

# The Surveys of the CPI: How They Fit Together and General Sampling Features

---

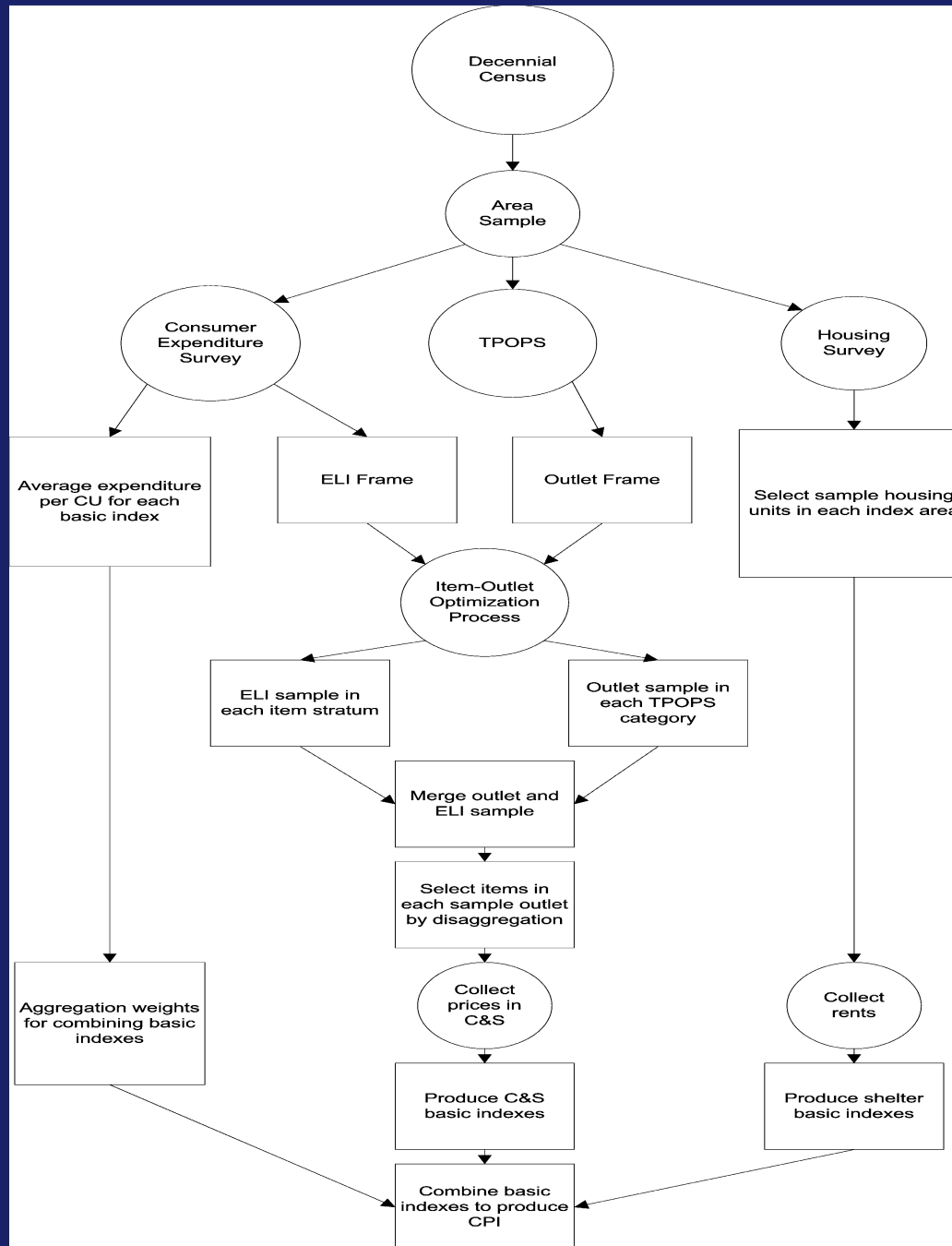
William Johnson  
Supervisory Mathematical Statistician  
DPSM  
October 28, 2014



