Using Scanner Data To Answer Food Policy Questions

Conference

Wednesday, June 1 - Thursday, June 2, 2011

Economic Research Service
1800 M Street, NW
Waugh Auditorium
Washington, DC

Using Scanner Data for the Public Good: Monitoring and measuring the US food and nutrition landscape

Shu Wen Ng

University of North Carolina at Chapel Hill

Using Scanner Data to Answer Food Policy Questions Conference • June 2, 2011

Our Food and Nutrition Landscape

Types of foods & beverages	Raw & Perishable	Packaged and processed	Prepared

‡ Poti & Popkin, JADA, in press
§ Ng, & Popkin, under review

Locations of purchase and consumption

Types of foods & beverages	Raw & Perishable	Packaged and processed	Prepared
Locations of purchase	Grocery, supermarket, tienda, farmer's market, cafeteria	Grocery, supermarket, tienda, convenience stores, vending machines	Quick service, Full service, grocery stores, cafeteria (school, work)
Locations of consumption	Home, cafeteria	Grocery stores, cafeteria , on-the-go, school, workplace, home	Quick service, Full service, grocery stores, cafeteria (school, work), on-the-go, home

Public monitoring of our Food Supply

Types of foods & beverages		Raw & Perishable	Packaged and processed	Prepared
Public data	Intake	NHANES Foods as reported	NHANES Foods as reported	NHANES 2007/08: 30% of calories prepared away from home 34% of calories eaten away from home ‡
	Food composition		USDA SR, FNDDS, MPED ~7,500 USDA food codes >60 nutrients	

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

‡ Poti & Popkin, JADA, in press
§ Ng, & Popkin, under review

UNCFRP
UNC Food Research Program
Robert Wood Johnson Foundation
University of North Carolina at Chapel Hill

Commercial measurement of our Food Supply

Types of foods & beverages		Raw & Perishable	Packaged and processed	Prepared
Public data	Intake	NHANES Foods as reported	NHANES Foods as reported	NHANES 2007/08: 30% of calories prepared away from home 34% of calories eaten away from home ‡
Public data	Food composition	USDA SR, FNDDS, MPED ~7,500 USDA food codes > 60 nutrients		
Commercial data	Sales/ purchase	Household panel until 2006 only	Retail scan & household panels 600,000 UPCs	Consumer panels
Commercial data	Nutrition	USDA SR	~ 200,000 unique formulations § Limited to NFP requirements	NDS-R: >1,800 Fast Food items/meals Limited to NFP requirements

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

UNCFRP
UNC Food Research Program
Robert Wood Johnson Foundation
University of North Carolina at Chapel Hill

Disconnect between measurements and with our Food Supply

Types of foods & beverages		Raw & Perishable	Packaged and processed	Prepared
Public data	Intake	NHANES Foods as reported	**NHANES Foods as reported**	NHANES 2007/08: 30% of calories prepared away from home 34% of calories eaten away from home ‡
	Food composition		**USDA SR, FNDDS, MPED ~7,500 USDA food codes > 60 nutrients**	
Commercial data	Sales/ purchase	Household panel until 2006 only	**Retail Scan & household panels 600,000 UPCs**	Consumer panels
	Nutrition	USDA SR	**~ 200,000 unique formulations § Limited to NFP requirements**	NDS-R: >1,800 Fast Food items/meals Limited to NFP requirements

‡ Poti & Popkin, JADA, in press
§ Ng, & Popkin, under review

Why turn to scanner data?

- Other limitations of publically available intake data
  - Sample design issues (e.g., North in summer, South in winter)
  - Does not measure usual intake (typically only 2-3 day recalls)
  - Significant lag time (e.g., 2 years for NHANES)
  - Possible rapid reformulations of the food supply
- Need to enhance our ability to monitor our food and nutrition environment, in order to create effective programs and policies across the areas of agriculture, economic development, welfare and public health.
  - Added information on prices with geographical and time variations
  - Possibly link with other geo-coded data on built environment
- Research groups working with scanner data
  - Marketing & business schools
  - ERS + grantees, CDC, FDA, ARS
  - Many agricultural economics departments in US, Canada and UK
  - UNC Food Research Program

THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL

UNCFRP
UNC Food Research Program
Robert Wood Johnson Foundation
University of North Carolina at Chapel Hill

