers (e.g., a 26% reduction for those who tended to make high-calorie purchases). There was some evidence of a price-per-item increase, and profits were not significantly affected. In a back-of-the-envelope macro calculation, the authors suggest that even a 6% reduction in calories at all chain restaurants could have meaningful population-level implications.

- Quote snippet:

  calorie posting caused food calories per transaction to fall by 14 percent, on average

- Citation anchors:
  - page=24, words=275-564, section=IV. Discussion, note=Food calories per transaction fall by 14% on average; three-quarters of the reduction from not purchasing a food item (extensive margin), one-quarter from substitution (intensive margin).
  - page=12, words=0-349, section=Results (Table 3) and cardholder data, note=Transaction data show −14% reduction in food calories per transaction; beverage calories −0.3%; total calories −5.8% (14.4 calories). Cardholder data show beverage calories not significantly different; food calories −11.2%; total −5.0%.
  - page=22, words=0-349, section=Table 6, note=High-calorie purchasers exhibit larger reductions (e.g., 26% reduction for high-calorie purchases).
  - page=30, words=0-349, section=Is the Effect Big Enough to Matter?, note=Crude macro estimate: a 6% reduction in calories across all chain restaurants could have meaningful population-level effects.

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

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `low (0.2836801199350919)`
- Retrieval method: `local`
- Evidence type: `Robustness checks in regression analyses: date fixed effects, alternative dependent variable, serial-correlation handling, and cross-dataset validation`
- Data source: American Economic Journal: Economic Policy (Bollinger et al., calorie posting in chain restaurants), 2011

- Table/Figure: Table 3
- Answer:

Key robustness checks include (1) adding date fixed effects as a robustness check (replacing week/day controls); results barely change (beverage calories: −0.004, food calories: −0.152, total calories: −0.063); (2) re-running with absolute calories as the dependent variable; results are almost identical to the main specification; (3) addressing potential serial correlation by aggregating pre/post posting data and testing differences; results are broadly consistent with the main findings; and (4) cross-dataset validation using cardholder data vs transaction data to isolate intensive-margin effects; results show the pattern persists, with beverage calories not significantly different from zero in cardholder data, but food calories and total calories still show reductions, indicating robustness of the calorie-posting effect though magnitudes vary by dataset.

- Quote snippet:

  > As a robustness check, we include date fixed effects in the above specification (and therefore drop the day-of-week and week fixed effects). The estimates based on the transaction data are barely changed.

- Citation anchors:
  - page=12, words=0-40, section=robustness_date_fixed_effects, note=As a robustness check, we include date fixed effects; results barely change.
  - page=12, words=41-90, section=robustness_absolute_calories, note=We repeated the analysis using absolute calories as the dependent variable; results are almost identical.
  - page=12, words=91-140, section=robustness_serial_correlation, note=To address serial correlation, we aggregate pre/post calorie posting data and test the difference; findings are broadly consistent.
  - page=12, words=141-190, section=robustness_cross_dataset, note=Cardholder data vs transaction data to isolate intensive-margin effects; beverage calories not significantly different from zero, but food calories and total calories show reductions.

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

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `low (0.21220639848696363)`
- Retrieval method: `local`
- Evidence type: `placebo tests`
- Data source: cardholder data and transaction data
- Answer:

Two placebo tests are reported: (1) a Seattle transaction-data placebo showing no significant effect of calorie posting on calories per transaction, and (2) a

parallel-trends placebo check showing no pre-trend differences between NYC and the control cities (Boston/Philadelphia) before posting.

- Quote snippet:

  As noted in Section III, regression analysis of the transaction data for Seattle (and controls) shows no significant impact of calorie posting on either food or beverage calories per transaction.

- Citation anchors:
  - page=25, words=275-599, section=Seattle placebo test (Seattle transaction data), note=Seattle regression shows no significant impact of calorie posting.
  - page=14, words=0-349, section=Parallel-trends placebo test, note=No evidence of pre-trend differences between NYC and Boston/Philadelphia.

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

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `low (0.25705277603599697)`
- Retrieval method: `local`
- Evidence type: `falsification_outcome`
- Data source: Milk order data; cardholder and transaction data
- Answer:

Milk usage by type and the frequency of cardholder purchases were tested as falsification/unaffected outcomes.

- Quote snippet:

  In all cases, there was no statistically significant impact of calorie posting.

- Citation anchors:
  - page=page 20, words=0-349, section=B. Substitution Effects, note=Milk usage data tested; no significant impact found.

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

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `medium (0.37475415756821695)`
- Retrieval method: `local`
- Evidence type: `summary of heterogeneity findings from Tables 2 and 5 and related text`
- Data source: American Economic Journal: Economic Policy, February 2011
- Answer:

Heterogeneity results reported include: (i) by income/education: the decrease in calories per transaction is larger in zip codes with higher income and in areas with more residents with college degrees; (ii) by size: beverage-size responses show heterogeneity—some customers switch to smaller sizes with lower calories per ounce, while others switch to larger sizes with higher calories per ounce; (iii) by baseline exposure: high-calorie and high-frequency customers show different magnitudes of response (e.g., high-calorie customers see a larger drop, while high-frequency customers show a slightly smaller drop than low-frequency customers); (iv) by region: similar patterns are observed in control cities (Boston and Philadelphia), suggesting the effects are not driven by regional differences in NYC alone.

- Quote snippet:

  "the decrease in calories per transaction was larger in zips with higher income and in zips with more education (i.e., more people with college degrees)."

- Citation anchors:
  - page=20, words=0-349, section=Heterogeneity in the Effect of Mandatory Calorie Posting, note=Income and education heterogeneity (Table 5)
  - page=11, words=0-349, section=Table 2, note=Size-based heterogeneity in beverage choices (Smaller/Same/Larger size)
  - page=21, words=0-349, section=Heterogeneity in the Effect of Mandatory Calorie Posting, note=Baseline exposure heterogeneity: high-calorie vs other groups
  - page=12, words=0-349, section=Control cities / Regional variation, note=Region: patterns similar in control cities (Boston/Philadelphia)

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

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `low (0.1427236921374667)`
- Retrieval method: `local`
- Evidence type: `Empirical mechanism tests using commuter vs non-commuter comparisons, knowledge surveys, margin decomposition (extensive vs intensive), heterogeneity, and control-city robustness.`
- Data source: Starbucks transaction data and cardholder data; NYC posting; control cities (Boston, Philadelphia); Seattle data; commuter vs non-commuter comparisons.
- Table/Figure: Figure 3
- Answer:

Mechanism tests performed include: (1) a salience vs learning test by comparing NYC posting effects to nonposting markets using commuters (regular purchasers

who shop outside NYC) to see if effects persist outside posting areas; they find effects in nonposting stores, implying learning rather than salience; (2) in-store surveys conducted before and after posting to gauge knowledge of calories and how posting affected that knowledge; (3) margin decomposition to separate extensive vs intensive effects, examining substitutions to lower-calorie items vs changes in purchase frequency/quantity; (4) heterogeneity analyses across quantiles of calories per transaction and across commuting status to assess who is most affected; (5) use of control cities (Boston, Philadelphia, Seattle controls) to distinguish time/seasonality from posting effects. Collectively, these tests imply a learning mechanism (information exposure changes choices beyond posting venues) with substitutions toward lower-calorie items and reductions in quantity-item purchases, rather than a pure salience effect or frequency change; higher-impact effects are seen among high-calorie purchasers and across markets, with commuter data suggesting cross-market learning.

