A Distributional PCE Price Index From Aggregate Data

Philipp Hochmuth
Stockholm University

Markus Pettersson
Stockholm University

Christoffer J. Weissert
Danmarks Nationalbank

DN-NB-BuBa conference
Structural Changes and the Implications for Inflation

May 7-8, 2024

Disclaimer: Views expressed are solely the authors' and do not represent Danmarks Nationalbank.

Focus of the paper: inequality in costs of living

Key question: Do costs of living change equally for everyone over time?

Large body of literature tries to answer this question

Amble and Stewart (1994), Garner et. al (1996), Crawford and Smith (2002), Hobijn and Lagakos (2005), McGranahan and Paulson (2006), Kaplan and Schulhofer-Wohl (2017), Argente and Lee (2021), Jaravel (2019, 2021), Klick and Stockburger (2021), Jaravel and Lashkari (2024), Baqaee, Burstein and Koike-Mori (2024) . . .

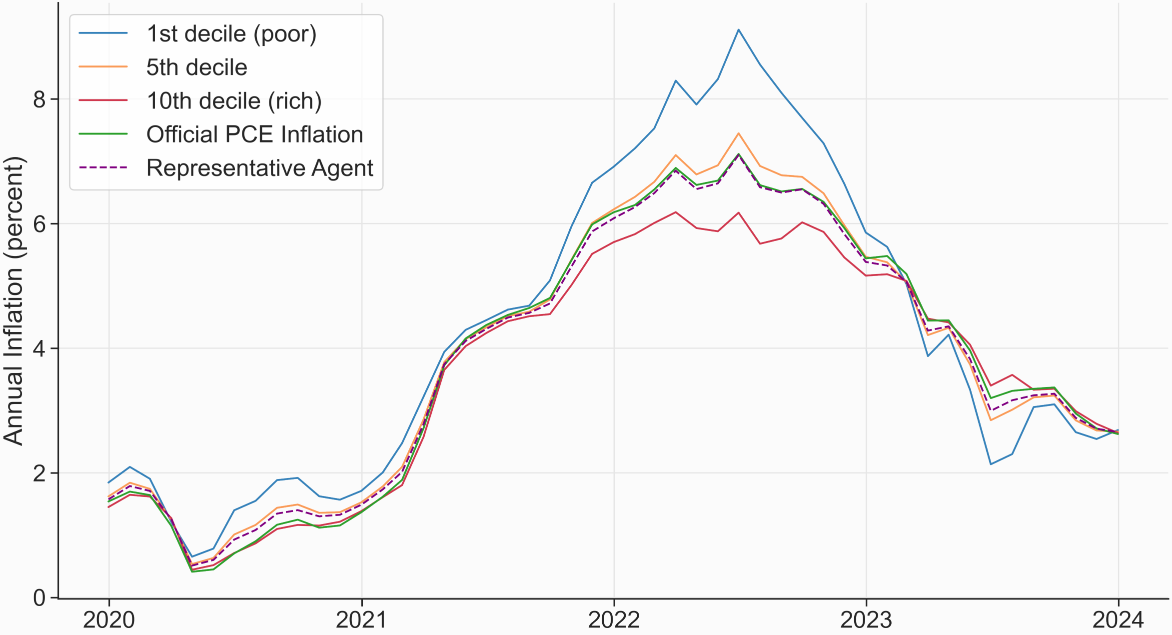
$\Longrightarrow$ all studies rely on detailed microdata

Microdata not an issue, at all, but some limitations:

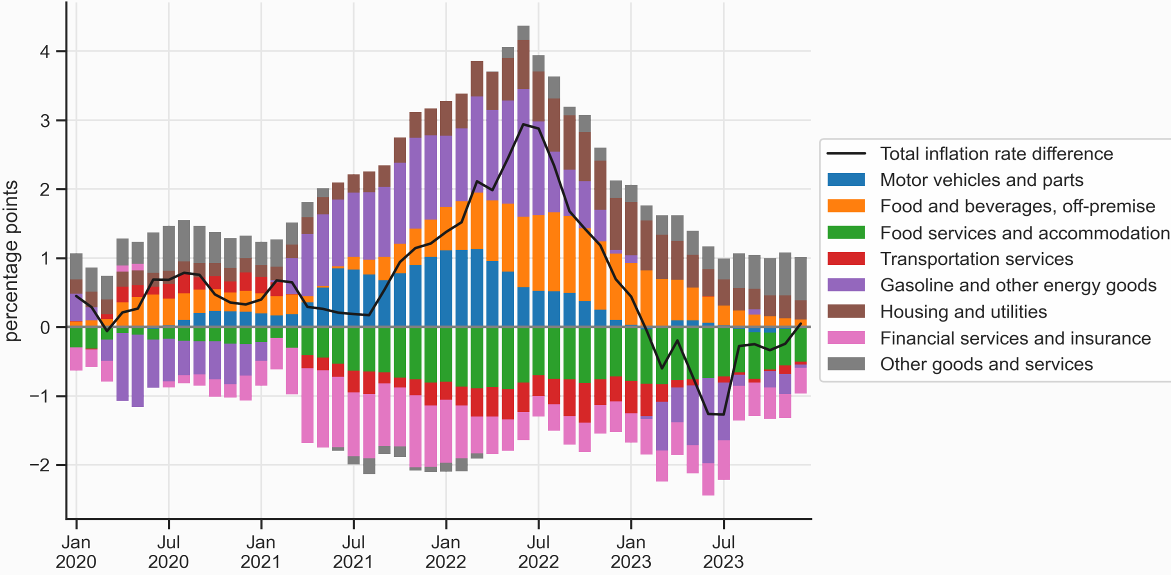
a) not available in many countries
b) not always available far back in time
c) not always readily available, only with some time lag