Considerations

- Representativeness
  - of total US
  - of subpopulations (esp. vulnerable subpopulations)
  - of usual intake(rather than snapshots)
  - How about selection and attrition?
- Level of detail
  - Unit of observation (individual, household, store, market, national)
  - Geography (county, state, MSA, US total)
  - Time (weekly, 4-weekly, quarterly, annual)
- Accuracy
  - Response bias/ responder burden
  - Updatedness
  - Rounding
- Time to access & longevity
- Cost (will depend on above)

Commercial scanner data

- Point-of-sale store scanner data
    - Sales (volume, dollar) at the UPC level for F&B products with UPC for each week-year collected from food/grocery stores, drug stores, mass merchandisers (F/D/M) and convenience stores; includes promotion & prices; can also get calculated measures on %sales change due to any promotion, price decrease, feature and/or display
    - Nielsen Scantrack & SymphonyIRI Total Store Advantage

- Household panel scanner data
    - Panel data of household purchases at the UPC level for F&B products with UPC for each shopping occasion; includes promotion & price paid
    - Nielsen Homescan (>50,000 households per year)

- Packaged food nutrition data
    - Label information at the UPC level for F&B products including Nutrition Facts Panel data and ingredient list: Gladson Nutrition Database
    - Basic nutritional attributes at the UPC level for newly introduced F&B products: Datamonitor Product Launch Analytics and Mintel GNPD
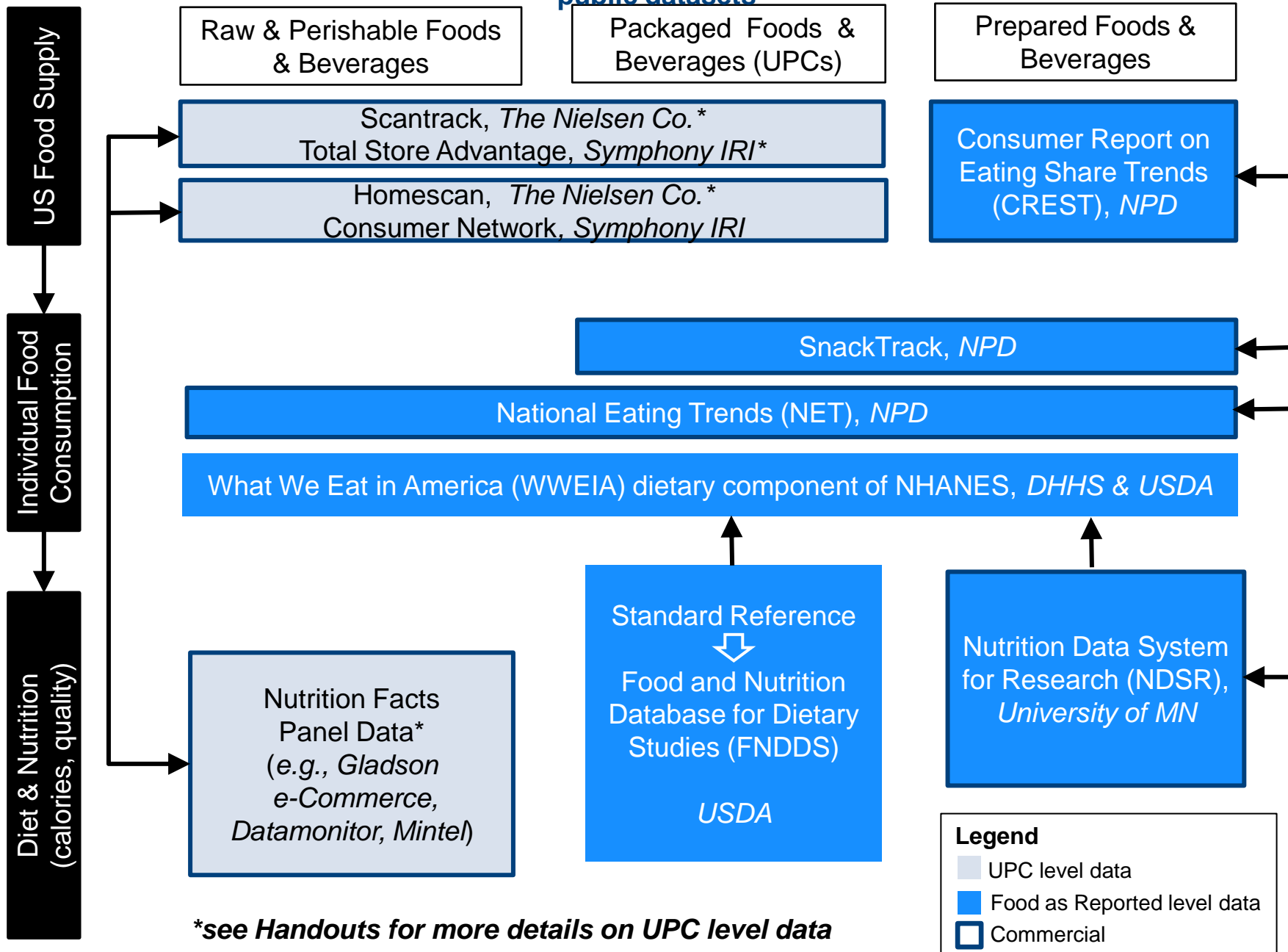
Major concerns with commercial scanner data *(see handouts)*