- Quote snippet:

  exposure to calorie information affects their choices even at nonposting (i.e., non-NYC) stores, which is consistent with a learning effect but inconsistent with the salience effect.

- Citation anchors:
  - page=3, words=275-564, section=Mechanism tests, note=Salience vs learning test: commuters' exposure effects in nonposting stores imply learning rather than salience.
  - page=7, words=275-566, section=Survey evidence, note=In-store surveys to measure knowledge of calories and impact of posting on knowledge.
  - page=16, words=0-349, section=B. Substitution Effects, note=Margin decomposition testing extensive vs intensive effects (substitution vs fewer items).
  - page=29, words=275-442, section=Commuter analysis, note=Commuters show calorie reductions in NYC and outside NYC; tests indicate cross-market learning patterns.

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

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `medium (0.3592342266694231)`
- Retrieval method: `local`
- Evidence type: `robustness checks across samples, bandwidths, and controls`
- Data source: Starbucks transaction data and cardholder data (NYC, Boston, Philadelphia)
- Table/Figure: Figure 2; Table 3; Table 9
- Answer:

The results are robust to alternative samples, bandwidths, and controls. Using

both transaction-level data and cardholder data yields consistent declines in calories per transaction after calorie posting. Robustness checks include comparing NYC against control cities, addressing potential serial correlation by aggregating pre/post data, and controlling for seasonality, holidays, and weather. Beverage calories show no significant change, while food calories per transaction fall by about 11.2% (total calories down about 5%), with similar patterns observed in transaction and cardholder analyses.

- Quote snippet:

  To address any concern over serial correlation, we aggregate all transaction data before calorie posting, and all transaction data after calorie posting, then test the difference between average calories per transaction before versus after.

- Citation anchors:
  - page=11, words=1-20, section=Methodology robustness, note=Estimates reported from both transaction data and cardholder data.
  - page=12, words=0-60, section=Serial correlation robustness, note=Aggregate pre/post posting data to test differences; results remain similar.
  - page=12, words=550-630, section=Cardholder data isolate intensive margin, note=Cardholder data isolates intensive-margin effects; beverage calories not significant; food calories −11.2%.
  - page=14, words=0-80, section=Figure 2 robustness, note=No pre-trend differences; drop around April 1, 2008; persists in both datasets.
  - page=29, words=0-100, section=Commuter analysis robustness, note=Table 9 shows commuters vs non-commuters; effects observed in NYC and outside NYC.

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

- Category: `D) Results, magnitudes, heterogeneity, robustness`
- Confidence: `low (0.2697093921016416)`
- Retrieval method: `local`
- Evidence type: `Conclusion bullets describing main findings and policy implications from the study.`
- Data source: Starbucks transaction data (city comparisons) and cardholder data; surveys.
- Answer:
- Mandatory calorie posting reduces average calories per transaction by about 6% at Starbucks, with the effect long-lasting.

- The reduction is mainly due to changes in food purchases; beverage calories show little change.
- The effect is larger among high-calorie consumers.
- Both learning and salience contribute to the observed behavior: exposure to calorie information increases sensitivity (salience) and prior NYC exposure can lead to reductions even at non-posting stores (learning).
- The overall impact on Starbucks profits is negligible on average; in some proximity to Dunkin Donuts, revenue may even rise.
- Direct impacts on obesity are likely small, but longer-run effects could be more meaningful if restaurants add more low-calorie options and public education benefits accrue; posting costs are low, which supports potential policy benefits.
- Quote snippet:

  We find that mandatory calorie posting causes average calories per transaction to fall by 6 percent at Starbucks.

- Citation anchors:
  - page=34, words=0-349, section=Conclusion, note=Main conclusion statements about 6% decrease, long-lasting effects, food vs beverage, learning and salience.
  - page=28, words=0-349, section=Conclusion/Discussion, note=Survey evidence suggests calorie posting reduces calories per transaction due to salience rather than learning.
  - page=30, words=0-349, section=Commuters, note=Learning and salience both play a role; effects observed in non-posting stores after NYC exposure.
  - page=31, words=0-349, section=Is the Effect Big Enough to Matter?, note=Back-of-the-envelope obesity impact discussion; small potential effect overall.
  - page=24, words=275-564, section=IV. Discussion, note=Profit impact negligible on average; higher profit margins via higher-priced items; implications for policy.

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

- Category: `E) Citations and related literature`
- Confidence: `low (0.27340975858791017)`
- Retrieval method: `local`
- Evidence type: `literature review`
- Answer:

The most important prior papers cited are: 1) Jacoby, Chestnut, and Silberman (1977) and the broader nutrition-information literature (Ippolito & Math-

ios 1990, 1995) showing that information about nutrition can influence consumer choices; 2) Bassett et al. (2008) and Downs, Loewenstein, and Wisdom (2009) and Wisdom, Downs, and Loewenstein (2010) providing empirical and strategic perspectives on calorie information at the point of purchase and how information design can promote healthier choices; 3) Currie et al. (2010) and related work (Anderson & Matsa 2011) connecting fast-food exposure and labeling to obesity outcomes, which frames what calorie-posting studies aim to influence; 4) DellaVigna (2009) and Chetty, Looney, and Kroft (2009) offering a theoretical framework on learning versus salience to interpret posting effects; 5) Variyam & Cawley (2006) and Jin & Leslie (2003, 2009) addressing signaling and the signaling value of labels in restaurant and grocery contexts. Collectively, these works establish that information provision at the point of purchase can affect choices, that the magnitude and interpretation depend on learning versus salience, and that obesity outcomes are an important but complex downstream question.

- Quote snippet:

  Two subsequent papers compare purchase data before and after calorie posting in NYC.

- Citation anchors:
  - page=25, words=275-599, section=Narrative, note=Mentions: 'Two subsequent papers compare purchase data before and after calorie posting in NYC' (Downs et al. 2009).
  - page=37, words=0-349, section=References, note=Contains Jacoby, Chestnut, and Silberman (1977) and Ippolito & Mathios (1990, 1995).
  - page=37, words=275-624, section=References, note=Cites Currie et al. (2010); DellaVigna (2009); Wisdom et al. (2010); Downs et al. (2009).
  - page=36, words=275-422, section=References, note=Includes Anderson & Matsa (2011); Currie et al. (2010); Chetty et al. (2009) relevant to salience/learning framing.
  - page=38, words=0-349, section=References, note=Cites Roberto et al. (2009); Variyam & Cawley (2006); Jin & Leslie (2003, 2009).