This paper: a distributional cost-of-living index from aggregate data

Latest inflation developments



Decomposition: poor-rich inflation rate gap



Outline

1) Brief idea behind the framework and data
2) Long-run inflation inequality trends covering last 65 years
3) Two additional insights

Framework

Nonhomothetic cost-of-living index

Nonhomothetic PIGL cost-of-living index is given by

$$P(u, \boldsymbol{p}_t, \boldsymbol{p}_s) = \left[(1 - w_{Ds})P_{Bt}^{\varepsilon} + w_{Ds}P_{Dt}^{\varepsilon}\right]^{\frac{1}{\varepsilon}} \left(\frac{P_{Ht}}{P_{Bt}}\right)^{\rho_{t,s}}$$

Implementation requires four inputs

(i) Three price aggregates: necessities P_D; luxuries P_B; homothetic goods P_H.

(ii) Expenditure share on necessities w_D.

(iii) Expenditure share on homothetic goods w_H (to get ρ),

(iv) Nonhomotheticity parameter ε

Nonhomothetic cost-of-living index

Nonhomothetic PIGL cost-of-living index is given by

$$P(u, \boldsymbol{p}_t, \boldsymbol{p}_s) = \left[(1 - w_{Ds}) P_{Bt}^{\varepsilon} + w_{Ds} P_{Dt}^{\varepsilon}\right]^{\frac{1}{\varepsilon}} \left(\frac{P_{Ht}}{P_{Bt}}\right)^{\rho_{t,s}}$$

Implementation requires four inputs

(i) Three price aggregates: necessities P_D; luxuries P_B; homothetic goods P_H.

(ii) Expenditure share on necessities w_D.

(iii) Expenditure share on homothetic goods w_H (to get ρ),

(iv) Nonhomotheticity parameter ε

- Price aggregates are obtained from a classification of goods and choice of aggregator

  ▸ Classification: estimation ▸ Nonhomothetic Törnqvist index

- Classification also gives (aggregate) expenditure shares

Nonhomothetic cost-of-living index

Nonhomothetic PIGL cost-of-living index is given by

$$P(u, \boldsymbol{p}_t, \boldsymbol{p}_s) = \left[(1 - w_{Ds})P_{Bt}^{\varepsilon} + w_{Ds}P_{Dt}^{\varepsilon}\right]^{\frac{1}{\varepsilon}} \left(\frac{P_{Ht}}{P_{Bt}}\right)^{\rho_{t,s}}$$

Implementation requires four inputs

(i) Three price aggregates: necessities P_D; luxuries P_B; homothetic goods P_H.

(ii) Expenditure share on necessities w_D.

(iii) Expenditure share on homothetic goods w_H (to get ρ),

(iv) Nonhomotheticity parameter ε

- Necessity expenditure share equation

$$w_D = \underbrace{\nu \left(\frac{G(\boldsymbol{p})}{e}\right)^{\varepsilon}}_{\to 0 \text{ as } e \to \infty}, \qquad \frac{\partial \ln w_D}{\partial \ln e} = -\varepsilon.$$

Nonhomothetic cost-of-living index

Nonhomothetic PIGL cost-of-living index is given by

$$P(u, \boldsymbol{p}_t, \boldsymbol{p}_s) = \left[(1 - w_{Ds}) P_{Bt}^{\varepsilon} + w_{Ds} P_{Dt}^{\varepsilon}\right]^{\frac{1}{\varepsilon}} \left(\frac{P_{Ht}}{P_{Bt}}\right)^{\rho_{t,s}}$$

Implementation requires four inputs

(i) Three price aggregates: necessities P_D; luxuries P_B; homothetic goods P_H.

(ii) Expenditure share on necessities w_D.

(iii) Expenditure share on homothetic goods w_H (to get ρ),

(iv) Nonhomotheticity parameter ε

- Necessity expenditure share equation

$$w_D = \nu \underbrace{\left(\frac{G(\boldsymbol{p})}{e}\right)^{\varepsilon}}_{\to 0 \text{ as } e \to \infty}, \qquad \frac{\partial \ln w_D}{\partial \ln e} = -\varepsilon.$$

Gives estimating equation for ε if microdata is available.
– but **what if only macrodata is available?**

Nonhomothetic cost-of-living index

Nonhomothetic PIGL cost-of-living index is given by

$$P(u, \boldsymbol{p}_t, \boldsymbol{p}_s) = \left[(1 - w_{Ds}) P_{Bt}^{\varepsilon} + w_{Ds} P_{Dt}^{\varepsilon}\right]^{\frac{1}{\varepsilon}} \left(\frac{P_{Ht}}{P_{Bt}}\right)^{\rho_{t,s}}$$

Implementation requires four inputs

(i) Three price aggregates: necessities P_D; luxuries P_B; homothetic goods P_H.

(ii) Expenditure share on necessities w_D.