# Surveys feeding into the CPI

---

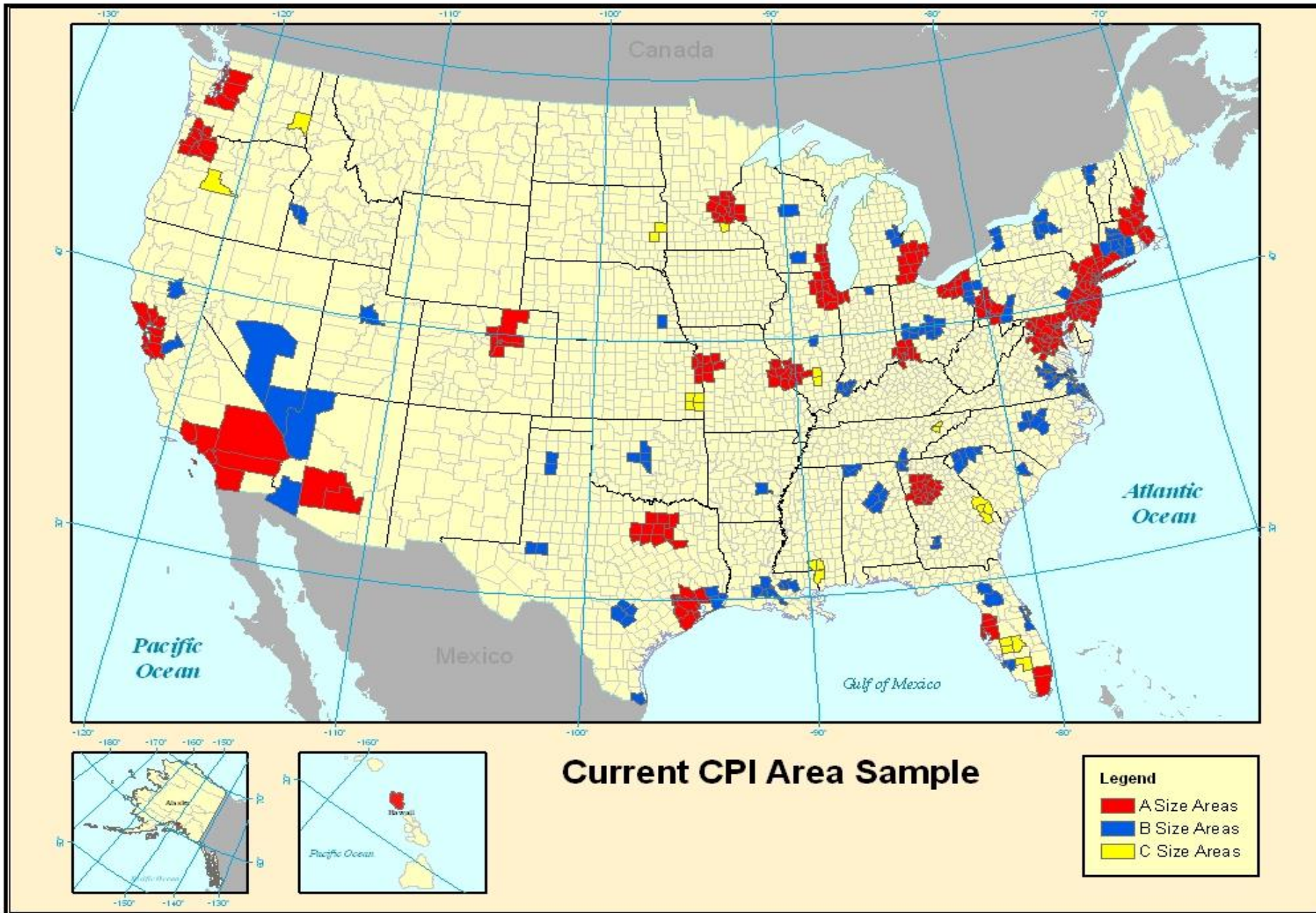
- Geographic sample of CBSAs
- Consumer Expenditure Survey (CE)
- Telephone Point of Purchase Survey (TPOPS)
- Housing Survey
- Commodities and Services Survey (C&S)



