

Is consumption getting more unequal?

Josh Clyne

e61 Institute

September 30, 2026

Introduction

- ▶ We should care about the **distribution of consumption** - standard economic models tell us that utility comes from consumption and leisure
- ▶ But measuring **long-run trends** in consumption inequality (usually) requires survey data, and one concern with survey-reported expenditure is that it contains measurement error from **misreporting**.
- ▶ Different treatments of this measurement error (often using the same survey data) has led to starkly different conclusions in the US

Today:

- ▶ **Is survey misreporting biasing our understanding of consumption inequality in Australia?**

US studies tend to disagree a lot about consumption inequality...



Figure: Cross-Sectional Consumption Inequality in the US from various studies. All series present 3 year moving averages and are normalised to Heathcote, Perri & Violante (2010) in 1980. Sources: Meyer and Sullivan (2023) and Attanasio and Pistaferri (2016)

Today: Is survey misreporting biasing our understanding of consumption inequality in Australia?

I look at two dimensions of misreporting:

1. **Aggregate misreporting**

- ▶ US literature documents **growing levels of under-reporting** in survey-measured expenditure for some items, using aggregate validation data from National Accounts
- ▶ I replicate Meyer and Sullivan (2023), using 'well measured' goods that better track national accounts to make inferences about trends in total consumption inequality

2. **Within household misreporting:**

- ▶ Unique feature of HILDA data - we see multiple reports of the same household-level expenditure
- ▶ How does this information change our understanding of 'well-measured' expenditure items?
- ▶ What are the implications for the evolution of consumption inequality?

HILDA Data

- ▶ Multiple household members independently report expenditure on the **same** household total
- ▶ Includes 16 **non-durable** expenditure categories from the Self-Complete Questionnaire (SCQ) in years 2006-2024
- ▶ Reports are recall-based, with frequency varying by good
- ▶ I concord majority of HILDA items to National Accounts COICOPs: using ABS HES to NA concordance (omitting education, medicines, and health insurance)

C5 For each type of expenditure below, write in your best estimate of the total amount spent on that item by all people in the household.

If you are unsure please make your best guess.

Do not include expenses associated with any businesses you may own.

Weekly Expenses		Any expenditure?		HOW MUCH PER WEEK? (on average)	
		NO	YES		
a	Groceries <i>(Include food, cleaning products, pet food and personal care products. Do <u>not</u> include alcohol or tobacco.)</i>	☐	☐	➔ \$	

Aggregate Misreporting: Replication of Meyer and Sullivan (2023) in HILDA

Target: Identifying changes in inequality

Reported expenditure by household h on good j in year t :

$$\ln x_{h,j,t} = \underbrace{\ln x_{h,j,t}^*}_{\text{true consumption}} + \underbrace{\psi_t^j}_{\substack{\text{systematic error} \\ \text{(common across HH)}}} + \underbrace{\nu_{h,j,t}}_{\text{idiosyncratic reporting error}} \quad (1)$$

To identify changes in true consumption, Meyer and Sullivan (2023) emphasise that the whole inequality literature must assume that the distribution of $\nu_{h,j,t}$ is:

- ▶ **Time invariant** (Assumption A), and
- ▶ **Independent of true consumption** (Assumption B)

Where Assumptions A and B enter

From well-measured items to total consumption

Idea from Meyer and Sullivan (2023): if we have a bundle of goods we believe is 'well-measured', we can make inferences about trends in *true total consumption* under two testable conditions.

1. **Unit elasticity:** well-measured consumption is a constant share of true total consumption

$$x_{h,t}^*$$

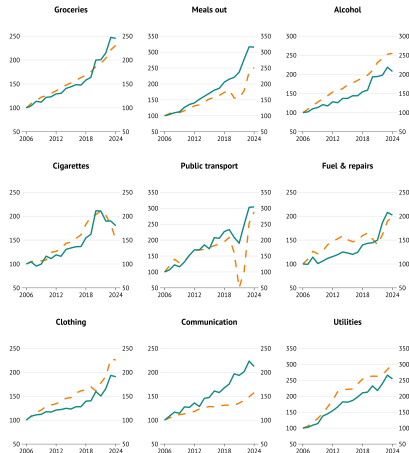
$$\ln x_{h,w,t}^* = \alpha + \ln x_{h,t}^* + \varepsilon_{h,w,t} \quad (3)$$

2. **Stable relative prices:** the price of the well-measured bundle does not move relative to the price of the whole consumption basket

Comparing HILDA expenditure categories to National Accounts

HILDA vs. National Accounts, indexed growth

Index, 2006 = 100, well-concorded categories



* Uses cross-sectional population weights (nhwt).