(iii) Expenditure share on homothetic goods w_H (to get ρ),

(iv) Nonhomotheticity parameter ε ▸ Estimating equation

Only macrodata? Consistent aggregation $\implies$ micro and macro behavior is tied together

$$\overline{w}_D \equiv \frac{1}{N} \int_0^N \frac{e_h}{\overline{e}} w_{Dh} \, dh = \widetilde{\nu} \kappa^{-\varepsilon} \left(\frac{P_F}{\overline{e}} \cdot \frac{P_D}{P_B}\right)^{\varepsilon}$$

- $\overline{w}_D$: Expenditure-weighted **average** necessity expenditure share
- $\widetilde{\nu}$: Scale parameter
- κ: Inequality measure
- $\overline{e}$: **Per-capita** expenditures
- P_F, P_D, P_B: Basket price indices

Empirical implementation

Data

Personal Consumption Expenditures

- Aggregate U.S. expenditures and prices
- January 1959 to December 2023
- 71 consumption categories

▸ Classification results

Distribution

▶ Garner *et al.* (2022):

- *Single* distribution of *expenditures* in 2019
- Point estimates: deciles, top 5 and 1 pct.

▸ Details

Data

Personal Consumption Expenditures

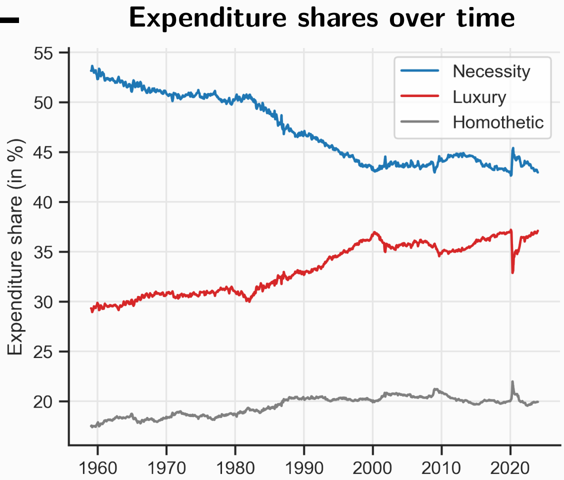
- Aggregate U.S. expenditures and prices
- January 1959 to December 2023
- 71 consumption categories

▸ Classification results

Distribution

- Garner *et al.* (2022):
  - *Single* distribution of *expenditures* in 2019
  - Point estimates: deciles, top 5 and 1 pct.

▸ Details



Data

Personal Consumption Expenditures

- Aggregate U.S. expenditures and prices
- January 1959 to December 2023
- 71 consumption categories

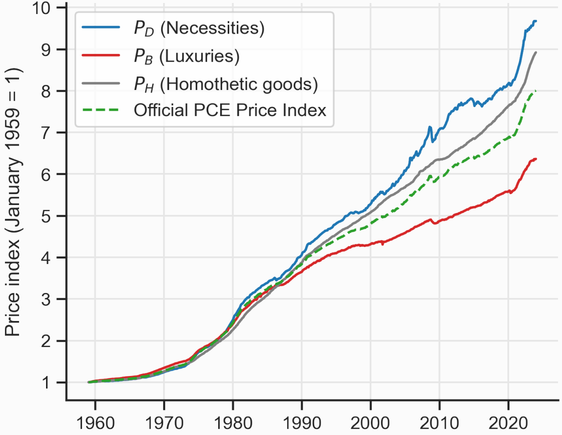
Classification results

Distribution

- Garner *et al.* (2022):
  - *Single* distribution of *expenditures* in 2019
  - Point estimates: deciles, top 5 and 1 pct.

Details

Basket prices over time



Data

Personal Consumption Expenditures

- Aggregate U.S. expenditures and prices
- January 1959 to December 2023
- 71 consumption categories

▸ Classification results

Clear rejection of homotheticity

Basline estimate of $\varepsilon = 0.702^{***}$

▸ ε estimation results

Distribution

▶ Garner *et al.* (2022):

- *Single* distribution of *expenditures* in 2019
- Point estimates: deciles, top 5 and 1 pct.

