Price indices and dynamical expenditure shares

Preliminary Estimation of Chained CPI-U

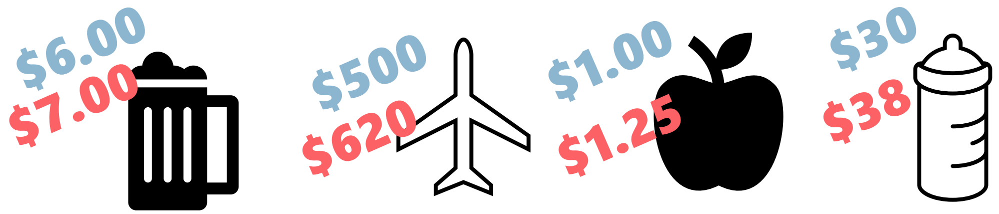
Kate Eckerle
Daniell Toth, Joshua Klick, Jeffrey Wilson
Bureau of Labor Statistics, OSMR, MSRC
FCSM
October 26th, 2023

Outline

- Price indices + consumer substitution
  - *Tornqvist Equation*
- BLS: two stage process
  - *Timeliness*
  - *Preliminary Chained CPI-U*
- New idea: Forecast monthly shares
  - *Vector timeseries model*
  - *With exogenous terms*

What is a price index?

- A price index is a number
- Should capture % change in cost of a set of items



- Inputs:

  -**prices** (what do people pay), before and now

  -**amounts** (how much do people buy)

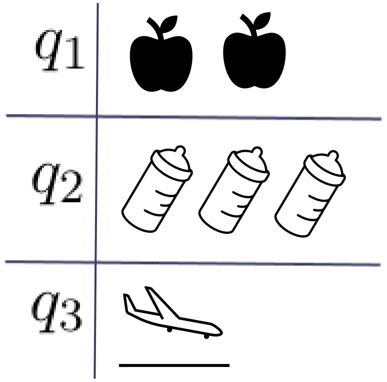
- How should you combine these into one number?

Example

- One way: **ratio of weighted averages** of prices

  How to weight by importance? Item quantities

$$\frac{q_1\ p_{\text{apple},t_1} + q_2\ p_{\text{formula},t_1} + q_3\ p_{\text{vacay},t_1} + \cdots}{q_1\ p_{\text{apple},t_0} + q_2\ p_{\text{formula},t_0} + q_3\ p_{\text{vacay},t_0} + \cdots}$$



- Expenditure = Quantity x Price

  total cost of a basket at today's prices, vs total cost of same basket at past prices

- Vector notation (items go down rows)

$$\frac{\mathbf{q} \cdot \mathbf{p}_t}{\mathbf{q} \cdot \mathbf{p}_{t-1}}$$

Lowe Index

$$\frac{\mathbf{q} \cdot \mathbf{p}_t}{\mathbf{q} \cdot \mathbf{p}_{t-1}}$$

If fixed quantity vector:

-is concurrent with time of past prices -> called Laspeyres

-is defined at or over a time period non-overlapping with times of prices, called ***Lowe***

we'll indicate this with a* b *subscript for "base period" (for us* ***precedes* t *and* t-1*)*

$$P_{t,t-1}^{\text{Lowe}} = \frac{\mathbf{q}_{\text{b}} \cdot \mathbf{p}_t}{\mathbf{q}_{\text{b}} \cdot \mathbf{p}_{t-1}}$$

Consumer Substitution

…but, when item price goes up, people tend to buy less than before.