** HILDA wave w is matched to the financial year starting in calendar year w + 1999.

*** Solid teal = HILDA; dashed orange = HFCE (National Accounts).

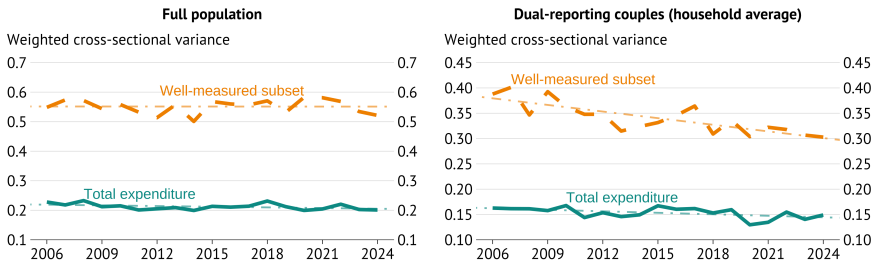
Sources: e61; HILDA Restricted Release 24

Does restricting to well-measured goods change the picture?

- ▶ Compare the correlation of year-on-year changes between HILDA and National Accounts
- ▶ Select categories whose correlation is statistically significant across the 18 years
- ▶ Categories selected: meals out, public transport, and vehicle fuel + repairs (approx. 21% of HILDA non-durable expenditure).

Well-measured expenditure as a guide to the total

Weighted cross-sectional variance of log equivalised value, by sample



* Uses the OECD-modified equivalence scale and cross-sectional population weights (hhwth).

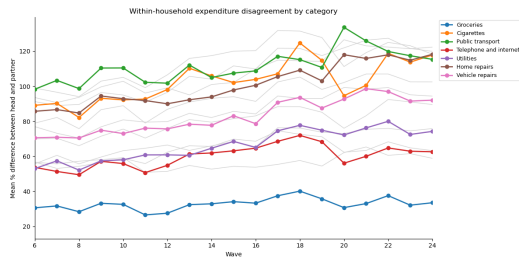
** Dot-dash lines are OLS trends over year.

*** 'Well-measured subset' uses only the categories identified as well-measured in Figure 5.

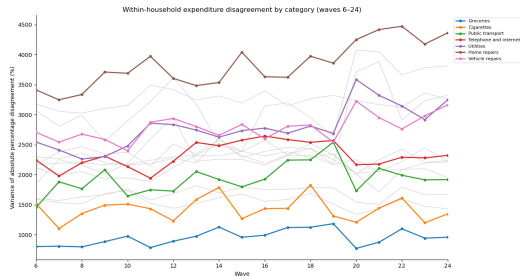
Sources: e61; HILDA Restricted Release 24

Within-household disagreements

Descriptive Evidence: We see the level and variance of intra-household disagreements change over time



(a) Mean Percentage Disagreement



(b) Variance Percentage Disagreement

Figure: The first two moments of within-household disagreements over time

$$\text{Percent Difference}_{h,j,t} = \frac{|x_{h,1,j,t} - x_{h,2,j,t}|}{(x_{h,1,j,t} + x_{h,2,j,t})/2}$$

Adding multiple reporters to the measurement error model

Reported expenditure by person $k \in \{1, 2\}$ of household h on good j in year t :

$$\ln x_{h,k,j,t} = \underbrace{\ln x_{h,j,t}^*}_{\text{true consumption}} + \underbrace{\psi_t^j}_{\substack{\text{systematic error} \\ \text{(common across HH)}}} + \underbrace{\nu_{h,k,j,t}}_{\substack{\text{idiosyncratic} \\ \text{reporting error}}}$$

Assumption A: The distribution of $\nu_{h,k,j,t}$ is constant over time

- Consider the difference between two household members' reports:

$$d_{h,j,t} = \ln x_{h,1,j,t} - \ln x_{h,2,j,t} = \nu_{h,1,j,t} - \nu_{h,2,j,t}$$

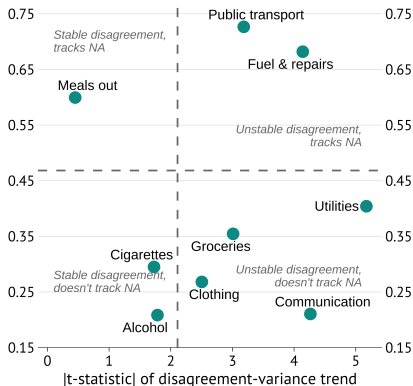
- We can test if $V(\nu_{h,1,j,t} - \nu_{h,2,j,t})$ is stable over time
- The test is rejected for all but three categories: meals out, cigarettes and alcohol

Test and results

Is disagreement instability linked to poor National Accounts tracking?