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

- Category: `E) Citations and related literature`
- Confidence: `low (0.23771165855886442)`
- Retrieval method: `local`
- Evidence type: `textual citation from related work section`
- Data source: NYC receipt data from fast-food patrons used in Downs, Loewenstein, and Wisdom (2009) and Elbel et al. (2009)
- Answer:

The work most directly builds on or extends the NYC calorie-posting receipt studies by Downs, Loewenstein, and Wisdom (2009) and Elbel et al. (2009).

- Quote snippet:

  Two subsequent papers compare purchase data before and after calorie posting in NYC.

- Citation anchors:
  - page=7, words=275-566, section=B. Related Research, note=Mentions two NYC calorie-posting receipt studies that are directly relevant prior work
  - page=8, words=0-349, section=Discussion of NYC calorie-posting and prior studies, note=Context describing prior NYC studies (Downs et al. 2009; Elbel et al. 2009) as the immediate predecessors

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

- Category: `E) Citations and related literature`
- Confidence: `low (0.26010566557050246)`
- Retrieval method: `local`
- Evidence type: `literature references / benchmark studies cited in the results`
- Answer:

The cited benchmarks/comparisons are the papers by Anderson & Matsa (2011) and Currie et al. (2010) as related work, and the NYC pre-post calorie-posting studies by Downs, Loewenstein, & Wisdom (2009) and Elbel et al. (2009).

- Quote snippet:

  Two recent papers by Michael L. Anderson and David A. Matsa (2011) and Janet Currie et al. (2010). Neither paper finds that fast food restaurants have a significant effect on obesity in general.

- Citation anchors:
  - page=8, words=Anderson-Matsa, section=Literature review / Background, note=Mentions benchmark papers Anderson & Matsa (2011) and Currie et al. (2010).
  - page=8, words=Currie-2010)., section=Literature review, note=Cites Currie et al. (2010) on obesity effects.
  - page=8, words=Downs-posting., section=NYC pre-post studies, note=Downs, Loewenstein, Wisdom (2009) pre-post calorie posting study used as comparison.
  - page=8, words=Elbel-NYC., section=NYC pre-post studies, note=Elbel et al. (2009) pre-post calorie posting study used as comparison.

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

- Category: `E) Citations and related literature`

- Confidence: `low (0.289029838272936)`
- Retrieval method: `local`
- Evidence type: `data sources`
- Data source: Starbucks transaction data; Starbucks cardholder data; milk order data; in-store customer surveys; Seattle/San Francisco survey data
- Table/Figure: Table 1
- Answer:

Cited datasets include: (1) Starbucks transaction data covering all 222 NYC stores and all 94 Boston/Philadelphia stores, with three months pre- and eleven months post-calorie posting; (2) Starbucks cardholder data (2.7 million anonymous individuals) used to analyze behavior changes and to separate intensive vs extensive margins; (3) Starbucks milk order data used to assess daily milk quantities replenished by location; (4) in-store customer surveys conducted in Seattle and San Francisco to test consumer knowledge of calories; and (5) related survey data and prior studies referenced for context and methodological comparison.

- Quote snippet:

  "Our transaction data cover all 222 Starbucks locations in NYC, and all 94 Starbucks locations in Boston and Philadelphia."

- Citation anchors:
  - page=5, words=0-349, section=Data Summary, note=Transaction data cover all Starbucks locations in NYC and Boston/Philadelphia, with pre- and post-period for calorie posting; includes item-level purchases and calories.
  - page=5, words=275-591, section=Data Summary, note=Cardholder data: 2.7 million anonymous individuals; used to isolate intensive vs extensive margins and compare with transaction data.
  - page=6, words=0-349, section=Table 1, note=Table 1 presents summary statistics for transaction data and cardholder data prior to policy change.
  - page=6, words=275-517, section=Table 1, note=Further notes on dataset comparability and confidentiality; supports description of data sources.
  - page=14, words=0-349, section=Results comparison, note=Figure 2 compares results for transaction data and cardholder data, showing NYC vs controls and pre/post patterns.
  - page=25, words=275-599, section=Seattle/San Francisco surveys, note=In-store customer surveys conducted in Seattle and San Francisco to supplement data with consumer responses.

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

- Category: `E) Citations and related literature`

- Confidence: `medium (0.3857005142588578)`
- Retrieval method: `local`
- Evidence type: `DiD/panel fixed-effects regression with within-city variation; references to randomized natural field experiments in the literature`
- Data source: Cited methodological references and the study's own transaction/cardholder data; bibliography includes works on field experiments and nutrition labeling
- Answer:

The study uses a difference-in-differences–style identification with fixed effects (city-week, store, week, etc.) and within-city/time variation to estimate the effect of calorie posting. It also cites methodological literature on field experiments, notably randomized natural field experiments.

- Quote snippet:

  > In both cases (transaction data and cardholder data), identification of the effect of calorie posting stems from within-city variation over time.

- Citation anchors:
  - page=11, words=275-522, section=II.A Calories Per Transaction, note=Presents the regression form and notes identification from within-city variation over time
  - page=12, words=0-349, section=II.A Calories Per Transaction, note=Discusses store fixed effects and individual fixed effects for identification
  - page=37, words=275-624, section=References, note=Cites Cai, Chen, and Fang (2009) Observational Learning: Evidence from a Randomized Natural Field Experiment

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

- Category: `E) Citations and related literature`
- Confidence: `low (0.2466639630086526)`
- Retrieval method: `local`
- Evidence type: `literature positioning against seminal works`
- Answer:

Yes. It positions itself against classic nutrition-information literature, notably Jacoby, Chestnut, and Silberman (1977) and Ippolito & Mathios (1990, 1995), treating them as seminal references whose claim that more nutrition information changes behavior is an old idea that the current study questions.

- Quote snippet:

The notion that increasing the provision of nutrition information may stimulate people to adopt healthier eating habits is an old idea

- Citation anchors:
  - page=7, words=275-566, section=B. Related Research, note=Mentions classic nutrition-information literature; cites Jacoby et al. (1977) and Ippolito & Mathios (1990, 1995).
  - page=8, words=0-349, section=B. Related Research, note=Notes earlier studies finding demand sensitive to nutrition information (Ippolito & Mathios; Kiesel & Villas-Boas; Variyam & Cawley).

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

- Category: `E) Citations and related literature`
- Confidence: `low (0.29701273332227734)`
- Retrieval method: `local`
- Evidence type: `Appendix reference (survey questionnaire)`
- Data source: Appendix (survey questionnaire)
- Answer:

Yes. The text cites an Appendix that contains the in-store survey questionnaire (two-page) used to measure calorie knowledge.