▸ Details

Main Results

Long-run inflation inequality

Inflation inequality in the last 65 years

850 pct.
Poorest

644 pct.
Richest

Long-run inflation inequality

Inflation inequality in the last 65 years

850 pct.
Poorest

644 pct.
Richest

$\Downarrow$

Long-run annual inflation rate inequality

3.53 pct.
Poorest

3.14 pct.
Richest

Long-run inflation inequality

Inflation inequality in the last 65 years

850 pct.
Poorest

644 pct.
Richest

$\Downarrow$

Long-run annual inflation rate inequality

0.39 pp.
Poor-rich gap

Long-run inflation inequality

Inflation inequality in the last 65 years

850 pct.
Poorest

644 pct.
Richest

$\Downarrow$

Long-run annual inflation rate inequality

0.39 pp.
Poor-rich gap

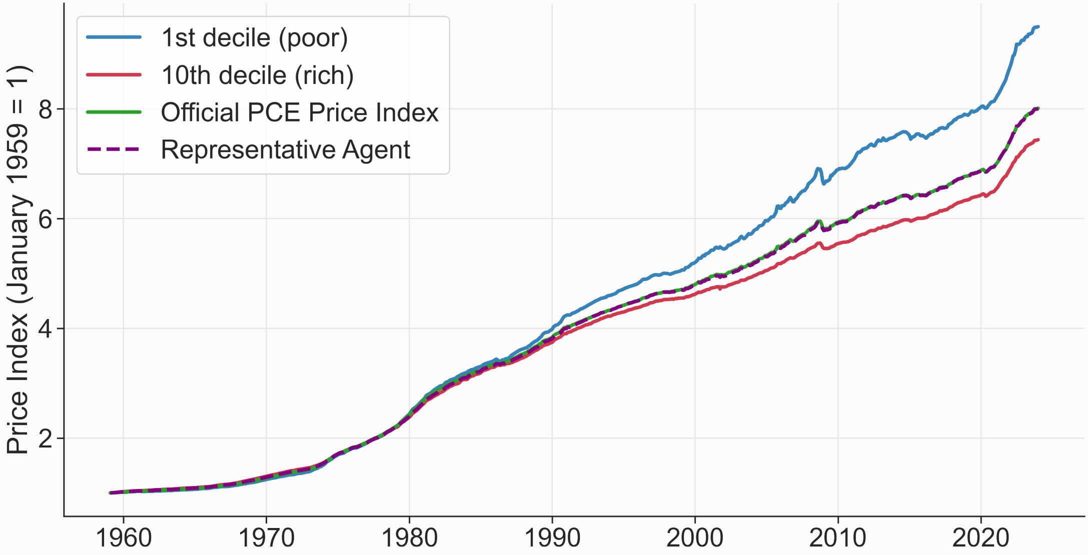
$\Downarrow$

PCE real consumption growth bias

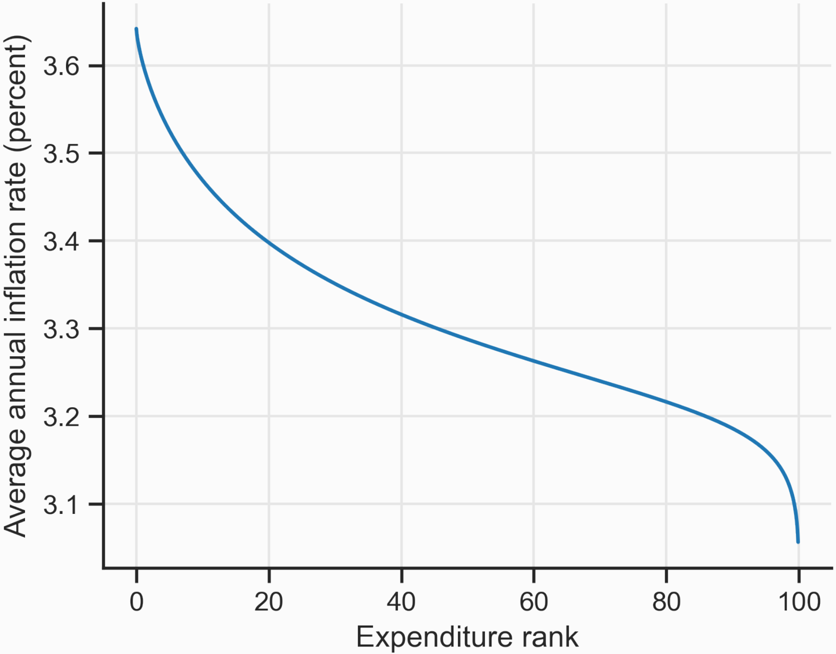
14 pct. lower
Poorest

8 pct. higher
Richest

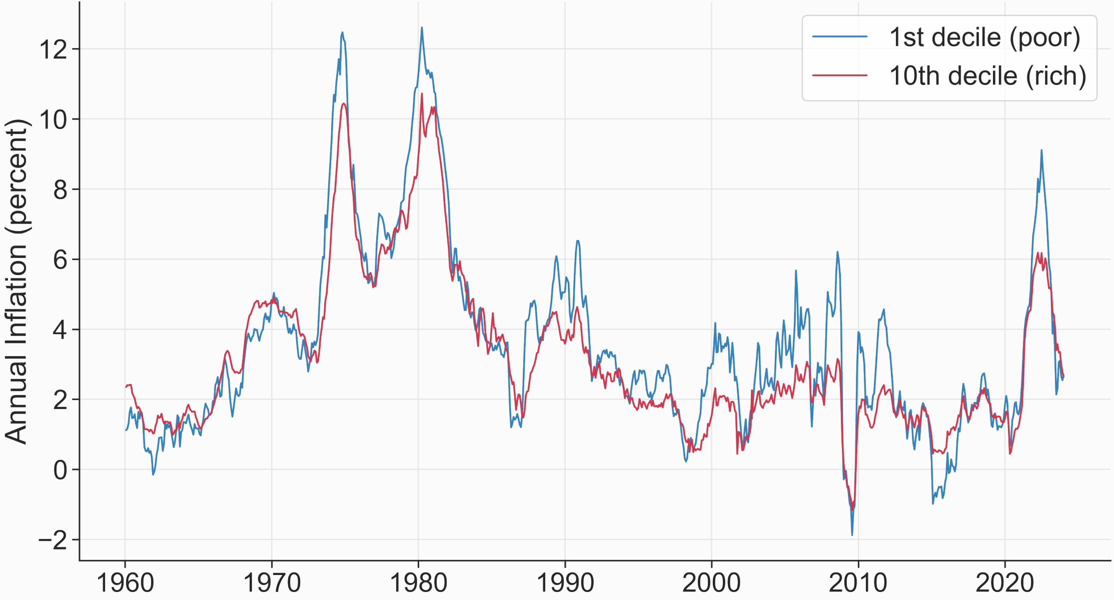
Long-run inflation inequality



Full distribution of long-run avg. annual inflation rates



Inflation Dynamics the last 65 years



Parsing inflation inequality

What matters for inflation inequality?