Disagreement-variance instability vs. National Accounts tracking

Correlation of HILDA/HFCE wave-on-wave % change



* Uses cross-sectional population weights (hhwth).

** Dashed lines: critical values for $p < 0.05$ (vertical: trend t-test, $n=19$; horizontal: correlation test, $n=18$).

Sources: e61; HILDA Restricted Release 24

Towards a noise-reduced measure of consumption inequality

Rather than a single partner's report, use the **average of the two partners' reports**. Each report is truth plus idiosyncratic noise:

$$\ln x_{h,1,j,t} = \ln x_{h,j,t}^* + \psi_t^j + \nu_{h,1,j,t}, \quad \ln x_{h,2,j,t} = \ln x_{h,j,t}^* + \psi_t^j + \nu_{h,2,j,t}$$

Averaging keeps true consumption and the systematic error unchanged:

$$\ln \bar{x}_{h,j,t} = \ln x_{h,j,t}^* + \psi_t^j + \frac{1}{2}(\nu_{h,1,j,t} + \nu_{h,2,j,t})$$

Compare the **average of the two individual reports' variances** to the **variance of the average report**:

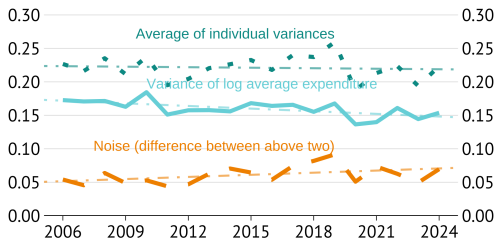
$$\underbrace{\frac{V(\ln x_{h,1,j,t}) + V(\ln x_{h,2,j,t})}{2}}_{\text{avg. of individual variances}} \quad \text{vs.} \quad \underbrace{V(\ln \bar{x}_{h,j,t})}_{\text{variance of the average}}$$

- ▶ **Large gap:** much of the spread across households in a single report is noise, not real differences in $x_{h,j,t}^*$
- ▶ **Gap moving over time:** the noise itself is changing, suggests that a single-reporter trend would mistake this for a change in true inequality

The noise-reduced measure vs. individual reports

Within-household expenditure disagreement

Weighted cross-sectional variance of log equivalised expenditure, by year
Weighted cross-sectional variance



* Uses the OECD-modified equivalence scale and cross-sectional population weights (hhwth).

** Dot-dash lines are OLS trends over year.

Sources: e61; HILDA Restricted Release 24

Conclusions

- ▶ Despite misreporting appearing to be a problem in two sources of validation data, methods to 'correct' for this error don't change the story much for consumption inequality in HILDA...
 - ▶ We see consumption inequality decreasing slightly more for the dual-reporting households subsample
- ▶ The influence of other misreporting error remains unknown (e.g. $\text{Cov}(\ln x_{h,j,t}^*, \nu_{h,k,j,t})$)

Future Directions:

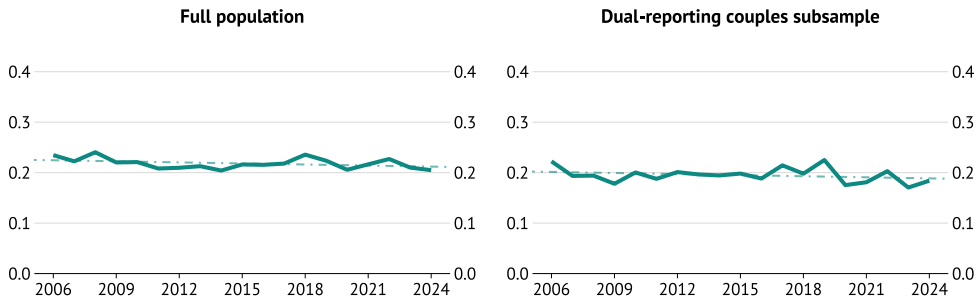
- ▶ Validate with expenditure dispersion in other data-sets (e.g. bank transactions)
- ▶ Add other forms of consumption to the mix (e.g. durables)

Appendix

What does the raw data tell us?