- Quote snippet:

  The questionnaire is shown in the Appendix.

- Citation anchors:
  - page=7, words=275-566, section=Methods / Surveys, note=The questionnaire is shown in the Appendix.

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

- Category: `E) Citations and related literature`
- Confidence: `low (0.2924731048267649)`
- Retrieval method: `local`
- Evidence type: `Limitations and gaps in existing literature as acknowledged by the authors in the study`
- Data source: Starbucks transaction data
- Answer:

Gaps in the literature highlighted by the authors include (1) no direct measurement of calorie-posting effects on obesity due to lack of BMI data, (2) limited generalizability because the study uses data from only one chain (Starbucks), and (3) uncertainty about whether changes in Starbucks purchases are offset by other eating decisions at home.

- Quote snippet:

First, we do not directly measure the effect of calorie posting on obesity itself. A second limitation is that we have data for only one chain (Starbucks).

- Citation anchors:
  - page=3, words=275-564, section=Limitations, note= authors note gaps: no obesity effect measurement; single-chain data; generalizability and potential home offset unobserved
  - page=8, words=0-349, section=Literature gaps / measurement, note= prior literature often relies on survey responses; limited use of actual sales data, signaling measurement gaps

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

- Category: `F) Replication and transparency`
- Confidence: `low (0.16797747380054723)`
- Retrieval method: `local`
- Evidence type: `not_present`
- Answer:

There is no indication of replication files or code being provided in the context; the excerpts reference an Appendix with a Stanford University survey but do not mention replication materials.

- Quote snippet:

  Appendix A 1 STANFORD UNIVERSITY RESEARCH SURVEY
  How many times per week do you typically come to Starbucks?

- Citation anchors:
  - page=35, words=0-161, section=Appendix A, note=Includes the Stanford University research survey; no replication files or code mentioned.

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

- Category: `F) Replication and transparency`
- Confidence: `low (0.2585968748290788)`
- Retrieval method: `local`
- Evidence type: `No pre-analysis plan or trial registration mentioned`
- Data source: Transaction data for Starbucks locations (NYC, Boston, Philadelphia) and cardholder data; plus in-store survey data.
- Answer:

No. The provided text does not mention a pre-analysis plan or registered trial. It describes data sources and a regression-based empirical strategy but does not reference preregistration or trial registration.

- Quote snippet:

  The basic impact of mandatory calorie posting on calorie consumption is evident without any regression analysis (no controls of any kind).

- Citation anchors:
  - page=11, words=350-420, section=II. Effect of Mandatory Calorie Posting on Calorie Consumption – Empirical Strategy, note=Describes regression specification with POS_Tc_t dummy and fixed effects.
  - page=14, words=0-349, section=Methods / Estimation, note=Figure 2 and regression-based analysis are presented.
  - page=9, words=0-349, section=Data & Approach, note=Outlines data sources (transaction and cardholder data) but no preregistration mention.

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

- Category: `F) Replication and transparency`
- Confidence: `low (0.275037519980488)`
- Retrieval method: `local`
- Evidence type: `Confidentiality and restricted data`
- Data source: Starbucks transaction data and cardholder data (confidential; restricted access)
- Answer:

Yes. The context discloses confidentiality constraints and restricted access for the data (confidential, cannot reveal further details).

- Quote snippet:

  Due to confidentiality requirements, we are unable to reveal any more details about these differences.

- Citation anchors:
  - page=5, words=0-349, section=A. Data Summary, note=To preserve confidentiality of competitively sensitive information, for both datasets
  - page=6, words=275-517, section=A. Data Summary, note=Due to confidentiality requirements, we are unable to reveal any more details about these differences.

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

- Category: `F) Replication and transparency`
- Confidence: `low (0.21844437303641645)`
- Retrieval method: `local`

- Evidence type: `procedural_description`
- Data source: Starbucks transaction data for NYC, Boston, Philadelphia and Starbucks cardholder data
- Table/Figure: Table 1—Summary Statistics for Transaction Data and Cardholder Data (Prior to policy change)
- Answer:

Yes. The paper documents key data cleaning and construction steps, including excluding stores not open for the entire data period to form a balanced panel, removing transactions with more than four units of any item, and normalizing values (NYC) to preserve confidentiality. It also notes the use of two datasets (transaction and cardholder) and provides a table summarizing data.

- Quote snippet:

  > we exclude transactions at stores that were not open during the entire data period (i.e., we analyze the balanced panel), and we exclude transactions that included more than four units of any one item

- Citation anchors:
  - page=6, words=0-40, section=Data cleaning, note=Excludes stores not open for the full period and excludes large-quantity transactions.
  - page=6, words=275-320, section=Data cleaning, note=Normalizes NYC values to preserve confidentiality.

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

- Category: `F) Replication and transparency`
- Confidence: `low (0.28419734314792333)`
- Retrieval method: `local`
- Evidence type: `explicit robustness check described (exclusion of weather controls).`
- Data source: Starbucks transaction data and cardholder data (NYC, Boston, Philadelphia).
- Answer:

Partially omitted.

- Quote snippet:

  > Our findings are unchanged if we exclude the weather controls.

- Citation anchors:
  - page=11, words=275-522, section=Robustness checks / Empirical strategy, note=Finding unchanged when weather controls are excluded; weather variables have an insignificant impact.

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

- Category: `G) External validity and generalization`
- Confidence: `low (0.28509778243723166)`
- Retrieval method: `local`
- Evidence type: `Textual evidence from the study highlighting commuter effects and limitations on generalizability to other chains.`
- Data source: Starbucks transaction and cardholder datasets (NYC, Seattle, Boston, Philadelphia).
- Answer:

Generalizable populations/settings are urban adults who are regular customers of large chain restaurants exposed to calorie labeling, especially commuters who split their visits between NYC and other locations; results may extend to other chain restaurants with posted calories, though generalization beyond Starbucks should be treated with caution due to the study focusing on a single chain.

- Quote snippet:

  > However, these cardholders may not be representative of Starbucks customers more generally, as we expect these individuals are above average in their loyalty to Starbucks.

- Citation anchors:
  - page=3, words=275-564, section=Main findings – Commuters, note=Demonstrates that exposure to calorie information affects commuters' choices across NYC and non-NYC stores, implying learning effects.
  - page=5, words=0-349, section=Limitations, note=Two important limitations are that there is data for only one chain (Starbucks), limiting generalization to other chains.
  - page=29, words=0-349, section=Commuters, note=Defines commuters and presents results for commuters' post-posting behavior in NYC vs non-NYC stores.

