



United States Department of Agriculture

IRI Consumer Network, MedProfiler, RxPulse: household-based scanner data

Robert Wood Johnson Foundation/USDA, Economic Research Service
Grants Competition Informational Session

March 1, 2022

Patrick McLaughlin

The analysis, findings, and conclusions expressed in this presentation should not be attributed to IRI.



Proprietary Scanner Data

- Consumer purchase and sales data
 - Retail point-of-sale data
 - See presentation on IRI InfoScan
 - **Household panel data**
 - Weekly COVID-19 response data
 - See separate presentation
- Used by industry for marketing research
- Used by USDA for
 - Research projects, including external *collaborations*
 - Program evaluations, regulatory impact analyses
 - Data products, both public facing and internal value-added

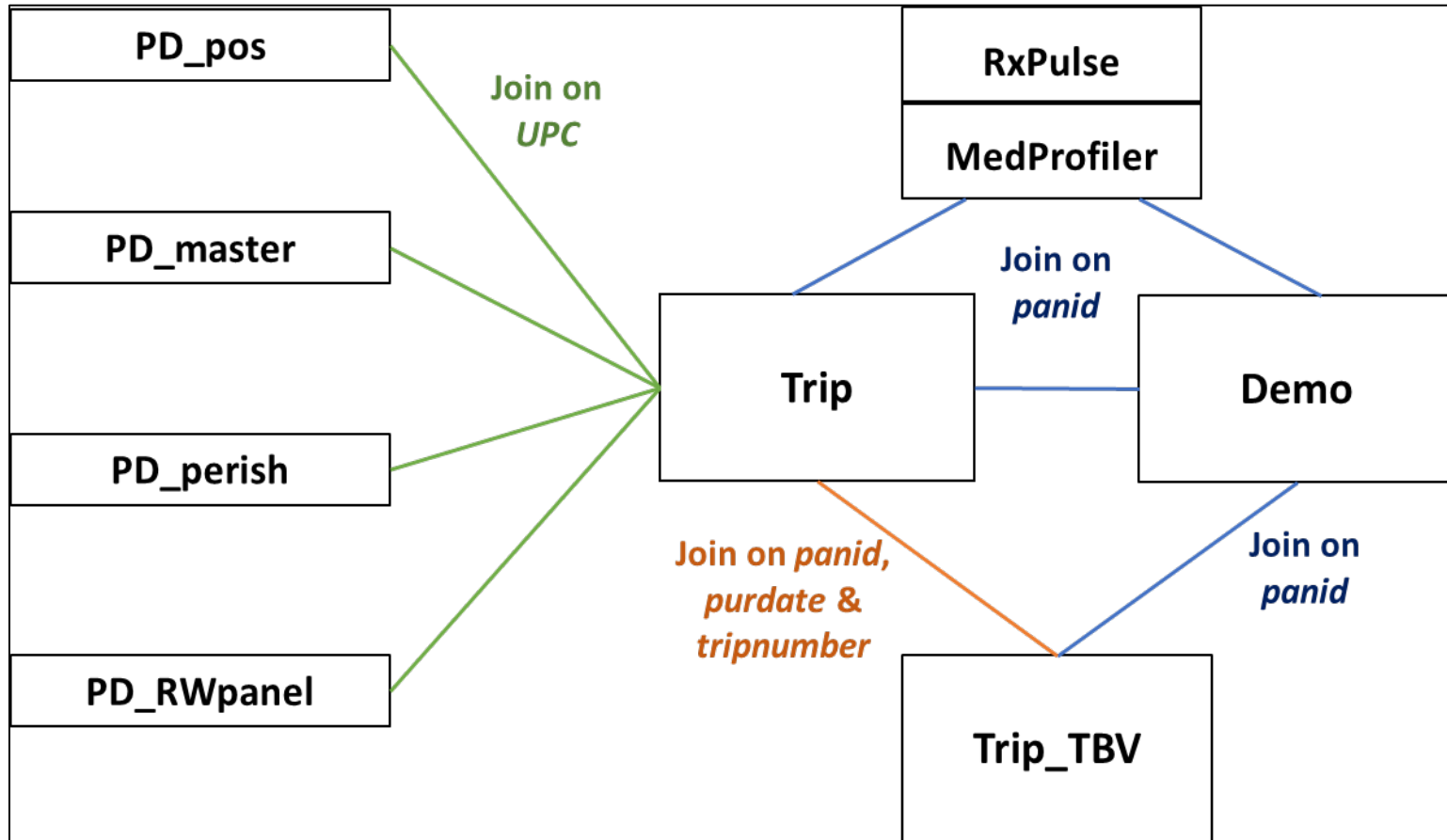


Household Scanner Data

- IRI Consumer Network household data (2008-2020)
 - National Consumer Panel
 - Joint venture between IRI and Nielsen
- What are households purchasing at food retailers?
 - Scan purchases at home
- Item-level food purchases for roughly 120K households
 - Around 60K households in more >1 year
 - Over 60 million purchase records per year
- Includes demographic characteristics of each household
- Subset linked with household health and prescription drug surveys
 - RxPulse and MedProfiler
- Detailed product information



Consumer Network Table Schema



Purchase information (“trip” table)

- Unique household ID number (“panid”)
- Store name
 - Chain name but don’t know which exact location
- Date of purchase
- Trip number (e.g., multiple trips to same store in one day)
- UPC of item purchased (join to product dictionaries)
 - UPC = universal product code
- Dollars paid
- Value of coupons used
 - Subtract from dollars paid for net expenditure