- Sales
  - Representativeness due to retailers and venues missed
  - Costs

- Purchase
  - Representativeness due to AFH foods
  - Bias
  - Costs

- Nutrition Facts Panel
  - Updated comprehensively to capture reformulations?
  - Limited to basic nutrient information; rounded values based on Nutrition Labeling Regulations; ±20% allowance
  - Problems will propagate to related UPCs
  - Limited prior data

Current measures of US food supply, consumption, diet and nutrition intake using commercial and public datasets

US Food Supply

Raw & Perishable Foods & Beverages	Packaged Foods & Beverages (UPCs)	Prepared Foods & Beverages

Scantrack, *The Nielsen Co.**
Total Store Advantage, *Symphony IRI**

Homescan, *The Nielsen Co.**
Consumer Network, *Symphony IRI*

Consumer Report on Eating Share Trends (CREST), *NPD*

Individual Food Consumption

SnackTrack, *NPD*

National Eating Trends (NET), *NPD*

What We Eat in America (WWEIA) dietary component of NHANES, *DHHS & USDA*

Diet & Nutrition (calories, quality)

Nutrition Facts Panel Data*
(*e.g., Gladson e-Commerce, Datamonitor, Mintel*)

Standard Reference
⇩
Food and Nutrition Database for Dietary Studies (FNDDS)