They *substitute* relatively cheaper alternatives, to differing degrees for different items.

$$P^{\text{Lowe}}_{t,t-1} = \frac{\mathbf{q}_{\text{b}} \cdot \mathbf{p}_t}{\mathbf{q}_{\text{b}} \cdot \mathbf{p}_{t-1}}$$

Fixed quantity -> tend to overstate rise in cost of living when prices rise

Two Extremes

Absolute price insensitivity

Fixed quantity

weighted arithmetic mean

$$P^{\text{Lowe}}_{t,t-1} = \frac{\mathbf{q}_{\text{b}} \cdot \mathbf{p}_t}{\mathbf{q}_{\text{b}} \cdot \mathbf{p}_{t-1}}$$

"upper bound"

Perfect price sensitivity

Fixed Expenditure Share

weighted geometric mean

$$P^{\text{Geo}}_{t,t-1} = \prod_j \left(\frac{p_{j,t}}{p_{j,t-1}} \right)^{s_{j,\text{b}}}$$

items → j

$$s_{j,\text{b}} \equiv \frac{E_{j,\text{b}}}{\sum_k E_{k,\text{b}}}$$

"lower bound"

Bridge Between Extremes

Lowe + Geo Means connected by a continuous family of indices

Lloyd-Moulton

One parameter: $\sigma \in [0, 1]$ "Elasticity of Substitution"

$$P_{t,b}^{\mathrm{LM}} = \left(\sum_j s_{j,\mathrm{b}} \left(\frac{p_{j,t}}{p_{j,\mathrm{b}}} \right)^{1-\sigma} \right)^{1/(1-\sigma)}$$

$\sigma = 0$

$$P_{t,b}^{\mathrm{Lowe}} = \sum_j s_{j,\mathrm{b}} \frac{p_{j,t}}{p_{j,\mathrm{b}}}$$

$\sigma \to 1$

$$P_{t,b}^{\mathrm{Geo}} = \exp \left(\sum_j s_{j,\mathrm{b}} \log \left(\frac{P_{j,t}}{P_{j,\mathrm{b}}} \right) \right)$$

Tornqvist Formula

***geometric mean* index but with *dynamical quantity* information**

***monthly* expenditure shares**

$$P_{t,t-1}^{\text{Tornq}} = \prod_j \left(\frac{p_{j,t}}{p_{j,t-1}} \right)^{\frac{1}{2}(s_{j,t}+s_{j,t-1})}$$

$$s_{j,t} \equiv \frac{E_{j,t}}{\sum_k E_{k,t}}$$

$$s_{j,t} = \frac{p_{j,t}\, q_{j,t}}{\mathbf{p}_t \cdot \mathbf{q}_t}$$

The idea is that this would capture the "true" amount of substitution, which is item and time dependent.

BLS: two-stage process

- Inputs: **Prices** (establishment) + **Expenditures** (household)

  Market basket: **243** basic items (hierarchical)
  Geography: **32** areas (Primary Sampling Units of CE)

  item-areas: $j = 1, 2, \ldots, N$

- 1st stage: compute a price index for each **item-area** con j ination

  → 243 x 32 = 7,776 *basic price indices* $P_{j,t}$ ← most with **Geo Means**

- 2nd stage: **aggregate** to reflect desired broader group

  Broadest level for CPI-U: All items, U.S. city average

Aggregation

Plug basic price indices $P_{j,t}$ and aggregation weights into **ANOTHER price index** formula.

months: $t = 1, 2, \ldots, T$

item-areas: $j = 1, 2, \ldots, N$

$E_{j,\mathrm{b}}$ $E_{j,t}$

Regular CPI

$$P^{\mathrm{Lowe}}_{t,t-1} = \frac{\mathbf{Q}_{\mathrm{b}} \cdot \mathbf{P}_t}{\mathbf{Q}_{\mathrm{b}} \cdot \mathbf{P}_{t-1}}$$

Chained CPI

$$P^{\mathrm{Tornq}}_{t,t-1} = \prod_j \left(\frac{P_{j,t}}{P_{j,t-1}} \right)^{\frac{1}{2}(s_{j,t} + s_{j,t-1})}$$

capture upper-level substitution

*both plutocratic, not democratic $Q^{\mathrm{eff}}_j \equiv \frac{E_j}{P_j}$

Timeliness problem

Hold-up: monthly shares, NOT prices

CE data comes with ~one year lag

$$\mathbf{s}_t \equiv \mathbf{E}_t / |\mathbf{E}_t|_{\ell_1}$$

$$P^{\text{Tornq}}_{t,t-1} = \prod_j \left(\frac{P_{j,t}}{P_{j,t-1}} \right)^{\frac{1}{2}(s_{j,t}+s_{j,t-1})}$$

receive ~one year late!

