What does consumer heterogeneity mean for measuring changes in the cost of living?

Robert S. Martin

Office of Prices and Living Conditions

FCSM Conference
3/9/2018

Disclaimer

The views expressed herein are those of the author and not necessarily those of the Bureau of Labor Statistics or the U.S. Department of Labor.

Contents

Motivation: COGI vs. COLI

- The Consumer Price Index (CPI-U) and Chained Consumer Price Index (C-CPI-U) aim to measure changes in the *cost of living*
- Cost of Living Index (COLI)
  - Inflation is measured as the change in expenditures required to maintain a fixed standard of well-being
- Cost of Goods Index (COGI)
  - Inflation is measured as the change in expenditure required to continue consuming a fixed basket of goods

Motivation: Nonhomotheticity

- Need (at least implicit) model of consumer preferences
- Tractability v. generality
- Homotheticity: assumption that substitution patterns do not vary with income
- Heterogeneity in expenditures well-documented
- Is homotheticity a decent approximation for aggregate indexes?
  - National Research Council (2002): literature tends to show low variation in inflation rates across groups in long run
  - Short run or more recent differences
  - Much existing research uses individual or group-specific category weights

- Hottman and Monarch (2018) relax homotheticity in economic model and describe differences in import price inflation for different income groups
- Potentially very important for consumer cost of living measurement
- I replicate HM's method of estimating nonhomothetic consumer demand using household scanner data
- Preview: Some indications against homotheticity but model may not be suited for my scanner data

Theory: Consumer preferences

- Constant Elasticity of Substitution (CES) utility over sectors s (closely related goods, i.e. "Coffee")

$$V_{ht} = \left[\sum_{s \in S} \varphi_{hst}^{\frac{\sigma-1}{\sigma}} Q_{hst}^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \qquad (1)$$

- Within sectors, generalized-CES over varieties v (i.e. "Folgers 24.2 oz Colombian - Ground") similar to Stone-Geary

$$Q_{hst} = \left[\sum_{v \in G_s} \varphi_{vt}^{\frac{\sigma^s-1}{\sigma^s}} \left(q_{hvt} - \alpha_v\right)^{\frac{\sigma^s-1}{\sigma^s}} \right]^{\frac{\sigma^s}{\sigma^s-1}}, \qquad (2)$$

- Like Stone-Geary, α_v loosely interpreted as subsistence quantity of good v, but allowed to be negative

Theory: Consumer preferences

- Constant Elasticity of Substitution (CES) utility over sectors s (closely related goods, i.e. "Coffee")

$$V_{ht} = \left[\sum_{s \in S} \varphi_{hst}^{\frac{\sigma-1}{\sigma}} Q_{hst}^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \tag{1}$$

- Within sectors, generalized-CES over varieties v (i.e. "Folgers 24.2 oz Colombian - Ground") similar to Stone-Geary

$$Q_{hst} = \left[\sum_{v \in G_s} \varphi_{vt}^{\frac{\sigma^s-1}{\sigma^s}} \left(q_{hvt} - \alpha_v\right)^{\frac{\sigma^s-1}{\sigma^s}} \right]^{\frac{\sigma^s}{\sigma^s-1}}, \tag{2}$$

- Like Stone-Geary, α_v loosely interpreted as subsistence quantity of good v, but allowed to be negative

- Model accommodates entering and exiting varieties (non-continuers)
- Regularity conditions for individual demand include $\alpha_v \leq 0$ for non-continuers
- Using aggregate data for a price index requires $\alpha_v = 0$ for non-continuers
- Demand curves, budget shares, and cost function (P_t) follow from maximizing Eq.'s (2) and (1) subject to budget constraint (which depends on prices p_{vt})

- Aggregate COLI from period t to $t+i$ in HM model:

$$\frac{P_{t+i}}{P_t} = \frac{\left(\sum_{s\in S} \varphi_{st+i}^{\sigma-1} P_{st+i}^{1-\sigma}\right)^{\frac{1}{\sigma-1}}}{\left(\sum_{s\in S} \varphi_{st}^{\sigma-1} P_{st}^{1-\sigma}\right)^{\frac{1}{\sigma-1}}} \left(\frac{Y_k - \sum_{s\in S}\sum_{v\in G_s}(\alpha_v n_t) p_{vt}}{Y_k}\right) + \frac{\sum_{s\in S}\sum_{v\in G_s}(\alpha_v n_{t+i}) p_{vt+i}}{Y_k} \quad (3)$$