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

- Category: `G) External validity and generalization`
- Confidence: `low (0.21427365593611017)`
- Retrieval method: `local`
- Evidence type: `explicit limitations on generalizability discussed in the study`
- Data source: Bollinger et al., Calorie Posting in Chain Restaurants, American Economic Journal: Economic Policy (2011).
- Answer:

The populations/settings least likely to generalize are (1) other restaurant chains

beyond Starbucks and (2) the broader Starbucks customer base, since the cardholder sample is highly loyal and not representative.

- Quote snippet:

  There are two important limitations to this research.

- Citation anchors:
  - page=3, words=There-research., section=Limitations, note=Only one chain (Starbucks) studied; cannot generalize to other chains.
  - page=5, words=However,-Starbucks., section=Data limitations, note=Cardholder sample may not reflect the broader Starbucks customer base; loyalty bias.
  - page=14, words=Seattle-2009., section=Cross-city robustness, note=Policy context differs across cities (Seattle vs NYC); pastry exemption affects generalizability.

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

- Category: `G) External validity and generalization`
- Confidence: `low (0.20192755605586474)`
- Retrieval method: `local`
- Evidence type: `limitations and boundary conditions discussion`
- Data source: Starbucks cardholder data (transaction data)
- Answer:

Yes. They discuss boundary conditions and scope limits, notably that obesity effects are not measured and the data come from only one chain (Starbucks), limiting generalizability to other chains.

- Quote snippet:

  There are two important limitations to this research. First, we do not directly measure the effect of calorie posting on obesity itself. A second limitation is that we have data for only one chain (Starbucks).

- Citation anchors:
  - page=3, words=275-564, section=Limitations and scope, note=Two important limitations: obesity not measured and data limited to Starbucks; generalizability to other chains is uncertain.

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

- Category: `G) External validity and generalization`
- Confidence: `medium (0.35759790340300607)`
- Retrieval method: `local`
- Evidence type: `Empirical study findings with cross-city analysis and discussion of limitations`

- Data source: Starbucks transaction and cardholder data (NYC vs. control cities) around calorie posting starting April 1, 2008
- Table/Figure: Figure 2
- Assumption flag: `True`
- Assumption notes: Assumes that learning effects observed in NYC/Starbucks could transfer to other markets or chains; acknowledges that generalization is unproven and dependent on similar exposure, consumer base, and competitive context.
- Answer:

Results may differ in different time periods or markets due to seasonality, holidays, and changes in consumer composition, as well as differences in market structure and exposure. The study shows an immediate drop in calories per transaction after posting (April 1, 2008) that persisted through February 2009 in NYC data, with a temporary decline in effectiveness around Christmas/New Year in transaction data but not in cardholder data, suggesting time-varying dynamics. Generalizing beyond Starbucks to other chains or markets is uncertain because the analysis uses data from a single chain and there may be varying competitive environments, menus, and consumer bases. In commuter vs. non-commuter settings, effects vary by location and context, indicating potential learning transfer but not universal applicability.

- Quote snippet:

  > The drop in calories per transaction occurred right around April 1, 2008, and persisted through February 2009.

- Citation anchors:
  - page=3, words=275-564, section=Limitations and salience vs learning, note=Two important limitations: data from one chain; uncertain generalizability to other markets.
  - page=12, words=275-624, section=Temporal pattern of calorie posting effects, note=Drop in calories per transaction occurred around April 1, 2008 and persisted through February 2009.
  - page=14, words=0-349, section=Time variation and holidays, note=Transaction data show reduced effectiveness around Christmas/New Year; cardholder data do not show this pattern.
  - page=29, words=0-349, section=Commuters vs non-commuters in different markets, note=Table 9 analyzes commuters' vs non-commuters' responses; demonstrates outside NYC effects and differential impacts.

H01: Are key variables measured directly or via proxies?

- Category: `H) Measurement validity`
- Confidence: `low (0.3080115570445991)`
- Retrieval method: `local`

- Evidence type: `Direct measurement from transaction/cardholder data; derived nutrition metrics; survey-based self-reports`
- Data source: Transaction data and cardholder data
- Table/Figure: Table 1
- Answer:

Key variables are measured directly from the available data sources (transaction-level data and cardholder data). Calories per transaction and related calorie metrics are computed from observed purchases, and milk-order data provide direct measurements of milk choices; survey data contribute self-reported information on calories and perceptions.

- Quote snippet:

  > Based on the trans- action data, we compute that, prior to calorie posting, in NYC: average drink calories per transaction were 143; average total calories per transaction were 247.

- Citation anchors:
  - page=5, words=275-591, section=Data overview, note=Introduces cardholder vs. transaction data and measurement of calorie-related outcomes
  - page=6, words=0-349, section=Table 1, note=Table 1 presents summary statistics derived from transaction/cardholder data prior to posting
  - page=11, words=275-522, section=Empirical strategy, note=Describes regressions for calories per transaction using data sources
  - page=12, words=275-624, section=Empirical strategy, note=Notes that cardholder data isolate intensive-margin changes; data sources for calorie measures

H02: What measurement error risks are acknowledged or likely?

- Category: `H) Measurement validity`
- Confidence: `low (0.3323464745797265)`
- Retrieval method: `local`
- Evidence type: `Limitations and measurement error risks in data collection and analysis.`
- Data source: Transaction data, cardholder data, and in-store surveys from Seattle, San Francisco, NYC (Starbucks).
- Answer:

Key measurement-error risks acknowledged or likely include: (1) reliance on data from a single chain (Starbucks), limiting generalizability and potentially biasing calorie estimates if Starbucks differs from other chains; (2) incomplete measurement of calories per transaction due to missing data on milk additions, customization, and items like soy milk, which can alter calorie counts; (3) inability to capture changes in the extensive margin (non-transactions) when mea-

suring calories per transaction; (4) survey-based sampling biases in the in-store surveys (e.g., under-representation of regular coffee consumers, potential wave-specific sampling differences, and timing biases); (5) potential nonrepresentativeness of cardholders, which could bias learning vs. salience interpretations; (6) potential confounds such as time trends or January effects affecting survey responses, making attribution to posting uncertain.

- Quote snippet:

  There are two important limitations to this research.

- Citation anchors:
  - page=3, words=275-290, section=Limitations, note=Two important limitations are discussed, including lack of obesity measurement and data from only one chain.
  - page=25, words=275-320, section=Survey methods, note=In-store surveys with potential sampling bias; under-representation of regular coffee consumers; possible wave-specific sampling bias.
  - page=12, words=275-320, section=Data/Methods, note=Cardholder data isolates intensive-margin effects but does not capture extensive-margin changes; measurement limitations acknowledged.
  - page=27, words=0-60, section=Survey results, note=Survey results may reflect sampling biases; second wave potentially sampled more informed consumers.
  - page=6, words=275-300, section=Data limitations, note=Confidentiality limits detail about differences in data; incomplete disclosure relevant to measurement interpretation.