- Quantity of items purchased



Purchase information (“trip” table), cont. 1

- Purchased on promotion? (product feature, price reduction)
- Method of payment (cash, credit card, Debit, SNAP, WIC, etc.)
- Price source
 - Roughly 2/3 of all purchase prices sourced from IRI InfoScan
 - See separate presentation on IRI InfoScan data
 - HH enters prices for about 1/3 of purchases
 - Around 2% from “price dictionary”
 - Proprietary average prices for UPCs by outlet types at the national level
 - Used if no price in InfoScan and HH-entered price is unrealistically high or low
- A analogous table for total food + non-food purchases (“trip_tbv”)
 - Link to “trip” based on unique household ID (“panid”), purchase date, and trip number
 - Can use to compute food’s share of expenditures on food retail shopping trips



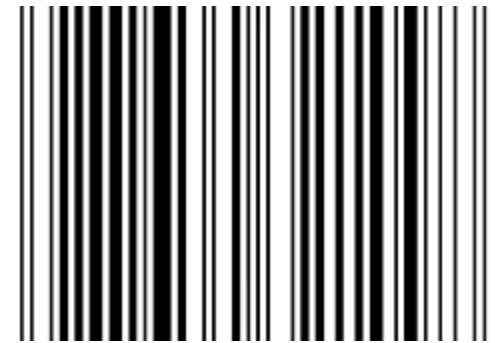
Purchase information (“trip” table), cont. 2

- Caveat on the types of food items and prices...
- Two broad types of food items
 - Point-of-sale (POS) or UPC food items (e.g., durable or shelf-stable packaged goods)
 - Perishables
 - Random weight (RW) - typically sold by weight (e.g., pounds of chicken breast) or by unit (e.g., number of lemons)
 - Fixed weight (FW) - packaged perishable items (e.g., 5 pound bag of apples, container of berries)
- Expenditures and quantity information for POS and FW perishables
- Expenditures only for RW perishables
- Detailed information in product dictionaries
 - Some variation by year



POS Product Dictionary, 2008 - 2018

- Product characteristics for over 1 million UPCs per year
- Key characteristics
 - Description
 - Brand
 - Size/weight
 - Product categorization
- Attributes, health, and nutrition-related claims
 - Type, flavor
 - Nutrition Facts Panel (limited)
 - Organic, gluten-free, type of sugar/artificial sweetener