- Sectoral price index P_{st} depends on σ^s and variety-level demand shifters φ_{vt}
- Compare to Laspeyres

$$\frac{P_{t+i}}{P_t} = \frac{\sum_{r\in S}\sum_{v\in G_s} p_{vt+i} q_{vt}}{\sum_{r\in S}\sum_{v\in G_s} p_{vt} q_{vt}} \quad (4)$$

- Or Törnqvist

$$\frac{P_{t+i}}{P_t} = \prod_{s\in S}\prod_{v\in G_s}\left(\frac{p_{vt+i}}{p_{vt}}\right)^{w_v} \quad (5)$$

where $w_v = \frac{1}{2}\left(\frac{p_{vt+i} q_{vt+i}}{\sum_{r\in S}\sum_{v\in G_s} p_{vt+i} q_{vt+i}} + \frac{p_{vt} q_{vt}}{\sum_{r\in S}\sum_{v\in G_s} p_{vt} q_{vt}}\right)$

- Aggregate COLI from period t to $t+i$ in HM model:

$$\frac{P_{t+i}}{P_t} = \frac{\left(\sum_{s\in S} \varphi_{st+i}^{\sigma-1} P_{st+i}^{1-\sigma}\right)^{\frac{1}{\sigma-1}}}{\left(\sum_{s\in S} \varphi_{st}^{\sigma-1} P_{st}^{1-\sigma}\right)^{\frac{1}{\sigma-1}}} \left(\frac{Y_k - \sum_{s\in S}\sum_{v\in G_s}(\alpha_v n_t)p_{vt}}{Y_k}\right) + \frac{\sum_{s\in S}\sum_{v\in G_s}(\alpha_v n_{t+i})p_{vt+i}}{Y_k} \quad (3)$$

- Sectoral price index P_{st} depends on σ^s and variety-level demand shifters φ_{vt}
- Compare to Laspeyres

$$\frac{P_{t+i}}{P_t} = \frac{\sum_{r\in S}\sum_{v\in G_s} p_{vt+i}q_{vt}}{\sum_{r\in S}\sum_{v\in G_s} p_{vt}q_{vt}} \quad (4)$$

- Or Törnqvist

$$\frac{P_{t+i}}{P_t} = \prod_{s\in S}\prod_{v\in G_s}\left(\frac{p_{vt+i}}{p_{vt}}\right)^{w_v} \quad (5)$$

where $w_v = \frac{1}{2}\left(\frac{p_{vt+i}q_{vt+i}}{\sum_{r\in S}\sum_{v\in G_s} p_{vt+i}q_{vt+i}} + \frac{p_{vt}q_{vt}}{\sum_{r\in S}\sum_{v\in G_s} p_{vt}q_{vt}}\right)$

Tradeoffs

Törnqvist index	HM index
▶ Flexible (see Diewert 1976)	▶ Assumes specific model
▶ No parameters to estimate	▶ Need to estimate σ, σ^s's, α_v's, φ_{vt}'s, φ_{hst}'s
▶ COLI interpretation requires homotheticity	▶ Explicitly allows nonhomotheticity

Theory: Completing the model

- Market demand with n_t consumers in period t

$$q_{vt} = (\alpha_v n_t) + \left(\frac{p_{vt}^{-\sigma^s} \varphi_{vt}^{\sigma^s - 1}}{\sum_{j \in G_s} p_{jt}^{1-\sigma^s} \varphi_{jt}^{\sigma^s - 1}} \right) \left(Y_t - \sum_{j \in G_s} (\alpha_j n_t) p_{jt} \right), \tag{6}$$

- Market structure for firms/varieties is monopolistic competition
- Price = Markup × Marginal Cost

$$p_{vt} = \frac{\varepsilon_{vt}}{\varepsilon_{vt} - 1} \delta_{vt} (1 + \omega_s) q_{vt}^{\omega_s} \tag{7}$$

where $\varepsilon_{vt} = \left(\frac{q_{vt} - \alpha_v n_t}{q_{vt}} \right) \sigma^s$ is the elasticity of demand perceived by the firm and $\frac{\varepsilon_{vt}}{\varepsilon_{vt} - 1}$ is the firm's markup

Theory: Estimation

- Carried out sector-by-sector using price (p_{vt}), sales ($p_{vt} q_{vt}$) data for each variety v and time period t
- To reduce dimensionality, HM model α_v as:

$$\alpha_v = \frac{1}{n_1} \left[\beta_s^j \min_t \left(q_{vt} | q_{vt} > 0 \right) \right], \quad \text{for } v \in G_s^j, j \in \{C, E\} \quad (8)$$

  where G_s^C = set of "continuing" varieties, and G_s^E = set of "non-continuing" varieties
- Two parameters to estimate: β_s^C and β_s^E
- Model regularity requires $\beta_s^C < 1$ and $\beta_s^E \leq 0$. Aggregation requires $\beta_s^E = 0$