Empirical finding: long-run annual inflation rate gap of 0.39 percentage points

- Excluding durable goods lowers long-run inflation inequality to 0.17 percentage points

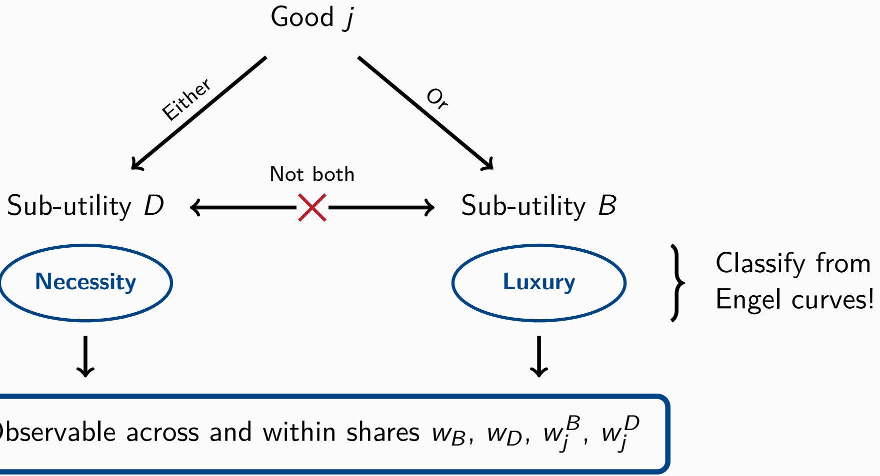
  $\Longrightarrow$ **Full** consumption basket matters
- Coarser product group aggregation (71 vs 15) lowers long-run inflation inequality to range between 0.15 to 0.20 percentage points
  - Consistent with Jaravel (2019, 2021)

  $\Longrightarrow$ **Broad** data is necessary

Thank you

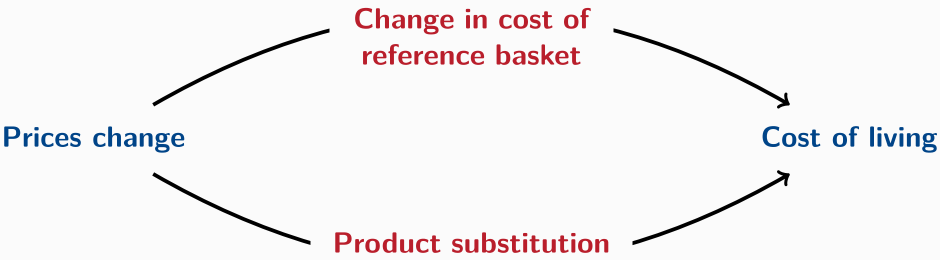
Appendix

Separability in a nutshell: Two-good case



Note: Quasi-separability groups *prices* of goods in the *expenditure function*, in contrast to direct separability which groups *quantities* of goods in the *utility function* (Gorman, 1996). ▸ Back

The role of product substitution



Reference baskets and **substitution behaviour** differ across the expenditure distribution!

Back to main presentation
Formal decomposition

A Nonhomothetic Törnqvist index

Proposition

Let $B(\boldsymbol{p})$, $D(\boldsymbol{p})$, and $H(\boldsymbol{p})$ be homogeneous translog expenditure functions. If $\varepsilon \to 0$ and $\sigma \to 1$, then the PIGL cost-of-living index becomes the standard Törnqvist index:

$$\frac{P(u, \boldsymbol{p}_t, \boldsymbol{p}_s)}{P(u, \boldsymbol{p}_{t-1}, \boldsymbol{p}_s)} = \prod_{j \in J} \left(\frac{p_{jt}}{p_{jt-1}} \right)^{\delta_{j,t,t-1}}, \qquad \delta_{j,t,t-1} = \frac{w_{jt} + w_{jt-1}}{2},$$

where $J = J_D \cup J_B \cup J_H$ is the full set of commodities available and $w_j = p_j q_j \,/\, e$ is the total expenditure share of commodity j.

▸ back

Classification: estimation

Classification from Engel curve slopes

(Wachter and Yogo, 2010; Orchard, 2022; Hochmuth, Pettersson and Weissert, 2023)

- Necessity if slope is negative; Luxury if positive; Homothetic if statistically insignificant

Regression

$$\overline{w}_{jgt} = \alpha_{jr} + \alpha_{jt} + \beta_{je} \ln \overline{e}_{gt} + \beta_{jp} \ln RPP_{jgt} + u_{jgt}. \quad (1)$$

- α_{jr} is region dummy: controls for permanent differences in consumption patterns across regions unrelated to nonhomotheticity
- α_{jt} is time fixed effect: controls for aggregate changes in relative prices between goods and for any other common macro shocks
- RPP_{jgt} price parity adjustment: controls for differences in relative prices across states and their evolution over time.