Perishables Product Dictionary, 2008 - 2018

- Product characteristics for roughly 30-50K “above-UPC” food items per year
- Key characteristics
 - Form (e.g., cut of meat)
 - Preparation (e.g., cut fruit)
 - Package size (if FW)
 - Product categorization
 - FW or RW
- Attributes, health, and nutrition-related claims
 - Claims like organic or kosher
 - Color



POS + Perishables Dictionary, 2019-2020

- Product characteristics for over 1 million POS UPCs per year and 150,000 perishable products
- Same columns and information as POS product dictionary, 2008-2018
 - New: brand information and description for perishable products
- Distinguish FW vs RW products
- Some fields not populated for perishables products
 - Closer to perishables dictionary, 2008-2018



RW Panel Dictionary, 2008-2020

- 175 broad RW perishable products
- Only contains product classification information
 - Dept = Fruit and category = Banana
 - Dept = Chicken and category = Thighs



Household Demographics (“demo” table)

- Male and female household head
 - Birthdays
 - Indicator who is survey respondent (the household head)
 - Education levels
 - Occupation
- Household members (up to 7)
 - Birthday
 - Relationship to household head
- Household size
- Household income (categorical, ranges)
- Race/ethnicity
- Age and presence of children (categorical)



Household Demographics (“demo” table), cont. 1

- Geographic location
 - Census Region, State, county, zip code, Census block group
- Vehicle ownership
- Participation in SNAP or WIC (starting in 2017)
 - SNAP = Supplemental Nutrition Assistance Program
 - WIC = Special Supplemental Nutrition Program for Women, Infants, and Children
 - Self-reported, surveyed in June



Household Demographics (“demo” table), cont. 2

- **Projection factors**

- National Consumer Panel households recruited to be representative at the national and Census Region level.
 - Key demographics: HH size, HH income, age, race/ethnicity, occupation, education, presence of children
- Roughly 60K out of 120K record enough purchases to be assigned a projection factor or sample weight
 - “Static panel” – HHs with projection factors often in the panel for multi-years
 - In general, best practice is to restrict to analyses to the static panel
- Can be used at national and Census Region level
 - Additional projection factors for some markets (e.g., metropolitan areas)
- Also projection factors for HHs participating in RxPulse and MedProfiler



MedProfiler, 2008-2020

- Opt-in survey on health concerns, medical conditions, diet, and lifestyle
 - Roughly 1/3 of HHs have at least one member respond
 - Fielded in October
 - Separate projection factors
- Self-reported height and weight
- Over 100 questions about having, treating, or concern over health conditions and concern about healthy/dietary behaviors. E.g.,
 - Rate overall health on Likert-type scale
 - Does the respondent have hypertension? Treat using OTC or Rx drugs?
- COVID-19 questions in 2020
 - Rate concern over COVID-19, precautions to avoid infection
 - Did the respondent have COVID?
 - How often was hand sanitizer used