USDA

Nutrition Data System for Research (NDSR), *University of MN*

Legend
- UPC level data
- Food as Reported level data
- Commercial

see Handouts for more details on UPC level data

Using commercial data in conjunction with public data

Linking scanner sales/purchase with nutrition facts panel data

| Scantrack *sales* Homescan *purchases* | By UPC | Gladson + PLA nutrition |

~400,000 UPCs /year
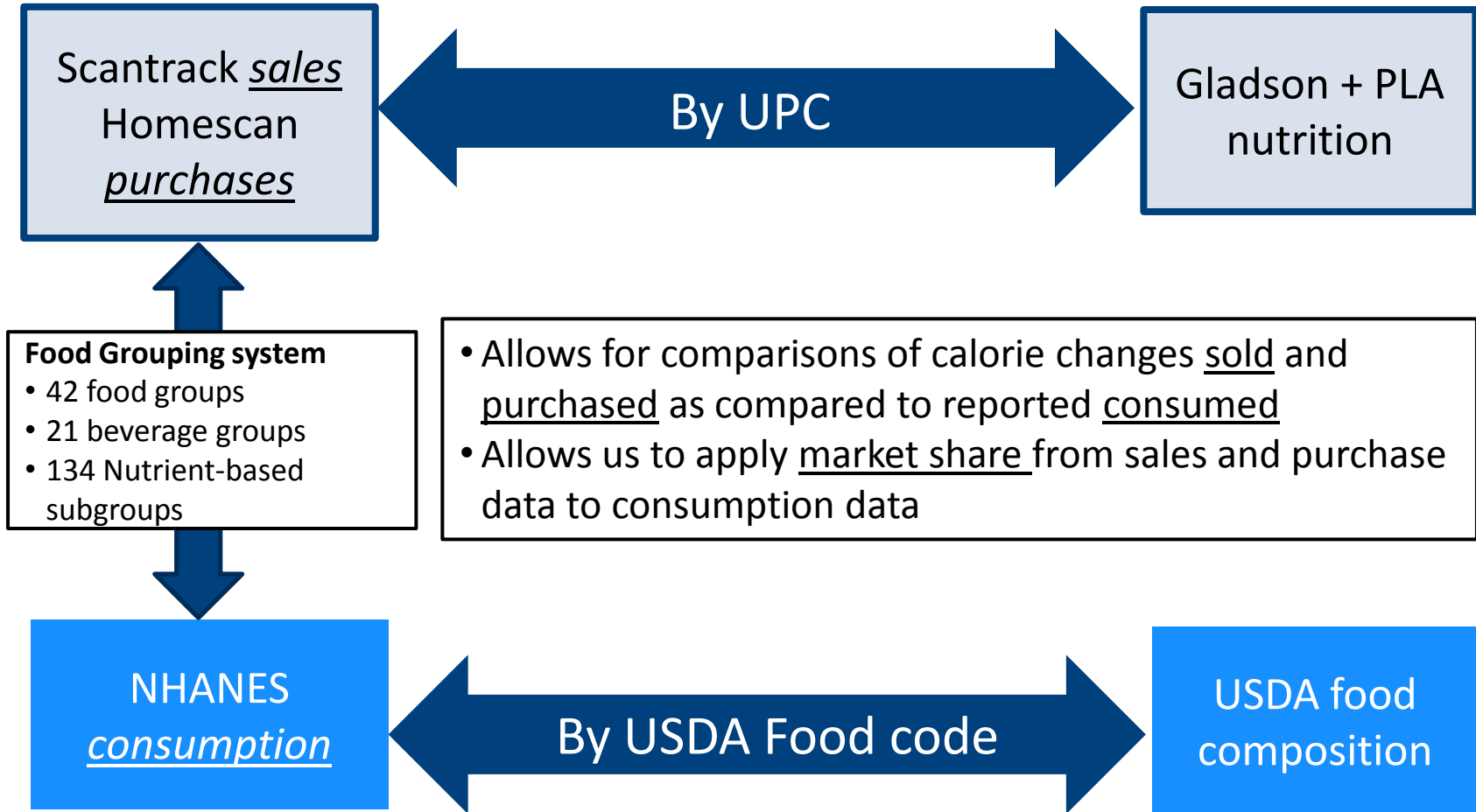
2007 Gladson: 140,000 UPCs
2010 Gladson: 170,000 UPCs
2009 PLA: added 2,500 UPCs

- Complex task
- 90% match in dollar and volume sales on average
- Ongoing with each additional year of sales, purchase & nutrition data
- After 2010, nutrient changes captured better