While CPI-U final upon release, C-CPI-U issued as preliminary in month t.

Preliminary Estimate

2002 through 2014:

-calculated using **Geometric Means** (downward bias!)

Since Jan 2015:

-via **constant elasticity of substitution model** (i.e. based on LM)

$$P_{t,t-1}^{\text{CES}} = \frac{\left(\sum_j \left(s_{j,\sigma}\frac{P_{j,t}}{P_{j,v}}\right)^{1-\sigma}\right)^{1/(1-\sigma)}}{\left(\sum_j \left(s_{j,\sigma}\frac{P_{j,t-1}}{P_{j,v}}\right)^{1-\sigma}\right)^{1/(1-\sigma)}}$$

$$\sigma = .6$$

Why .6?

Originally based on work of Greenlees (2010)

- sigma from Feenstra-Reinsdorf model:

$$\Delta \log(s_{j,\mathrm{Y}}) = \alpha + (1-\sigma)\Delta \log(r_{j,\mathrm{Y}}) + \epsilon_j$$

$$\Delta \log(s_{j,\mathrm{Y}}) \equiv \log(s_{j,\mathrm{Y}}) - \log(s_{j,\mathrm{Y}-1})$$

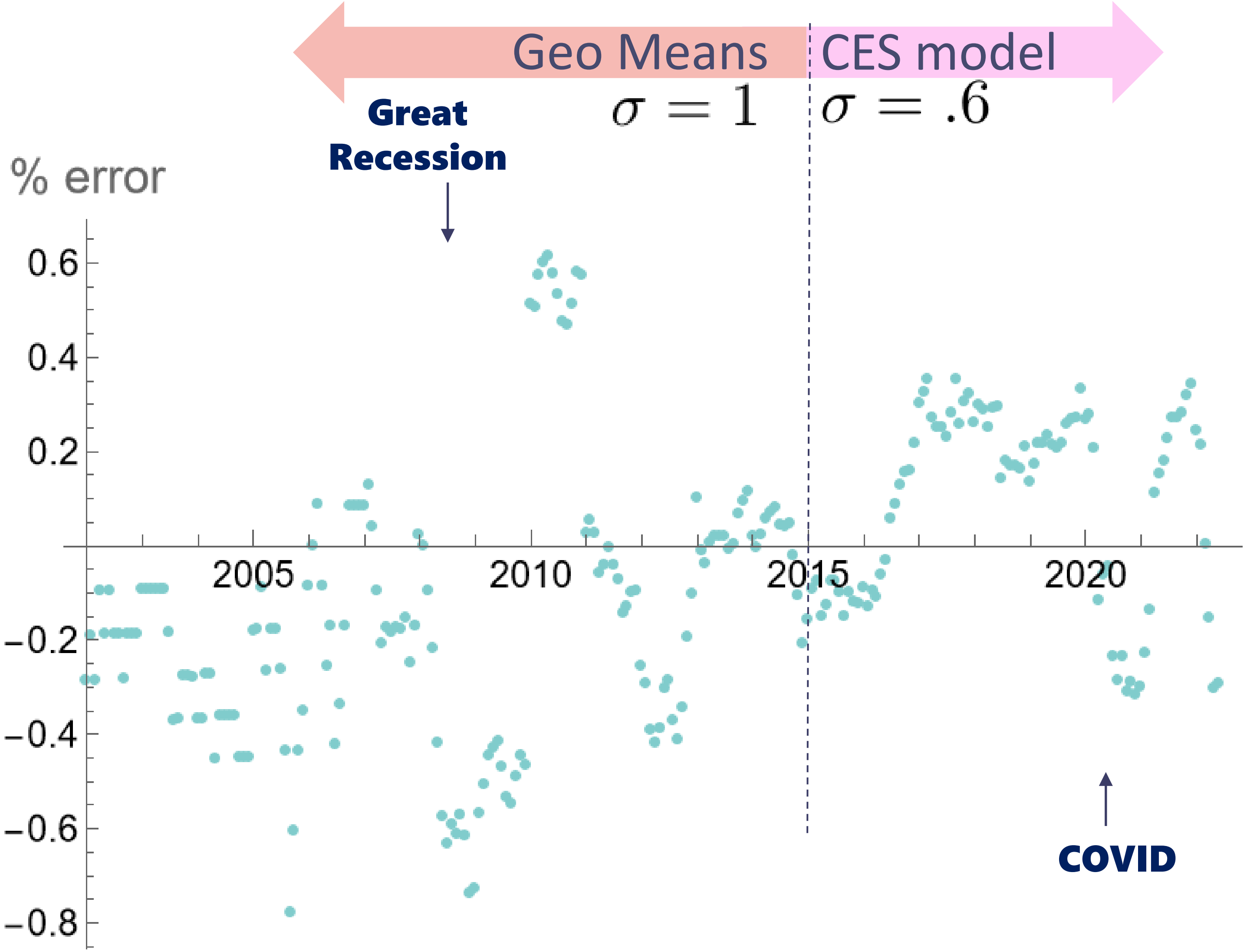
$$r_{j,\mathrm{Y}} \equiv \frac{P_{j,\mathrm{Y}}}{P_{j,\mathrm{Y\text{-}1}}}$$