Product j, state g, year t, state-level aggregate expenditure share $\overline{w}_{jgt}$, per-capita consumption expenditure $\overline{e}_{gt}$ back

Classification: results

	(1)	(2)	(3)
New motor vehicles, 2.1 %	D	D	B
Used motor vehicles, 1.2 %	D	D	D
Motor vehicle parts/accessories, 0.6 %	D	D	D
Furniture and furnishings, 1.5 %	B	B	B
Household appliances, 0.5 %	D	D	D
Glassware, tableware & utensils, 0.3 %	B	B	B
Tools for house and garden, 0.3 %	D	D	D
Video/audio/photo equipment, 1.9 %	B	B	B
Sporting equipment and guns, 0.6 %	D	D	B
Sports and recreational vehicles, 0.5 %	H	H	H
Recreational books, 0.2 %	H	H	D
Musical instruments, 0.0 %	B	B	B
Jewelry and watches, 0.6 %	B	B	B
Therapeutic appliances/equipment, 0.5 %	D	D	D
Educational books, 0.1 %	B	B	B
Luggage and similar items, 0.2 %	B	B	B
Telephone and related equipment, 0.2 %	B	B	B
Food and nonalcoholic beverages, 6.9 %	D	D	D
Alcoholic beverages, 1.2 %	B	B	B
Food produced/consumed on farms, 0.0 %	D	D	D
Women's and girls' clothing, 1.4 %	H	H	H
Men's and boys' clothing, 0.8 %	D	D	D
Children's and infants' clothing, 0.2 %	D	D	D
Other clothing and footwear, 0.7 %	D	D	D
Gasoline, 2.8 %	D	D	D
Fuel oil and other fuels, 0.2 %	D	D	H
Pharmaceutical products, 3.4 %	D	D	D
Recreational items, 1.4 %	D	D	D
Household supplies, 1.1 %	D	D	D
Personal care products, 1.1 %	B	B	H
Tobacco, 0.9 %	D	D	D
Magazines/newspapers, stationery, 0.5 %	B	B	H
Rental nonfarm dwellings, 3.7 %	D	B	D
Owned nonfarm dwellings, 12.2 %	H	B	H
Rental value of farm dwellings, 0.2 %	B	D	D
Group housing, 0.0 %	B	H	B

	(1)	(2)	(3)
Water supply and sanitation, 0.8 %	D	D	D
Electricity, 1.5 %	D	D	D
Natural gas, 0.5 %	D	D	D
Physician services, 4.1 %	H	D	H
Dental services, 1.0 %	D	D	B
Paramedical services, 2.7 %	B	B	H
Hospitals, 7.8 %	D	D	D
Nursing homes, 1.4 %	D	D	D
Motor vehicle maintenance/repair, 1.4 %	H	H	B
Other motor vehicle services, 0.7 %	B	B	H
Ground transportation, 0.4 %	B	B	D
Air transportation, 0.8 %	B	B	B
Water transportation, 0.0 %	D	H	D
Membership clubs, 1.5 %	B	B	B
Video/audio/photo equip services, 1.0 %	B	B	B
Gambling, 1.1 %	B	B	B
Other recreational services, 0.5 %	B	B	B
Purchased meals and beverages, 5.6 %	B	B	D
Food to employees inc military, 0.2 %	D	D	D
Hotels, 1.0 %	B	B	B
Financial services no payment, 2.4 %	B	B	B
Financial services charges/fees, 2.6 %	B	B	B
Life insurance, 0.7 %	B	B	B
Household insurance, 0.1 %	B	B	H
Health insurance, 1.5 %	B	B	B
Motor vehicle insurance, 0.6 %	B	B	B
Telecommunication services, 1.3 %	B	B	B
Postal and delivery services, 0.1 %	D	H	H
Higher education, 1.4 %	D	D	H
Elementary and secondary schools, 0.3 %	H	H	H
Commercial/vocational schools, 0.4 %	B	B	B
Professional and other services, 1.5 %	D	D	D
Personal care/clothing services, 1.1 %	B	B	B
Social services and religion, 1.5 %	B	B	B
Household maintenance, 0.6 %	B	B	B

Specifications

Column (1): baseline

Column (2): w/o RPP, 1997–2022

Column (3): Controlling for age

Baseline results

30 necessities

34 luxuries

7 homothetic goods

$\Longrightarrow$ consistent with e.g. Wachter and Yogo (2010), Orchard (2022) and Hochmuth *et. al* (2023)