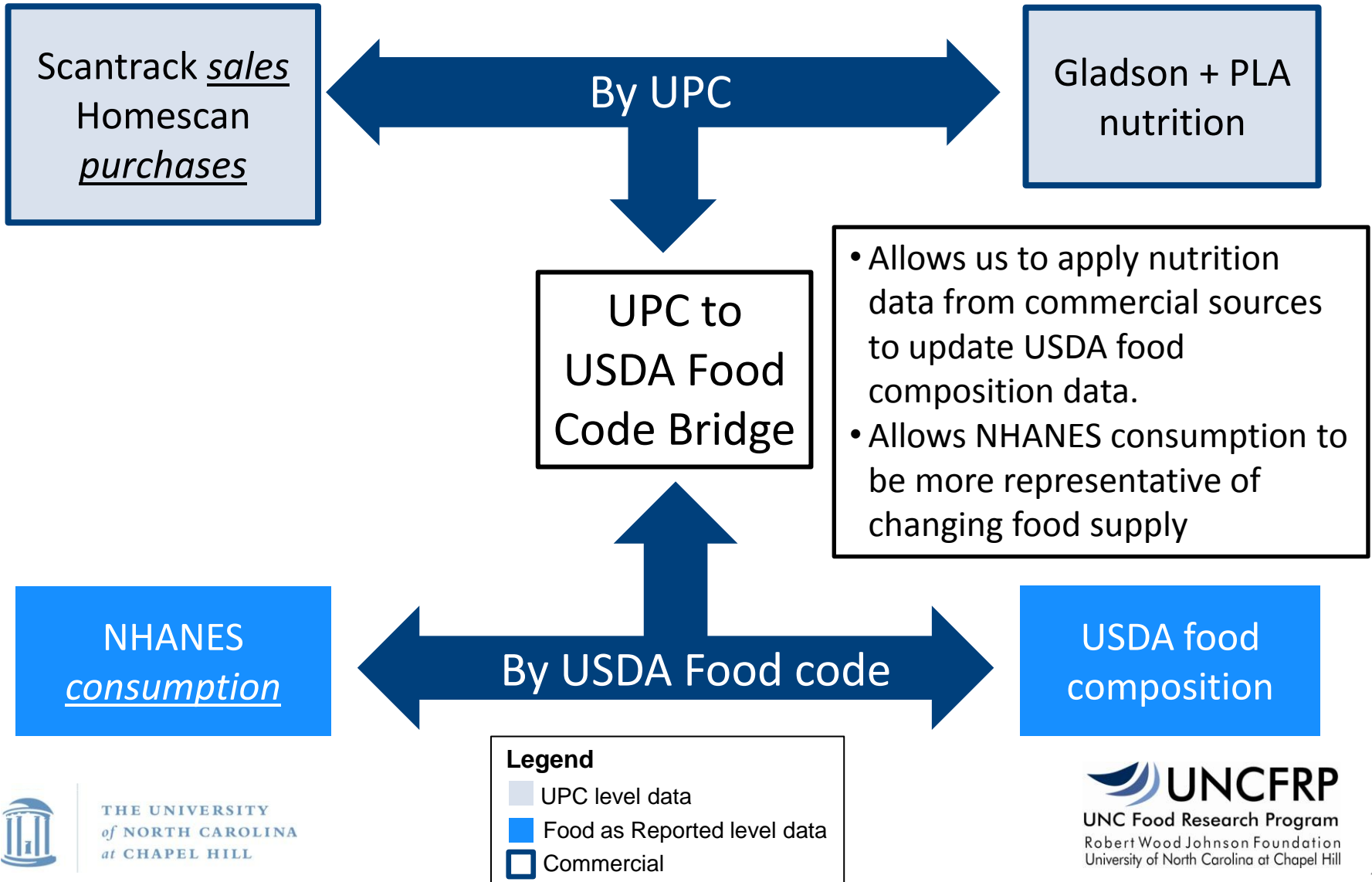# How to estimate proportion of reported calories from product groups

How to capture product reformulations, introductions in NHANES

Data and knowledge integration

- To maintain relevance, needs to be updated annually

- Significance and application

    – Cross-validations: Constructing multiple years of these datasets will allow for trend analyses to determine whether both kinds of datasets produce consistent findings.

    – Improve on existing FCTs to make them more reflective of the rapidly changing food supply using information from the commercial NFP label data sources

    – E.g., Gladson NFP label data includes the full ingredient list, whereas the FNDDS recipe file often does not. We used ingredients lists of 68,477 products from the 2007 and 2010 Gladson NFP label data to identify ready-to-eat cereals, bars and cookies reported in NHANES 2007/08 that contain fruit juice concentrate as an added sweetener.

How might the use of scanner data contribute to public health and welfare?

- Monitoring and measuring changes: "What gets measured, gets changed"
- Promote self-regulation by food manufacturers, retailers and food service
  - Nutritional content of products
  - Marketing or pricing strategies
- Encourage marketing companies to collect information useful for public health research
  - Typical clientele: food companies, retailers and service; marketing & business schools
  - Public Health is a new market and Nielsen (and perhaps others) are definitely interested
- Measure home, workplace, school food environment
- General public starting to use these (e.g., via iPhone apps)

THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL

UNCFRP
UNC Food Research Program
Robert Wood Johnson Foundation
University of North Carolina at Chapel Hill

17

Limitations still exist

- If using only UPC data, will miss:
  - Away-from-home foods (restaurants, cafeteria, concessions, vending)
  - Foods without UPCs (e.g., random weight items, farmer market purchases)
- Important considerations that are unknown and out of direct control of researchers:
  - Sampling frame of commercial data
  - Comprehensiveness of how updated commercial NFP data is
- Costs can be significant
  - Federal agencies should explore ways to share access to data

Thanks!

Contact: Shu Wen Ng

shuwen@unc.edu