

Info Session on Nielsen Scanner Data

benfaber@econ.berkeley.edu

UC Berkeley, Jan 29 2014

Today's Info and Data Manuals

- Will be uploaded here:
 - http://emlab.berkeley.edu/eml/nielsen_data/

Accessing the Data

- Submit a request for a registration code:
 - <https://eml.berkeley.edu/cgi-bin/datareg.cgi>
 - For questions: rowilma@econ.berkeley.edu
- Use the registration code to register with the Kilts Center:
 - <https://kiltsnielsen.chicagobooth.edu/ApplyWebApp/ApplyAppAccessPage.aspx>
 - Type in the registration code in the “SUBSCRIBED” section.
 - Note that your advisor will have to be registered too.
- Data are not stored on a Berkeley server:
 - You will receive instructions on downloading and installing “Globus” which will allow you to download the data to your own computer.

Two Main Datasets

1. Consumer Panel Data (“Home scanner”)
2. Retail Scanner Data

General Overview

- Barcode level information on prices and quantities purchased:
 - ~3.2 million UPC codes (barcodes, “universal product code”).
 - 1075 product modules (e.g. PREGNANCY TEST KITS, CELLULAR PHONE).
 - 125 product groups (e.g. HEALTH AIDS, ELECTRONICS).
 - 10 departments (e.g. HEALTH & BEAUTY, GENERAL MERCHANDISE). [▶ View](#)
- “Product groups accounting for about 30% of US expenditure shares”.
- Period covered: 2004-2011.
- Frequencies: daily info (home scanner); weekly info (retail scanner).
- Finest geography: 5 digit zip code (households), counties (stores).
- Store identifiers and UPC codes can be linked between the two datasets.
- Store identities (store names) are anonymized. [▶ View](#)

Home Scanner Data

- Size Unzipped: ~20 GB
- ~40,000 households for 2004-07, ~60,000 for 2007 onwards.
- Demographic variables:
 - household size, age of children, type of residence, age, race, education, employment, occupation, marital status.
 - 20 income categories (e.g. \$35k-39.9k).
 - WIC assistance indicator.
- Purchase details:
 - Date, store id, price, number of units, coupon usage, sales information (part of multi-pack, etc.).
- Products:
 - Brand names and attributes (e.g. low fat, low salt, organic, shape, flavor, scent, strength).
- Retailers:
 - 66 different types of retail outlets (e.g. grocery vs hypermarket).

Retail Scanner Data

- 2006-2011
- Size Unzipped: ~2.5 TB
 - For downloading you will submit extracts of data by year and/or product segments.
- ~55,000 stores covered
- More than half of US sales in food and drugs, and 1/3 mass merchandise.
- Weekly frequencies of universe of sales and prices of any given store id by UPC codes.
 - Observing product entry, exit, promotions, etc.

Note on Data Format

- Tab-delimited format with double quote characters inside text fields
 - Importing should be done in Stata 13 (new import delimited command).
 - Earlier versions requires workaround (documentation provides some sample codes).

What Could these Data Be Linked to?

- UPC codes:
 - Nutritional information (e.g. Aviv Nevo, Emily Oster).
 - Domestic vs imported (e.g. John Romalis).
 - What about other firm characteristics? [View](#)
- Geography:
 - 5-digit zip codes and counties for households.
 - Counties and 3-digit zip codes for stores.

Some Recent Work in Different Fields

- Macro:
 - “Product Creation and Destruction: Evidence and Price Implications”, Broda, Leibtag, Weinstein, AER, 2010.
 - “Consumer Spending in the Great Recession”, Vavra, work in progress.
- Urban:
 - “Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across U.S. Cities”, Handbury, R&R Emtrca, 2013.
 - “Neighborhood Demographics and The Local Availability and Affordability of Groceries: Evidence from. Housing Project Demolitions”, Diamond, work in progress.
- Trade:
 - “China and Walmart, Champions of Equality”, Broda & Romalis, 2008.
- Public Finance:
 - “The impacts of alcohol taxes: Evidence from scanner data”, Griffith et al., 2013.
 - “The impact of banning advertising on the market for crisps”, Griffith et al., 2013.
- IO:
 - “The Evolution of Brand Preferences: Evidence from Consumer Migration”, Bronnenberg, Dubé, Gentzkow, AER, 2012.
 - Many more ...

Thanks for coming!

Tables

Which Kinds of Products?

Code/ Value	Description	Approximate Number of Product Groups
0	Health and Beauty Aids	21 (e.g. baby care, cosmetics, cough & cold remedies, deodorant, hair care, oral hygiene, pain remedies, skin care, shaving)
1	Dry Grocery	41 (e.g. baby food, baking mixes, bottled water, candy, carbonated beverages, cereal, coffee, condiments, crackers, pet food, prepared foods, snacks, soup, canned vegetables)
2	Frozen Foods	12 (e.g. ice cream, frozen pizza, frozen vegetables)
3	Dairy	12 (e.g. cheese, eggs, yogurt)
4	Deli	1
5	Packaged Meat	1
6	Fresh Produce	1
7	Non-Food Grocery	12 (e.g. detergent, diapers, fresheners/deodorizers, household cleaners, laundry supplies, pet care)
8	Alcohol	4 (e.g. beer, wine, liquor, coolers)
9	General Merchandise	19 (e.g. batteries/flashlights, candles, computer/electronic, cookware, film/cameras, insecticides, lawn & garden, motor vehicle, office supplies)
99	Magnet Data	1

▶ Back