- First stage of HM's procedure extends the approach of Feenstra (1994) to estimate the σ^s, ω_s, β_s^C, and β_s^E
- Simultaneity bias
- For market demand equation (6) and firm pricing (supply) equation (7), multiply each by p_{vt}, take logs, and difference both over time and with respect to a reference variety k within the same sector s

$$\Delta^{k,t} \ln \left(p_{vt} q_{vt} - \alpha_v n_t p_{vt}\right) = (1 - \sigma^s)\Delta^{k,t} \ln p_{vt} + \nu_{vt} \tag{9}$$

$$\Delta^{k,t} \ln p_{vt} = \frac{\omega_s}{1+\omega_s}\Delta^{k,t} \ln \left(p_{vt} q_{vt}\right) + \frac{1}{1+\omega^s}\Delta^{k,t} \ln \left(\frac{\varepsilon_{vt}}{\varepsilon_{vt} - 1}\right) + \kappa_{vt} \tag{10}$$

- $\nu_{vt} = (1 - \sigma^s)\left[\Delta^t \ln \varphi_{kt} - \Delta^t \ln \varphi_{vt}\right]$ and $\kappa_{vt} = \frac{1}{1+\omega^s}\left[\Delta^t \ln \delta_{vt} - \Delta^t \ln \delta_{kt}\right]$ are the unobserved errors

- Identifying assumption: for each v, double differenced supply and demand errors ν_{vt} and κ_{vt} are orthogonal $\Rightarrow$ moment equation for each v:

$$G(\beta_s) = \mathbb{E}\left[u_{vt}(\beta_s)\right] = 0, \tag{11}$$

where $\beta_s = \left(\sigma^s, \beta_s^C, \beta_s^E, \omega_s\right)'$, and $u_{vt} = \nu_{vt}\kappa_{vt}$.

- Need at least four varieties in a sector to identify β_s
- Generalized Methods of Moments (GMM): Stack sample analogs of varieties into vector $G^*(\beta_s)$, then

$$\hat{\beta}_s = \underset{\beta_s}{\operatorname{argmin}} \left\{ G^*(\beta_s)' W G^*(\beta_s) \right\} \tag{12}$$

with weighting matrix W

- Writing Eq. (11) in terms of observables and re-arranging, we have for each variety v:

$$
\begin{aligned}
\mathbb{E}\left[\left(\Delta^{k,t}\ln p_{vt}\right)^2\right] = & \frac{\omega_s}{1+\omega_s}\mathbb{E}\left[\Delta^{k,t}\ln\left(p_{vt}q_{vt}\right)\Delta^{k,t}\ln p_{vt}\right] \\
& -\frac{1}{\sigma^s-1}\mathbb{E}\left[\Delta^{k,t}\ln p_{vt}\Delta^{k,t}\ln\left(p_{vt}q_{vt}-\alpha_v n_t p_{vt}\right)\right] \\
& +\frac{\omega_s}{(1+\omega_s)(\sigma^s-1)}\mathbb{E}\left[\Delta^{k,t}\ln\left(p_{vt}q_{vt}\right)\Delta^{k,t}\ln\left(p_{vt}q_{vt}-\alpha_v n_t p_{vt}\right)\right] \\
& +\frac{1}{1+\omega_s}\mathbb{E}\left[\Delta^{k,t}\ln p_{vt}\Delta^{k,t}\ln\left(\frac{\varepsilon_{vt}}{\varepsilon_{vt}-1}\right)\right] \\
& +\frac{1}{(1+\omega_s)(\sigma^s-1)}\mathbb{E}\left[\Delta^{k,t}\ln\left(p_{vt}q_{vt}-\alpha_v n_t p_{vt}\right)\Delta^{k,t}\ln\left(\frac{\varepsilon_{vt}}{\varepsilon_{vt}-1}\right)\right]
\end{aligned}
\tag{13}
$$