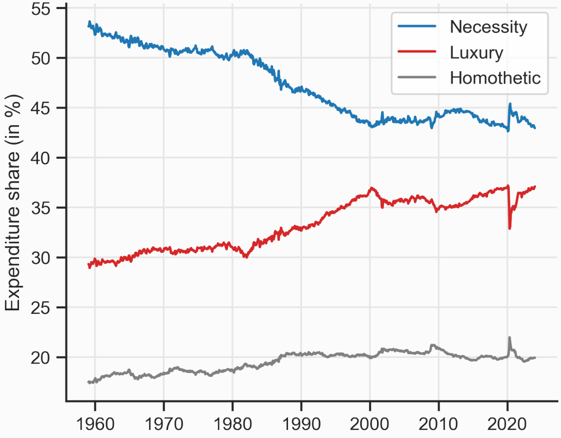
Goods are broadly necessities

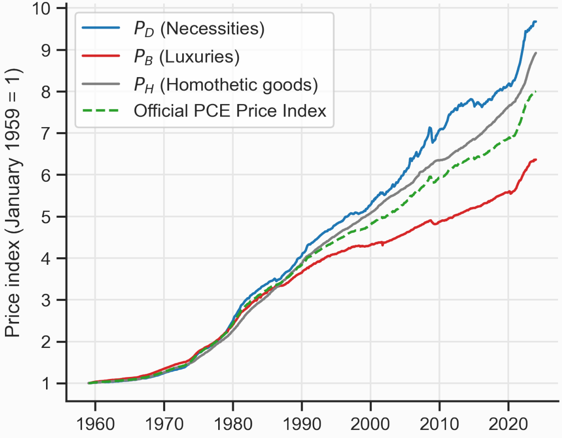
Services are broadly luxuries

$\Longrightarrow$ consistent with macro evidence on structural change

▸ back

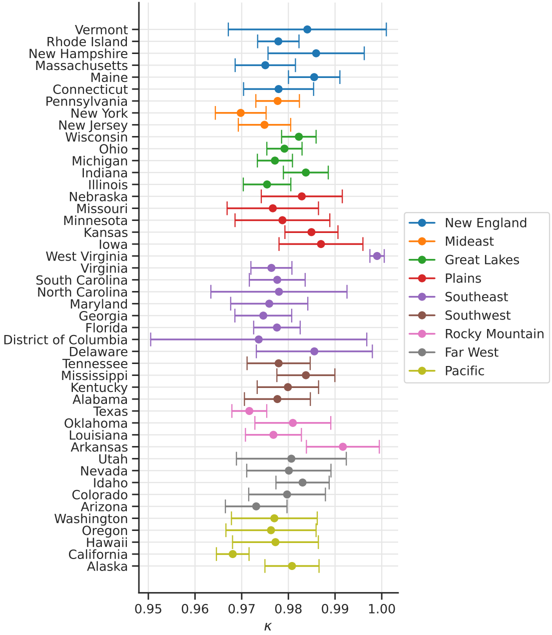
Expenditure shares and basket prices





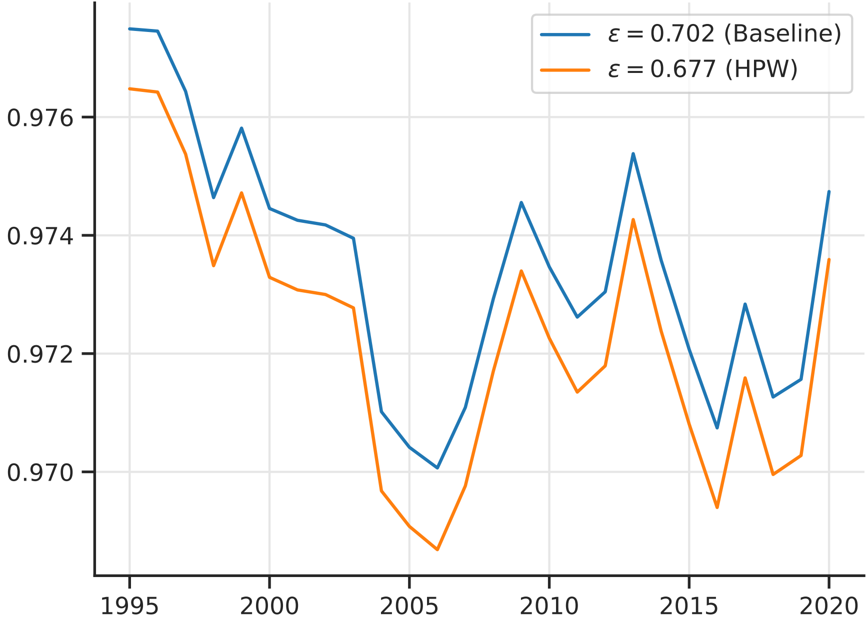
▸ back

AGGREGATION FACTOR κ BY U.S. STATES



back

Aggregation factor κ over time



back

Estimation of ε using aggregate state-level data

Taking logs of aggregate necessity expenditure share equation yields linear fixed-effects regression

$$\ln \overline{w}_{Dgt} = \alpha_r + \alpha_t + \varepsilon \ln \left[\frac{P_{Fgt}}{\overline{e}_{gt}} \cdot \frac{P_{Dgt}}{P_{Bgt}} \right] + u_{gt},$$

Identification of ε is obtained from U.S. cross-state variation

- Aggregation is also consistent within states $\Longrightarrow$ no aggregation bias
- Compute state and category-specific prices by adjusting subcategory price indices with RPPs
- Apply PIGL formulas using state-level expenditure shares
- α_r captures region fixed effects
- α_t captures time fixed effects
- g denotes state

▸ κ across states ▸ κ across time ▸ Estimation results

▸ back

Estimated preference parameters from US state-level data

	(1)	(2)	(3)	(4)
ε	0.702*** (0.061)	0.726*** (0.062)	0.791*** (0.018)	0.512*** (0.038)
Durable goods[a]	✓		✓	✓
RPP controls	✓	✓		✓
Age controls				✓
Observations	765	765	1,326	765
RMSE	0.053	0.053	0.038	0.039
Adjusted R^2	0.290	0.286	0.708	0.536

Notes. RMSE denotes the root mean square error. Robust standard errors in parentheses. *, **, and *** denote statistical significance at the 5 percent, 1 percent, and 0.1 percent levels. Columns (3) and (4) use the classification without RPPs and with age controls, respectively.

[a] Motor vehicles and parts, furnishings and durable household equipment, recreational goods and vehicles, and other durable goods.

▸ back

Full distribution of inflation rates

PIGL cost-of-living index:

$$P(u, \boldsymbol{p}_t, \boldsymbol{p}_s) = \left[(1 - w_{Ds}) P_{Bt}^{\varepsilon} + w_{Ds} P_{Dt}^{\varepsilon}\right]^{\frac{1}{\varepsilon}} \left(\frac{P_{Ht}}{P_{Bt}}\right)^{\rho_{t,s}}$$

- Basket prices, the homothetic expenditure share and the estimate of ε is sufficient to compute the PIGL cost-of-living index *for some* base-period necessity expenditure share w_{Ds}.
- We already have *one* interesting candidate: the aggregate/representative w_{Ds}
- We can also study hypothetical individuals such as 'a person with 50 pct. of the average´

What about the actual distribution?