# Area Sample

---

- All surveys feeding the CPI are conducted in selected areas
- Normally done every 10 years based on data from the most recent decennial Census
- Sample units are CBSAs (formerly MSAs)
- Employs overlap maximization and controlled selection
- Provides represented populations for use in weights



# Consumer Expenditure

---

- The CE consists of two separate surveys –a quarterly interview and a diary:
  - ▶ the quarterly interview surveys (over 5 consecutive quarters) are used to ask consumers about their major purchases.
  - ▶ Additionally, consumers keep a diary of their purchases over a single two week period.

# Consumer Expenditure

---

- These surveys are used to create the expenditure weights that are used in constructing the CPI market basket.
- The expenditure weights are updated every two years.
- The current weights used in the CPI are based on consumer expenditures in 2011-2012.

# Consumer Expenditure

---

- CE is used to calculate the relative importance of Entry Level Items (ELI) within Item Stratum by region
- CE is also used to calculate percent of POPS, one of the factors in the C&S final quote weight
- CE microdata is used in constructing models of Owners' Equivalent Rent used in selecting the housing sample



# Telephone Point of Purchase Survey

---

- Random digit dialing survey conducted by Census
- Separate frames for landlines and cell phones
- Each area is asked about 1/8 of item categories every six months
- All items are typically rotated in an area every 4 years

# Telephone Point of Purchase Survey

---

- Respondents are asked how much they spent on various items during the recall periods for the items
- We ask for the expenditure broken down by outlet where the good or service was purchased
- We ask for the name and street of outlets
- Provides an estimate of the total daily expenditure for the POPS category in the PSU (called the *basic weight*)

# Commodities and Services

---

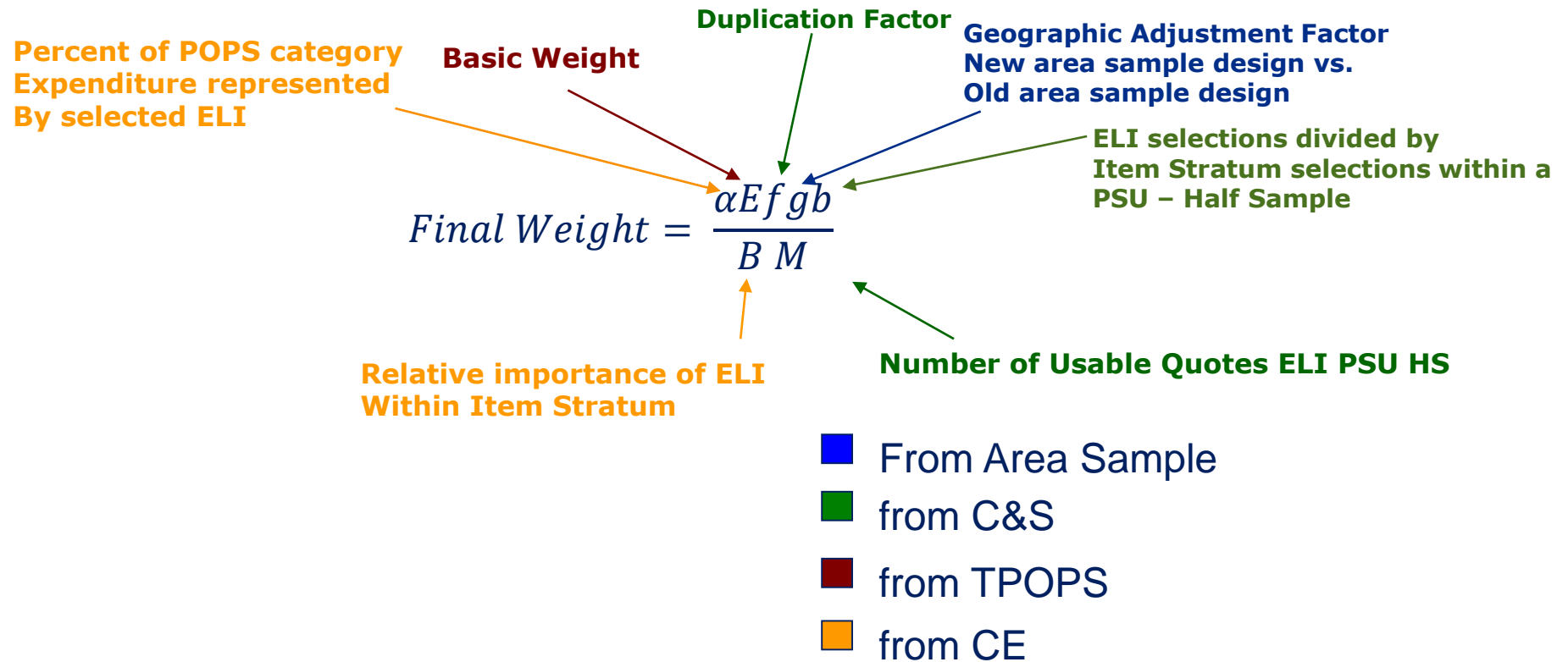
- The Item-Outlet Optimization Process specifies the number of item stratum selections and outlet selections
- Outlet samples are drawn from TPOPS frames
- ELIs are sampled within Item Strata using region level relative importances from CE
- Selected ELIs are matched to TPOPS categories producing a sample of items and outlets