H03: Are there validation checks for key measures?

- Category: `H) Measurement validity`
- Confidence: `low (0.24878415846921556)`
- Retrieval method: `local`
- Evidence type: `Robustness checks / validation checks`
- Data source: transaction data and cardholder data
- Answer:

Yes. The paper reports several validation/robustness checks for key measures, including weather controls robustness, holiday/day-of-week fixed effects, and cross-dataset validation across transaction and cardholder data (and Seattle post-law data) to ensure results are not driven by particular assumptions or data sources.

- Quote snippet:

  we include day-of-week dummies and holiday dummies that differ for NYC and the controls, finding no difference in the estimates.

- Citation anchors:

  - page=11, words=275-522, section=Weather controls robustness, note=Weather variables insignificant; results unchanged if weather controls are excluded.
  - page=14, words=275-590, section=Fixed effects and holidays robustness, note=Day-of-week and holiday dummies included; estimates unchanged when holidays are dropped.
  - page=14, words=275-590, section=Cross-dataset validation, note=Similar calorie-posting effects observed in both transaction and cardholder data.
  - page=14, words=275-590, section=Seattle robustness check, note=Seattle post-law data used as an additional robustness check.

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

- Category: `H) Measurement validity`
- Confidence: `low (0.20302754159074737)`
- Retrieval method: `local`
- Evidence type: `explicit discussion of limitations but no construct validity discussion`
- Answer:

No. The authors do not discuss construct validity for core outcomes in the provided context.

- Quote snippet:

  > However, there are two important limitations to this research. First, we do not directly measure the effect of calorie posting on obesity itself.

- Citation anchors:
  - page=3, words=275-564, section=Limitations, note=Limitations section notes that obesity effects are not directly measured; no mention of construct validity discussion.

I01: What policy counterfactuals are considered or implied?

- Category: `I) Policy counterfactuals and welfare`
- Confidence: `low (0.3421445345053721)`
- Retrieval method: `local`
- Evidence type: `textual evidence`
- Data source: American Economic Journal: Economic Policy, 2011; Starbucks calorie posting study
- Table/Figure: Table 3; Figure 3
- Answer:

Policy counterfactuals considered or implied: (1) The within-city counterfactual using NYC stores with posting vs control cities (Boston and Philadelphia) where posting did not occur, to identify what would have happened absent posting. (2)

A nationwide policy counterfactual: if calorie posting were implemented across all chain restaurants, with a crude assumption of a 6% reduction in calories, what impact could that have on obesity. (3) Generalizability counterfactual: whether the observed Starbucks effects would hold for other chains, since the study uses data from a single chain and cannot confirm effects elsewhere. (4) Timing/menu-change counterfactual: whether posting would drive changes in menu offerings or simply shift consumer choices, noting pre-existing trends toward low-calorie options and potential lags in menu changes.

- Quote snippet:

  To control for other factors affecting transactions, we also observe every transaction at Starbucks company stores in Boston and Philadelphia, where there was no calorie posting.

- Citation anchors:
  - page=2, words=0-40, section=Dataset description / Counterfactual control cities, note=Control cities Boston and Philadelphia used as counterfactual for NYC.
  - page=3, words=275-564, section=Limitations, note=Only one chain; cannot know if effects generalize to other chains.
  - page=32, words=0-80, section=B. Is the Effect Big Enough to Matter?, note=Policy counterfactual: 6 percent reduction at all chain restaurants.

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

- Category: `I) Policy counterfactuals and welfare`
- Confidence: `low (0.31018180399339024)`
- Retrieval method: `local`
- Evidence type: `empirical findings from a field study of mandatory calorie posting in NYC Starbucks`
- Data source: American Economic Journal: Economic Policy (Feb 2011) – Calorie Posting in NYC Starbucks
- Assumption flag: `False`
- Answer:

Mandatory calorie posting reduced calories per transaction, mainly by discouraging food purchases (extensive margin) with little impact on visit frequency or beverage calories. About 14% fewer food calories per transaction overall, and roughly three-quarters of the total calorie reduction comes from not buying food items. Reductions are larger among higher-income, more-educated areas and among women; higher-calorie purchasers saw larger absolute reductions, and reductions tend to be larger at the upper end of the calorie distribution. There are some costs to policy implementation (menu-board changes, potential legal costs), and profits were largely unaffected overall (prices per item rose slightly, possibly shifting demand toward higher-margin items).

- Quote snippet:

  three quarters of the total calorie reduction can be attributed to people opting not to buy food items (i.e., the extensive margin of food demand).

- Citation anchors:
  - page=18, words=275-541, section=Table 4 bottom-panel findings on calories per item vs items per transaction, note=Notes that nearly three quarters of the total calorie reduction is due to not buying food items (extensive margin).
  - page=24, words=275-564, section=Discussion of overall calorie effects and profits, note=States that calorie posting reduced food calories per transaction by 14% and discusses profit implications and price per item changes.
  - page=20, words=0-349, section=Frequency of cardholder purchases, note=Reports no statistically significant change in transaction frequency; purchases frequency largely unaffected.
  - page=20, words=0-349, section=Heterogeneity by zip socioeconomic status, note=Finds larger calorie reductions in higher-income and more-educated zips.
  - page=20, words=275-551, section=Demographic heterogeneity in cardholder data, note=Female cardholders more responsive to calorie posting than males.
  - page=22, words=0-349, section=Distributional pattern of calorie reductions, note=Calorie reductions tend to be larger in the top half of the distribution; percent reductions are roughly similar from 75th to 99th percentile.
  - page=22, words=0-349, section=Cost considerations, note=Notes direct costs of changing menu boards and potential higher legal costs.

I03: Are cost-benefit or incidence analyses provided?

- Category: `I) Policy counterfactuals and welfare`
- Confidence: `low (0.31306873732199253)`
- Retrieval method: `local`
- Evidence type: `Descriptive discussion of costs and absence of formal cost-benefit/incidence analysis in the article.`
- Data source: Starbucks transaction data in NYC and control cities; in-store surveys; Seattle control data
- Assumption flag: `False`
- Answer:

No formal cost-benefit or incidence analyses are provided. The paper discusses costs (e.g., a cited $2,000 per store compliance cost for Wendy's and other potential indirect costs) but does not perform a formal cost-benefit or incidence analysis; it mainly reviews revenue/profit implications and discusses potential

public benefits.

- Quote snippet:

  One news report indicated the cost of compliance for the Wendy's chain was about $2,000 per store.

- Citation anchors:
  - page=4, words=275-576, section=Costs/Implementation discussion, note=Mentions costs of posting and cites Wendy's compliance cost (~$2,000 per store) and indirect costs.
  - page=34, words=275-351, section=IV. Discussion, note=Concludes with cost-benefit considerations and potential benefits; references costs and potential outweighing benefits.

I04: What policy recommendations are stated or implied?