Consumption inequality since 2006

Weighted cross-sectional variance of log equivalised expenditure, by year



* Uses the OECD-modified equivalence scale and cross-sectional population weights (hhwth).

** Dot-dash line is the OLS trend over year.

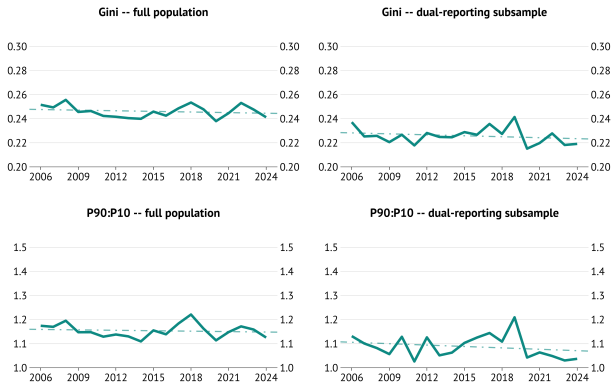
*** Dual-reporting subsample: couple households where both partners separately report every expenditure category (see Figure 2); shows the head's own reported total.

Sources: e61; HILDA Restricted Release 24

Consumption, income and wage inequality: Gini and P90:P10

Consumption inequality: Gini and P90:P10

Full population vs. dual-reporting subsample, by year



* Uses the OECD-modified equivalence scale and cross-sectional population weights (hwhth).

** Dot-dash line is the OLS trend over year.

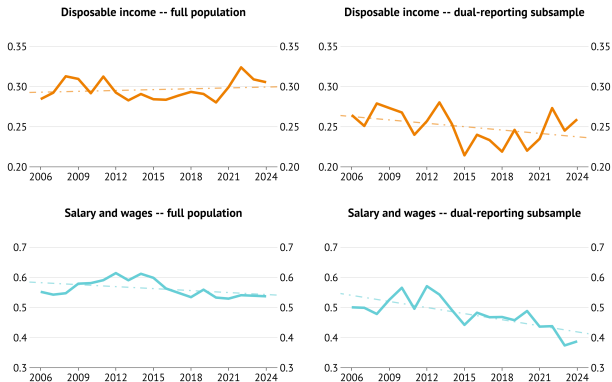
*** Dual-reporting subsample: couple households where both partners separately report every expenditure category (see Figure 2); shows the head's own reported total.

Sources: e61; HILDA Restricted Release 24

Consumption, income and wage inequality: Gini and P90:P10

Income and wage inequality

Weighted cross-sectional variance of log equivalised value, full population vs. dual-reporting subsample



* Uses the OECD-modified equivalence scale and cross-sectional population weights (hwhth).

** Dot-dash line is the OLS trend over year.

*** Dual-reporting subsample: couple households where both partners separately report every expenditure category (see Figure 2); shows the household's own reported value.

Sources: e61; HILDA Restricted Release 24

Note that dual-reporting households are a subset

Table: Demographic characteristics by sample and reporting status

	Full sample	Couples	
		Both report	One reports
N (household-years)	95,782	28,043	20,971
Avg. observations per year	5,041	1,476	1,104
Age (years)	49.6	46.6	48.6
University educated	30.4%	34.3%	36.8%
Household income (2025\$, equiv.)	55,334	65,070	66,081
Employed	65.0%	72.3%	70.7%

Note: Full sample includes all household types, not just couples. Income is equivalised using the OECD-modified scale. Demographic characteristics are those of the household head. Unweighted sample means/shares.

- Sample includes individuals with > 5k \$2025 total non-durable spending report, and trims top 1%

Meyer and Sullivan: where Assumptions A and B enter

Substituting the elasticity condition (3) into the error model (1) and exponentiating:

$$x_{h,w,t} = x_{h,t}^* \cdot \underbrace{e^{\alpha + \psi_{w,t}}}_{\text{common across } h} \cdot \underbrace{e^{\lambda_{h,w,t}}}_{\text{idiosyncratic}}, \quad \lambda_{h,w,t} \equiv \varepsilon_{h,w,t} + \nu_{h,w,t}$$

Identifying changes in inequality: variance of the log. Taking the variance of $\ln x_{h,w,t}$ across households:

$$V(\ln x_{h,w,t}) = V(\ln x_{h,t}^*) + V(\lambda_{h,w,t}) + 2 \text{Cov}(\ln x_{h,t}^*, \lambda_{h,w,t})$$

Recovering *changes* in true inequality requires two things, now applied to $\lambda_{h,w,t}$, to be stable over time:

Assumption A (time-invariance):

$$V(\lambda_{h,w,t}) - V(\lambda_{h,w,t'}) = 0$$

Assumption B (weaker than independence):

$$\text{Cov}(\ln x_{h,t}^*, \lambda_{h,w,t}) - \text{Cov}(\ln x_{h,t'}^*, \lambda_{h,w,t'}) = 0$$

Testing Assumption A: F-tests on wave dummies

For each category j , estimate:

$$d_{h,j,t}^2 = \alpha_j + \sum_{t \in \mathcal{T} \setminus \{t_0\}} \gamma_{jt} \mathbf{1}\{t\} + \varepsilon_{h,j,t}$$

whether the wave dummies are jointly zero:

$$H_0^{\text{var}} : \gamma_{jt} = 0 \quad \forall t$$

Rejection of either the mean or variance null is sufficient to conclude Assumption A fails for category j .

Back

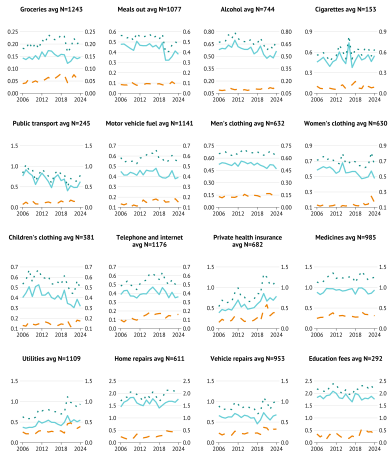
Assumption A: rejected for total expenditure and most categories

Category	F_{var}	p	Result
Total expenditure	2.90	.000	Rejected
Groceries	3.17	.000	Rejected
Meals out	2.03	.006	Rejected
Alcohol	1.37	.135	Not rejected
Cigarettes	1.29	.186	Not rejected
Public transport	2.35	.001	Rejected
Motor vehicle fuel	3.95	.000	Rejected
Men's clothing	1.39	.125	Not rejected
Women's clothing	2.54	.000	Rejected
Children's clothing	1.06	.392	Not rejected
Telephone & internet	4.69	.000	Rejected
Private health ins.	14.30	.000	Rejected
Medicines	3.20	.000	Rejected
Utilities	10.18	.000	Rejected
Home repairs	9.43	.000	Rejected
Vehicle repairs	7.29	.000	Rejected
Education fees	5.38	.000	Rejected

The noise-reduced measure, by category

Within-household expenditure disagreement, by category

Weighted cross-sectional variance of log equivalised expenditure, by year



* Uses the OECD-modified equivalence scale and cross-sectional population weights (hwhth).
** Solid = variance of average; dotted = average of individual variances; dashed = noise.
Sources: eSL; HILDA Restricted Release 24

HILDA vs. National Accounts: year-on-year changes

HILDA vs. National Accounts, wave-on-wave % change

Per cent, well-concorded categories (correlation shown per panel)



* Uses cross-sectional population weights (h1wh1).

** HILDA wave w is matched to the financial year starting in calendar year w + 1999.

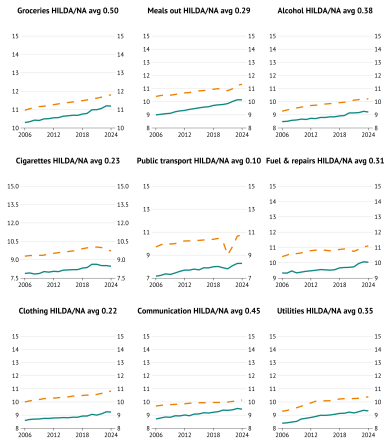
*** Solid teal = HILDA; dashed orange = HFCCE (National Accounts).

Sources: e61; HILDA Restricted Release 24

HILDA vs. National Accounts: log levels

HILDA vs. National Accounts, log levels

log(\$M), well-concorded categories



* Uses cross-sectional population weights (hwhth).

** HILDA wave w is matched to the financial year starting in calendar year $w + 1999$.

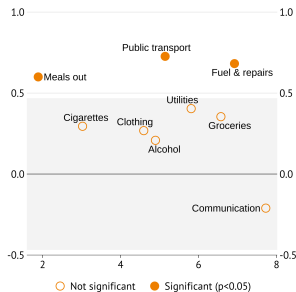
*** Solid teal = HILDA; dashed orange = MFCE (National Accounts).