# Commodities and Services

---

- Disaggregation is performed at initiation to select unique items which will be followed
- Unique items are priced every one or two months as long as possible or until the PSU/Item Stratum is rotated again

## FINAL SAMPLE WEIGHT FORMULA FOR C&S:



# Housing Survey

---

- The total available segments are allocated to areas.
- Segments (typically block groups) are sampled within an area proportional to rent expenditure plus OER (Owners' Equivalent Rent) expenditure.
- Note that only renters are sampled but they are used to represent both Rent and OER
- We get addresses for these segments from an address vendor.

# Housing Survey

---

- We estimate the percentage of renter occupied units and determine the number of addresses needed to obtain 5 renters in each segment
- We contract with a pre-screening vendor who does a mail survey of addresses – those who respond they own their property are screened out.
- We contact remaining addresses and initiate renter occupied housing units.

# Housing Survey

---

- We divide the sample into six panels.
- Each panel is priced twice a year at six month intervals.
- The collected rent is subject to adjustments and the economic rent is used with rent weights to produce Rent indexes.
- The rent is adjusted to the pure rent and combined with OER weights to produce OER indexes.



# Housing Survey

- The measure of size for segments is
- Size =  $1/(\text{PSU Probability of selection})^*$   
(Number of renters \* average rent + Number of owners \* OER)
- The estimate of OER is calculated by a model produced using data from the CE survey.
- Sample weight is apportioned to rent and OER according to percent of total cost in the segment for rent and the total cost in the segment for OER

# Themes in sampling

---

- The primary sampling method used throughout the CPI is systematic sampling probability proportional to size with a random start.
- This is often combined with sorting the units to be sampled in a specific order to create an implicit stratification.
- Sampling is done proportional to some measure of size wherever possible.

# Systematic Sampling PPS

---

- Used in selecting a PSU from each Stratum – State – Previously Selected cell chosen by controlled selection of the area sample.
- Used to select outlets from TPOPS frames. A sort order is used of largest, smallest, 2<sup>nd</sup> largest, 2<sup>nd</sup> smallest, etc.
- Used to select segments within an area. A sort order of county and average rent level is used.

# Systematic Sampling PPS

---

- Used in selecting addresses within a segment.
- Used in selecting ELIs within an Item Stratum

# Contact Information

---

William Johnson  
Supervisory Mathematical Statistician  
DPSM/Survey Research and Analysis  
202-691-6921  
[Johnson.bill@bls.gov](mailto:Johnson.bill@bls.gov)