- Writing Eq. (11) in terms of observables and re-arranging, we have for each variety v:

$$\mathbb{E}\left[\left(\Delta^{k,t}\ln p_{vt}\right)^2\right] = \frac{\omega_s}{1+\omega_s}\mathbb{E}\left[\Delta^{k,t}\ln\left(p_{vt}q_{vt}\right)\Delta^{k,t}\ln p_{vt}\right]$$
$$-\frac{1}{\sigma^s-1}\mathbb{E}\left[\Delta^{k,t}\ln p_{vt}\Delta^{k,t}\ln\left(p_{vt}q_{vt}-\alpha_v n_t p_{vt}\right)\right]$$
$$+\frac{\omega_s}{(1+\omega_s)(\sigma^s-1)}\mathbb{E}\left[\Delta^{k,t}\ln\left(p_{vt}q_{vt}\right)\Delta^{k,t}\ln\left(p_{vt}q_{vt}-\alpha_v n_t p_{vt}\right)\right]$$
$$+\frac{1}{1+\omega_s}\mathbb{E}\left[\Delta^{k,t}\ln p_{vt}\Delta^{k,t}\ln\left(\frac{\varepsilon_{vt}}{\varepsilon_{vt}-1}\right)\right]$$
$$+\frac{1}{(1+\omega_s)(\sigma^s-1)}\mathbb{E}\left[\Delta^{k,t}\ln\left(p_{vt}q_{vt}-\alpha_v n_t p_{vt}\right)\Delta^{k,t}\ln\left(\frac{\varepsilon_{vt}}{\varepsilon_{vt}-1}\right)\right] \tag{13}$$

Replication: Data

- HM use supplier trade data from Linked-Longitudinal Firm Trade Transaction Database (LFTTD) based on customs information
- I use the Homescan database from Nielsen covering 2008Q1 to 2010Q4
- ∼60,000 households in the sample per year who scan barcodes to record transactions
  - Nielsen provides frequency weights for projections to U.S. population
- Prices, quantities, bar codes (UPC), brand information, some characteristics
- Demographic info for households

- Coverage is mainly packaged food and general merchandise
  - Only about 13% of U.S. consumer expenditures (BLS Consumer Expenditure Survey)
- Roughly 700,000 UPC's per year in 122 product groups (sectors)

Replication: Method

- I estimate the model both unconstrained and with $\beta_s^C < 1$ and $\beta_s^E \leq 0$ imposed
- HM implement other constraints/penalties in estimation which I have not yet replicated

Results

- With different data, I do not expect to replicate HM's results exactly
- E.g. there should be more substitution between varieties at UPC level than between varieties at HS10-firm level

Table: Summaries of σ^s, ω_s estimates

	HM	Replications with Homescan	
	(1)	(2)	(3)
σ^s			
10%	3.06	17.97	14.38
Median	4.93	25.27	21.59
90%	8.59	46.77	44.82
ω_s			
10%	0.16	0.56	0.50
Median	0.44	0.74	0.67
90%	1.59	1.12	1.00
Sectors	980	113	111
Constraints	Yes	No	$\beta_s^C < 1$, $\beta_s^E \leq 0$

Table: Summaries of β_s^C, β_s^E estimates

	HM	Replications with Homescan	
	(1)	(2)	(3)
β_s^C			
10%	9.96E-05	0.66	0.56
Median	0.33	0.84	0.79
90%	0.39	0.95	0.99
β_s^E			
10%	-5.97E-05	0.83	-1.285E-15
Median	-2.55E-09	0.90	0.00
90%	-1.08E-10	0.97	0.00
Sectors	980	113	111
Constraints	Yes	No	$\beta_s^C < 1$, $\beta_s^E \leq 0$

BLS

Table: Summaries of sales-weighted average elasticities and markups

	HM	Replications with Homescan	
	(1)	(2)	(3)
P.E.D. (ε_{vt})			
10%	.	10.36	9.39
Median	.	14.63	14.67
90%	.	30.32	29.35
Markups ($\frac{\varepsilon_{vt}}{\varepsilon_{vt}-1}$)			
10%	1.13	1.05	0.18
Median	1.25	1.12	1.10
90%	1.48	1.18	1.15
Sectors	980	113	111
Constraints	Yes	No	$\beta_s^C < 1$, $\beta_s^E \leq 0$

Discussion

- Homescan results indeed imply greater substitution between UPCs within same product group
- The estimated P.E.D. from the Homescan data are still somewhat high relative to other work (i.e. Broda and Weinstein 2010)
- $\beta_s^E \geq 0$ unless otherwise constrained
- I also find evidence against homothetic CES, but $\beta_s^E \geq 0$ doesn't fit generalized-CES
- Alternative classification of continuing and non-continuing varieties?

Conclusion

- More work is needed to address variety-level substitution heterogeneity in COLI framework
- HM also find cross-sector nonhomotheticity may play a larger role in import price index heterogeneity
  - Perhaps worth revisiting group or household-specific expenditure weights for the CPI
- Broader issue: what to do when "the model is fighting the data" in terms of theoretical constraints

CONTACT INFORMATION

Robert S. Martin
Division of Price and Index Number Research
Bureau of Labor Statistics
2 Massachusetts Ave, NE
Washington, DC 20212
Email: Martin.Robert@bls.gov