The Numerator Data

KELLOGG RESEARCH SUPPORT

Feb 24, 2025

Northwestern | Kellogg

- Vast database of offline and online purchases of a large panel of 4 million consumers, including a 'static panel' of 150k consumers
- Data start 01/01/2018, quarterly updates
- Not just classic retail: digital purchases, food delivery....
- Multiple data sources: Scanned receipts, emailed receipts, direct information from retailer through portal access
- "New Age Panel"
- Research Potential from linked consumer surveys

Parameter Estimates -4 -3 -2 Month to Adoption

Figure 6: Change in Grocery Spending Post GLP-1 Adoption

Notes: Changes in grocery spending using TWFE, based on 2,623 users and 4,072 corresponding matched non-adopters. The dependent variable is log of spending on groceries.

Source: Hristakeva et al. (2024), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5073929

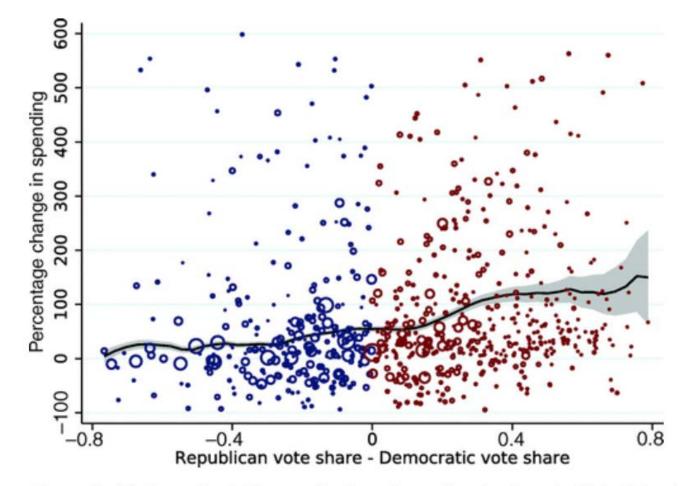


Figure 6. (Color online) Change in Goya Spending by County Vote Margin

Source: Liaukonyte et al. (2022), https://pubsonline.informs.org/doi/full/10.1287/mksc.2022.1386

WHERE ARE THE DATA FROM?



How It Works

Rewards

Blog

Support

Log In

Get rewarded every time you shop.

Upload your receipts and pad your piggy bank with extra cash.





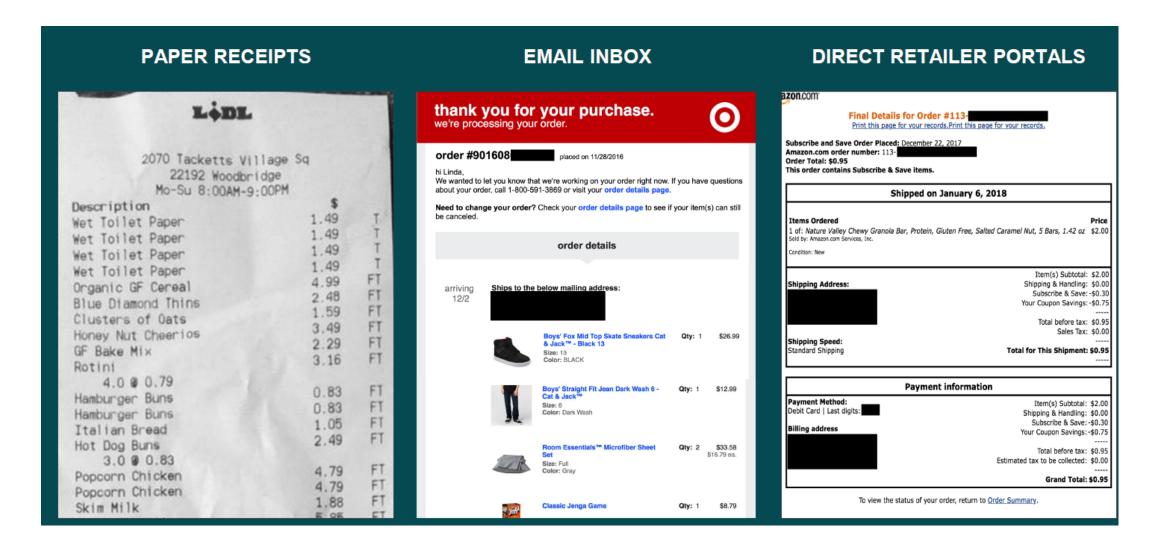








WHERE ARE THE DATA FROM?



WHERE ARE THE DATA FROM?