annualized shares + prices relatives to obtain **stable** estimates (Y=year)

- Fixed value selected:

$$\sigma = .6$$

Can we do better?



Surely an improvement, but assumes stability of this effective/net sigma…

- ❖Errors correlated
- ❖Shocks
  - Great Recession
  - COVID economy
- ❖Seasonal effects

Alternative: Forecast monthly shares

If one drops Feenstra-Reinsdorf, optimal sigma varies enormously year-to-year, and month-to-month. (Robert Cage + Joshua Klick)

❖Cage + Wilson (2009):

Forecast monthly budget shares -> plug into Tornqvist

- Univariate time series models (ARIMA)
- Worked well for highly seasonal

…but, this will capture budget share dynamics that can be predicted *without* knowing about prices, or what other shares are doing.

Our idea

❖Whole idea behind superlative indices is that the state of relative prices changes affects allocation of spending across items…

*Build an expenditure **forecast model** that **incorporates** item-area **price** indices (which recall are not lagged).*

1. Avoid constraints

   -model item-area expenditures, not shares

2. Allow interactions among item-area expenditures

   *-univariate -> **vector** timeseries model*

3. Allow couplings to item-area prices (external)

   -exogenous terms

Vector ARMAX

$$B\ y_{j,t} = y_{j,t-1}$$

$$\Delta = I - B$$

model $y_{j,t} \equiv \Delta \log(E_{j,t})$ **with inputs** $x_{j,t} \equiv \Delta \log(P_{j,t})$

$$\underbrace{\left(I - \sum_{k=1}^{p} \boldsymbol{\phi}_p B^p\right)}_{\hat{O}_1} \mathbf{y}_t = \underbrace{\left(\sum_{k=0}^{r} \boldsymbol{\beta}_k B^k\right)}_{\hat{O}_2} \mathbf{x}_t + \underbrace{\left(I - \sum_{k=1}^{q} \boldsymbol{\theta}_k B^k\right)}_{\hat{O}_3} \boldsymbol{\epsilon}_t$$

$$\mathbf{y}_t = \hat{O}_1^{-1}\hat{O}_2\ \mathbf{x}_t + \hat{O}_1^{-1}\hat{O}_3\ \boldsymbol{\epsilon}_t + \boldsymbol{\xi}_t$$

$$\xi_{j,t} = \sum_{n=1}^{6} A_{j,n} \sin\left(2\pi \frac{nt}{12} + \delta_{j,n}\right)$$

Contact Information

Kate Eckerle
Research Mathematical Statistician
MSRC, OSMR

eckerle.kate@bls.gov