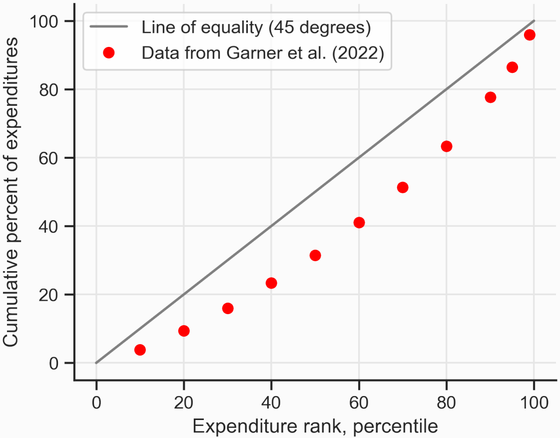
Full distribution of inflation rates: Lorenz curve

Garner *et al.* (2022) distribute 2019 aggregate PCE spending across U.S. households

- Point estimates: deciles, top 5 and 1 pct.

Model offers direct link between

i) aggregate expenditure share,
ii) overall distribution and,
iii) household-level necessity exp. shares.



$\ell(x)$ is the Lorenz curve, x is expenditure rank. Evaluated at x_h, it holds that $\ell'(x_h) = e_h/\overline{e}$.

Individual and aggregate necessity expenditure shares then imply

$$w_{Dh} = \frac{w_{Dh}}{\overline{w}_D}\overline{w}_D = \left(\frac{e_h}{\overline{e}\kappa}\right)^{-\varepsilon}\overline{w}_D = \left(\frac{\ell'(x_h)}{\kappa}\right)^{-\varepsilon}\overline{w}_D.$$

▸ Lorenz aggregation factor κ

Full distribution of inflation rates: Lorenz curve

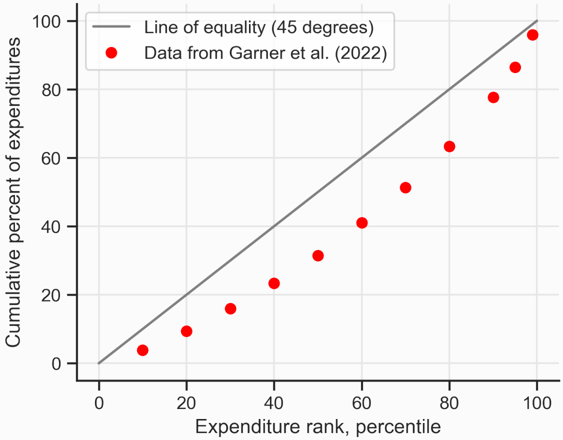
$$w_{Dh} = \left(\frac{e_h}{\overline{e}\kappa}\right)^{-\varepsilon} \overline{w}_D = \left(\frac{\ell'(x_h)}{\kappa}\right)^{-\varepsilon} \overline{w}_D.$$

Sufficient data input

- Lorenz curve, $\ell(x)$,
- empirically observed aggregate expenditure share, $\overline{w}_D$,
- preference parameter, ε.

How to get $\ell(x)$?

Use Garner *et al.* (2022) point estimates



Back to main data frame

Full distribution of inflation rates: Lorenz curve

$$w_{Dh} = \left(\frac{e_h}{\overline{e}\kappa}\right)^{-\varepsilon} \overline{w}_D = \left(\frac{\ell'(x_h)}{\kappa}\right)^{-\varepsilon} \overline{w}_D.$$

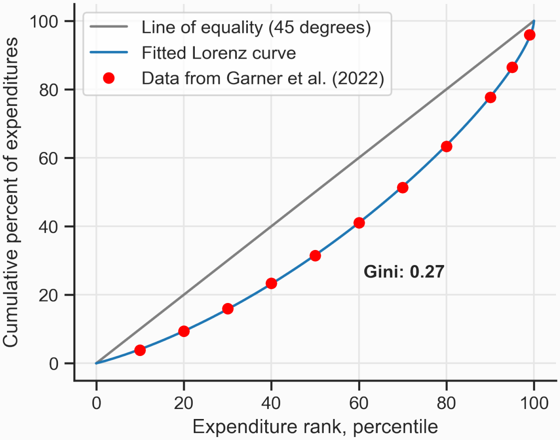
Sufficient data input

- Lorenz curve, $\ell(x)$,
- empirically observed aggregate expenditure share, $\overline{w}_D$,
- preference parameter, ε.

How to get $\ell(x)$?

Use Garner *et al.* (2022) point estimates and parameterize $\ell(x)$ following Sitthiyot and Holasut (2021) Parameterization of $\ell(x)$

Back to main data frame



PARAMETERIZATION OF $\ell(x)$

Sitthiyot and Holasut (2021) propose to parametrize $\ell(x)$ as a weighted average between an exponential function and the functional form implied by a Pareto distribution:

$$(1-\omega)x^{\eta}+\omega(1-(1-x)^{1-\eta}),$$

where ω and η are parameters to estimate.

Fitting this function to the distributional PCE data by Garner *et al.* (2022) yields an R^2 of 0.9999.

▸ back ▸ Back to main data frame