Sources: e61; HILDA Restricted Release 24

Mean-disagreement instability vs. National Accounts tracking

Mean-disagreement instability vs. National Accounts tracking

|t-statistic| of mean-disagreement trend vs. rho of
HILDA/HFCE wave-on-wave % change
rho: correlation of HILDA/HFCE wave-on-wave %
change



* Uses cross-sectional population weights (hwhwt).

** Shaded band: rho statistically indistinguishable from zero
($p > 0.05$, $n = 18$).

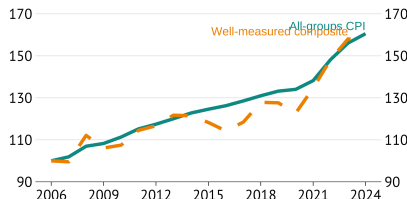
Sources: e61; HILDA Restricted Release 24

Selecting the 'well-measured' bundle and checking two testable assumptions

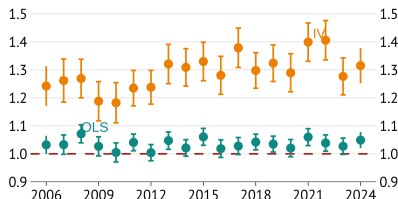
The testable assumptions

Well-measured composite vs. all-groups CPI price level, and OLS/IV expenditure elasticity, by year

Price index (2006 = 100)



Elasticity of well-measured w.r.t. total expenditure



* Well-measured composite is a fixed-weight index using 2006 (population-weighted) HILDA expenditure shares across the categories identified as well-measured in Figure 5.

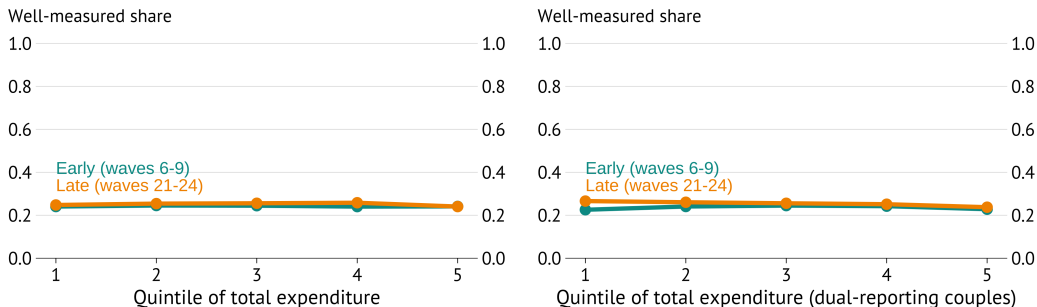
** IV instrument: log household disposable income.

Sources: e61; HILDA Restricted Release 24

Validity check: is the well-measured share constant across the distribution?

Is the well-measured share constant across the distribution?

Well-measured share of total expenditure, by quintile and period



* Uses the OECD-modified equivalence scale and cross-sectional population weights (hhwth).

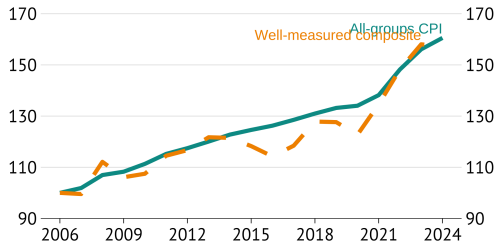
Sources: e61; HILDA Restricted Release 24

Validity check: deflation choice and expenditure elasticity

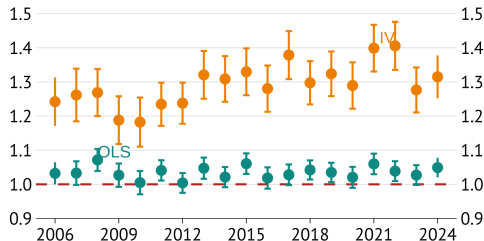
The testable assumptions

Well-measured composite vs. all-groups CPI price level, and OLS/IV expenditure elasticity, by year

Price index (2006 = 100)



Elasticity of well-measured w.r.t. total expenditure



* Well-measured composite is a fixed-weight index using 2006 (population-weighted) HILDA expenditure shares across the categories identified as well-measured in Figure 5.

** IV instrument: log household disposable income.

Sources: e61; HILDA Restricted Release 24

Data Disclaimer

This paper uses unit record data from Household, Income and Labour Dynamics in Australia Survey [HILDA] conducted by the Australian Government Department of Social Services (DSS). The findings and views reported in this paper, however, are those of the authors and should not be attributed to the Australian Government, DSS, or any of DSS' contractors or partners. DOI: 10.26193/GMYVPW.