- Category: `I) Policy counterfactuals and welfare`
- Confidence: `low (0.3366218290825945)`
- Retrieval method: `local`
- Evidence type: `Policy recommendation`
- Data source: Starbucks NYC transaction data (with controls in Boston and Philadelphia), Starbucks cardholder data, and in-store surveys.
- Assumption flag: `True`
- Assumption notes: Assumes that the benefits of calorie posting (calorie reductions, potential obesity impact, and public education) outweigh costs; assumes low costs of posting; assumes effects generalize beyond NYC and could have long-run effects.
- Answer:

Implement or maintain mandatory calorie posting on restaurant menus; benefits likely outweigh costs due to low posting costs, potential modest calorie reductions, public-education gains, and possible long-run menu innovation; expansion nationwide is plausible.

- Quote snippet:

  We find that mandatory calorie posting causes average calories per transaction to fall by 6 percent at Starbucks... the costs of calorie posting are very low—so even these small benefits could outweigh the costs.

- Citation anchors:
  - page=33, words=275-320, section=V. Conclusion, note=Policy recommendation that mandatory calorie posting is beneficial if benefits outweigh costs; low costs; potential for innovation and education; calls for expansion.
  - page=34, words=0-60, section=Conclusion, note=Empirical finding: 6% reduction in calories; costs are very low; long-run effects possible;

policy implications.

J01: What missingness or attrition patterns are reported?

- Category: `J) Data quality and integrity`
- Confidence: `low (0.3156909266007003)`
- Retrieval method: `local`
- Evidence type: `Data exclusions and sampling patterns (balanced panel, bulk-purchase exclusions, percentage exclusions, sample representativeness).`
- Data source: Starbucks transaction data and cardholder (anonymous) purchase data.
- Answer:

The study reports several missingness/attrition patterns via data exclusions and sample selection: (1) they exclude transactions from stores not open during the entire data period to maintain a balanced panel; (2) they exclude transactions that included more than four units of any one item (bulk purchases); (3) these exclusions amount to 2.2% of all transactions; (4) the fraction of excluded transactions is unknown but believed to be well under 5%; and (5) the cardholder sample may not be representative of Starbucks customers more generally, which could affect generalizability.

- Quote snippet:

  > We exclude transactions at stores that were not open during the entire data period (i.e., we analyze the balanced panel).

- Citation anchors:
  - page=5, words=0-349, section=Data Summary, note=Exclude transactions at stores not open during the entire data period to maintain a balanced panel.
  - page=5, words=0-349, section=Data Summary, note=Exclude transactions that included more than four units of any one item (bulk purchases).
  - page=5, words=0-349, section=Notes, note=The excluded transactions represent 2.2 percent of all transactions.
  - page=5, words=0-349, section=Notes, note=The fraction of excluded transactions is unknown, but we believe it to be well under 5 percent.
  - page=5, words=275-591, section=Data Summary, note=Cardholders may not be representative of Starbucks customers more generally.

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

- Category: `J) Data quality and integrity`
- Confidence: `low (0.31128811353338065)`
- Retrieval method: `local`
- Evidence type: `Data filtering / exclusions for outliers`

- Data source: Transaction data and Cardholder data (Starbucks dataset)
- Answer:

Outliers are handled by excluding certain transactions and stores rather than applying statistical winsorization. Specifically, the authors exclude transactions that include more than four units of any one item (to remove bulk purchases) and exclude stores that were not open for the entire data period (balanced panel). These exclusions account for about 2.2% of all transactions.

- Quote snippet:

  > We exclude transactions at stores that were not open during the entire data period (i.e., we analyze the balanced panel), and we exclude transactions that included more than four units of any one item because we consider these purchases to be driven by fundamentally different processes (bulk purchases for an office, say). The excluded transactions represent only 2.2 percent of all transactions.

- Citation anchors:
  - page=5, words=275-591, section=Data filtering and exclusions, note=Exclusion criteria: balanced panel stores and purchases with more than four units; 2.2% of transactions excluded.

J03: Are there data audits or validation steps described?

- Category: `J) Data quality and integrity`
- Confidence: `low (0.24336161630175762)`
- Retrieval method: `local`
- Evidence type: `Data validation and robustness checks`
- Data source: Transaction data; Cardholder data; Milk order data; Seattle data
- Table/Figure: Figure 1 and Figure 2; robustness checks referenced in text
- Answer:

Yes. The paper describes multiple data validation and robustness steps, including (1) creating a balanced panel and excluding incomplete or extreme transactions, (2) pre-trend checks showing no differential trends between NYC and controls, (3) robustness checks using day-of-week and holiday dummies, (4) analysis with two data sources (transaction data and cardholder data) for cross-validation, and (5) external checks such as Seattle's post-law data and control comparisons.

- Quote snippet:

  > as a robustness check, we include day-of-week dummies and holiday dummies that differ for NYC and the controls, finding no difference in the estimates.

- Citation anchors:

  - page=14, words=0-349, section=Robustness checks, note=Robustness checks with day-of-week/holiday dummies; pre-trend checks; within-individual variation.
  - page=6, words=0-349, section=Data cleaning, note=2.2% of transactions excluded; balanced panel.
  - page=5, words=275-591, section=Data sources / Balance, note=Cardholder vs. transaction data; balanced panel.
  - page=14, words=275-590, section=Placebo / external validation, note=Seattle placebo test with pastry exemption.
  - page=20, words=275-551, section=Heterogeneity analyses, note=Gender, frequency, and other demographics effects.
  - page=25, words=275-599, section=Survey / Additional data, note=In-store surveys; data collection methods.

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

- Category: `J) Data quality and integrity`
- Confidence: `low (0.3048953057287519)`
- Retrieval method: `local`
- Evidence type: `sampling bias / selective sample acknowledgment in methods`
- Data source: Starbucks transaction data (NYC, Boston, Philadelphia) and anonymous cardholder data; Seattle survey data
- Assumption flag: `False`
- Answer:

Yes. The authors acknowledge selective sample inclusion and potential reporting bias in their data sources, noting that the cardholder sample may not be representative of Starbucks customers overall (likely more loyal) and that the survey sample may under-represent certain consumer types (e.g., those ordering regular coffee).

- Quote snippet:

  > However, these cardholders may not be representative of Starbucks customers more generally, as we expect these individuals are above average in their loyalty to Starbucks.

- Citation anchors:
  - page=5, words=275-591, section=A. Data Summary, note=Cardholder data may not be representative of Starbucks customers more generally.