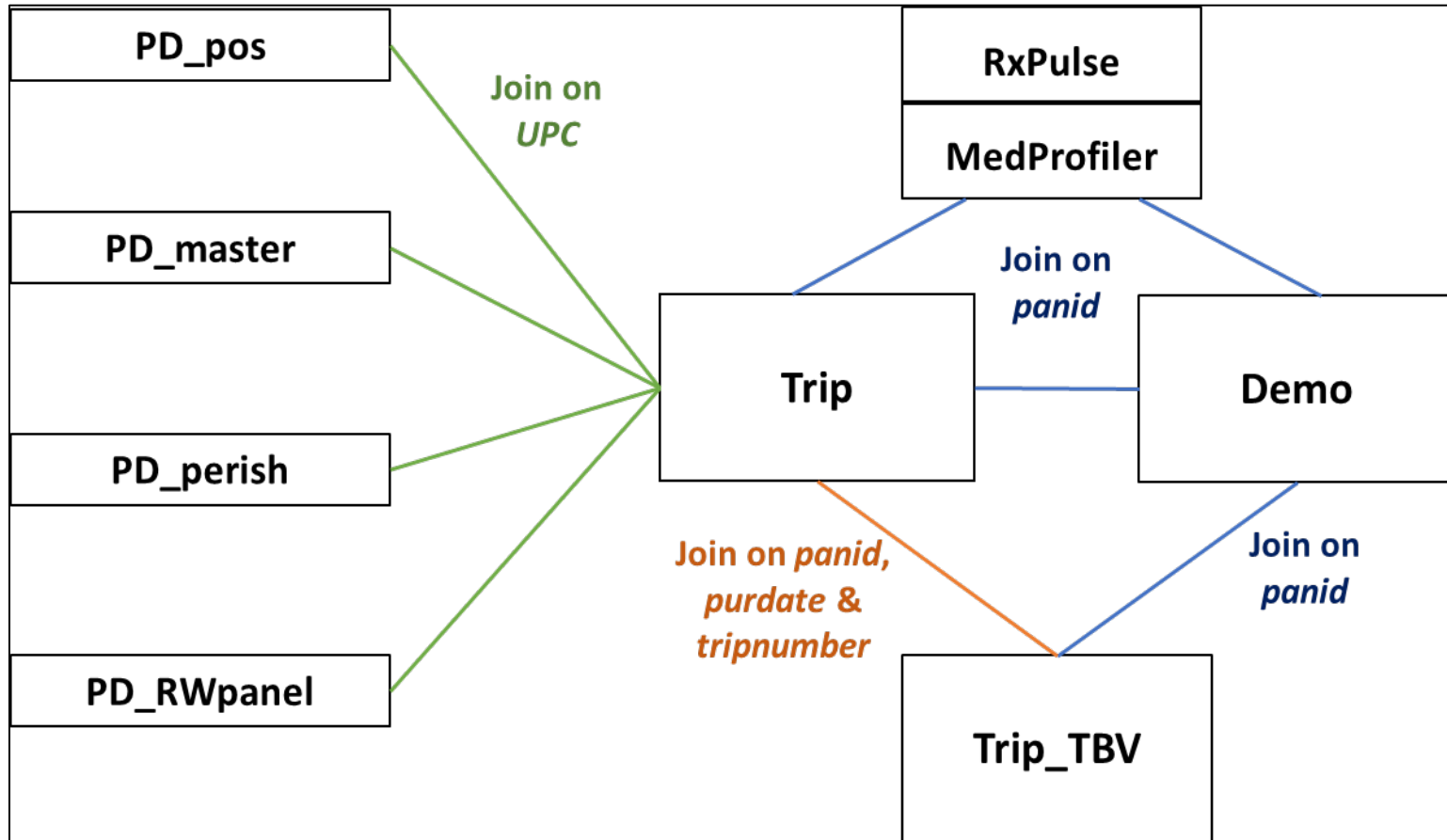


RxPulse, 2008-2020

- Opt-in, ongoing survey on prescription
 - Historically roughly 1/6 of households have at least one member respond
 - Separate projection factors
- Prescription purchase variables
 - HH identifier (“panid”) and birthday of respondent
 - Week of purchase
 - Prescription brand, condition treated, form of drug
 - Number of refills, insurance info, out of pocket \$\$
 - Location prescription was filled
 - Type of outlet
 - Reason outlet was chosen
 - Mail order?



Consumer Network Table Schema



For more information

- [Using Scanner Data](#) on the ERS website
 - Legal language for third-party access agreement (TPA)
 - Business offices **must** sign as is, no exceptions
- Technical bulletins
 - [Understanding IRI Household-Based and Store-Based Scanner Data](#) (TB-1942) by Mary Muth (RTI) and others
 - [Food-at-Home Expenditures: Comparing Commercial Household Scanner Data From IRI and Government Survey Data](#) (TB-1946)
 - [Examining Food Store Scanner Data: A Comparison of the IRI InfoScan Data with Other Data Sets, 2008–2012](#) (TB-1949)





United States Department of Agriculture

ERS Scanner Team

Patrick McLaughlin

Patrick.McLaughlin@usda.gov

Alex Stevens

Alexander.Stevens@usda.gov

Xiao Dong

Xiao.Dong@usda.gov

[Using Scanner Data](#)