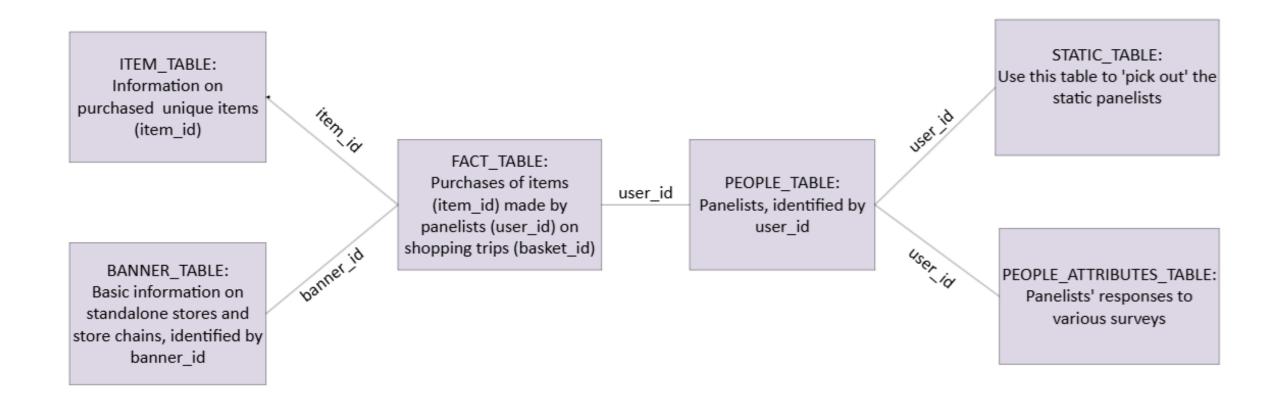
You're in for something fresh." 10200 W. GRAND FRANKLIN PARK, IL 50131 Phone # (847) 455-7711 Store Director - Pail Alsot		
12/26/13		20:05:39
GROCEIR'Y KLG SCBY DOO 122 KLG SCBY DOO 122 KLG SCBY DOO 122	3800057211 3300057211 3800057211	2.48 F1 2.48 F1 2.48 F1
You Save *ESEDY/JWLWPS18CT	.50 4130301356	1.99 T2
You Save *R*NLD FOIL 3201 You Save	1090000383	.37 T2
1 0 10, 10.00 *D/NNON OTKOS AF5 You Save	3563201865	1.00 F1
1 @ 10,10.00 *00K0S YOG 5.3 0Z	3563203218	1.00 F1
You Save CM LIO 640Z JWL 1% 128Z FROZEIN	5000032852 4128011435	4.69 F1 2.79 F1





- Weights available to make static panel representative and scale up to national market
- Records where order was placed (in store or online), and how it was delivered/collected (in store, shipping, download)
- Contrast to Nielsen data:
 - Numerator is available to all Kellogg researchers
 - Larger panel (150k vs 60k)
 - Uses data from scanned receipts instead of having panelists scan barcode of each item at home less churn
 - App contact point allows surveying of panelists
 - Frequent (quarterly) updates
 - Retailer names not encrypted
 - "Omnichannel coverage": better capture of online purchases, food delivery, etc.

DATABASE STRUCTURE



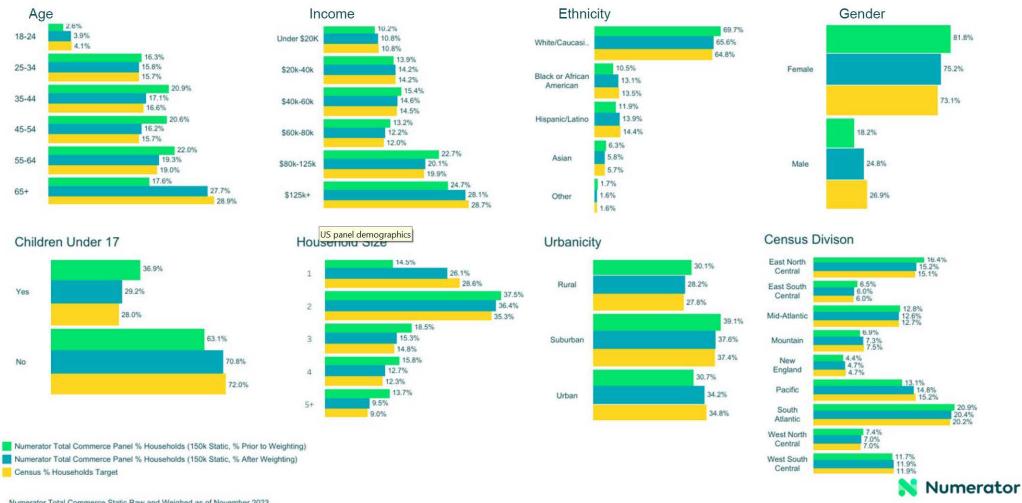
HANDLING QUARTERLY UPDATES, REPRODUCIBILITY

Superseding / Overlapping quarterly updates

- E g:
 - 4/18/2025 update: Q1 2025 data
 - 7/18/2025 update: Q1-Q2 2025 data
 - 10/17/2025 update: Q1-Q3 2025 data
 - 1/16/2026 update: CY 2025 data
- Manage by filtering on batch date variable
 - Best practices can vary by table!
- Use CTEs to keep code modular

WEIGHTING/SCALING

Total Commerce Static Demographics



Numerator Total Commerce Static Raw and Weighed as of November 2023

ACCESSING NUMERATOR

- Two main options:
 - Through Redivis:
 - "Workflows"/SQL console
 - Python ('redivis' package)
 - R ('redivis' package)
 - CLI interface to run SQL queries (kdcutils)
 - Through Athena/Kellogg Data Cloud:
 - SQL console
 - Python via pyodbc
 - R via odbc and DBI libraries
 - CLI interface to run SQL queries (kdcutils)