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

- Category: `K) Model fit and diagnostics`
- Confidence: `low (0.2471733819564274)`
- Retrieval method: `local`

- Evidence type: `Tables and figures notes (Table 4, Figure 2, Table 2)`
- Table/Figure: Table 4 (R^2 values) and Figure 2 (95% confidence intervals); Table 2 (p-value 0.11).
- Answer:

Reported goodness-of-fit and diagnostic metrics include R-squared values for the regression specifications (e.g., R^2's ranging from 0.27 to 0.82; 0.26–0.37 for cardholder data), standard errors in parentheses for coefficient estimates, and 95% confidence intervals (Figure 2 notes). A Pearson chi-square p-value (0.11) is also reported for a related test of equality of proportions.

- Quote snippet:

  R^2's ranging from 0.27 to 0.82.

- Citation anchors:
  - page=page 18, words=0-349, section=Table 4 notes, note=R^2 values reported for regression specifications (R^2's ranging from 0.27 to 0.82; cardholder data 0.26–0.37).
  - page=page 15, words=0-254, section=Figure 2 notes, note=95 percent confidence intervals (dashed lines) shown in regression figures.
  - page=page 11, words=0-349, section=Table 2 notes, note=Pearson's chi-square test with p-value 0.11.

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

- Category: `K) Model fit and diagnostics`
- Confidence: `low (0.2680523417435153)`
- Retrieval method: `local`
- Evidence type: `Regression specifications incorporating logs, levels, and nonlinear weather term`
- Table/Figure: Table 9; Table 2; Figure 3
- Answer:

Yes. The authors test both log and level functional forms: they regress with logged dependent variables (e.g., log daily sales; log(calories per item purchased); log(calories + 1)) and also use level outcomes (e.g., number of items per transaction). They also include nonlinear elements in controls (temperature and temperature squared) rather than only linear specifications, though they do not report testing other nonlinear functional forms for the main outcomes.

- Quote snippet:

  "we regress log daily sales on an indicator for calorie posting"

- Citation anchors:
  - page=page 16, words=0-349, section=Regression specifications (logs and levels), note=regress log daily sales

  - page=page 16, words=275-558, section=Regression specifications (logs and levels), note=dependent variables include number of items per transaction and log(calories per item purchased)
  - page=page 29, words=0-349, section=Table 9 description, note=The dependent variable is log(calories + 1)
  - page=page 11, words=0-349, section=Model specification / Weather controls, note=temperature and temperature squared (nonlinear term) included

K03: Are residual checks or specification tests reported?

- Category: `K) Model fit and diagnostics`
- Confidence: `low (0.23556498944397872)`
- Retrieval method: `local`
- Evidence type: `textual_evidence`
- Data source: Transaction data and Cardholder data (calorie posting study)
- Answer:

No. The text describes regression specifications and robustness checks (alternative specifications, controls, etc.) and reports robust standard errors, but it does not present residual checks or formal specification tests.

- Quote snippet:

  An alternative approach is to modify equation (1) to include separate week dummies for NYC and the control cities, and to exclude the POST variable.

- Citation anchors:
  - page=page 11, words=0-349, section=Variations in specification (alternative specification approach), note=Describes an alternative specification with separate week dummies and exclusion of POST

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

- Category: `K) Model fit and diagnostics`
- Confidence: `medium (0.37660521918656825)`
- Retrieval method: `local`
- Evidence type: `Robustness checks and alternative estimators`
- Data source: Transaction data and cardholder data
- Answer:

Results are robust to alternative specifications and estimators. Robustness checks show that including date fixed effects, dropping some controls, and using a conservative serial-correlation approach do not materially change the estimated effects; results are unchanged when weather controls are removed; using

absolute calories as the dependent variable yields almost identical results; and findings from transaction data are corroborated by cardholder-data analyses.

- Quote snippet:

  To address any concern over serial correlation, we aggregate all transaction data before calorie posting, and all transaction data after calorie posting, then test the difference between average calories per transaction before versus after.

- Citation anchors:
  - page=page 12, words=0-349, section=Robustness checks (date fixed effects; serial correlation; weather controls), note=Date fixed effects included; results barely changed; weather controls do not drive conclusions; serial-correlation robustness test.
  - page=page 12, words=0-349, section=Robustness checks (alternative dependent variable), note=Repeating the analysis using absolute calories as the dependent variable yields almost identical results.
  - page=page 12, words=275-624, section=Data-source robustness, note=Cardholder-data results corroborate transaction-data findings, isolating intensive-margin effects.

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

- Category: `P) Previous questions`
- Confidence: `n/a`
- Retrieval method: `n/a`
- Evidence type: `n/a`
- Answer:
- Objective and data
  - Analyzes the impact of NYC's mandatory calorie posting on chain restaurants using two datasets (store-day transaction data and cardholder data) in NYC vs control cities (Boston/Philadelphia). POST = 1 after April 1, 2008; results rely on NYC vs controls and fixed effects/controls. (page 13, 275-528)
- Overall impact on calories
  - Calorie posting reduces average calories per store-day by about 4.6 calories, with the reduction in calories per transaction smaller, consistent with an increase in total transactions after posting. (page 13, 275-528)
- Seattle placebo comparison (external validation)

  - In Seattle, beverage calories per transaction fell by 4.6 calories, while food calories per transaction rose by 0.8 calories, suggesting beverage-focused effects and little noise from the posting policy elsewhere. (page 15, 0-254)

- Item- and per-item effects
  - Table 4 (transaction data) shows: number of beverages per transaction increases modestly (0.005***); number of food items per transaction falls (−0.029***); calories per beverage and per food item also decline (log beverage calories per beverage: −0.008***; log food calories per food item: −0.039***). (page 18, 0-349)
  - Interpretation: the reduction in calories is driven both by fewer items purchased (extensive margin with food items) and by lower calories per item (intensive margin). Food items contribute notably to the effect. (page 18, 0-349)

- Main channel: extensive margin dominates
  - About three quarters of the total calorie reduction is attributed to people opting not to buy food items (extensive margin of food demand). (page 18, 0-349)

- Distributional and heterogeneity patterns
  - Quantile evidence: calorie reductions are larger for higher quantiles; the percent change is roughly stable around 5–6 percent from 75th to 99th percentile, though absolute decreases are larger at higher quantiles. (page 22, 0-349)
  - Heterogeneity: effects differ by characteristics (e.g., income, education, age, gender, customer type); e.g., higher-income and more-educated groups show additional differential effects (Posting × median income, Posting × percent with college degree, etc.). (page 21, 0-349)

- Summary takeaway
  - Mandatory calorie posting reduces calories per transaction primarily through reduced food purchases (extensive margin) and, to a lesser extent, lower calories per item; effects are larger for higher-calorie consumers and more pronounced at higher consumption quantiles. (page 18, 0-349; page 21-22, 0-349)

Sources used:

- NYC calorie posting effects and overall magnitude: page 13, words 275-528
- Seattle placebo example: page 15, words 0-254
- Table 4 item-level and margin findings: page 18, words 0-349
- Extensive-margin interpretation: page 18, words 0-349
- Distributional/heterogeneity results: pages 21-22, words 0-349