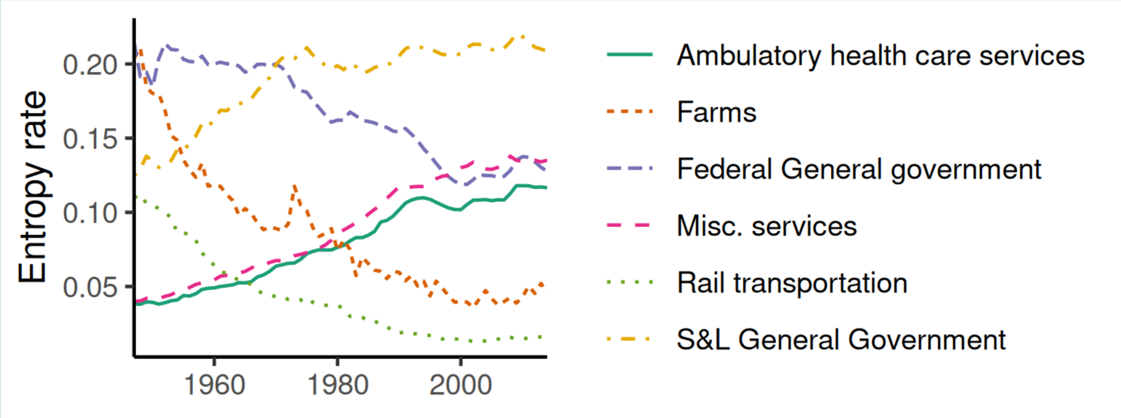
Accounting for variety

Julian Winkler
julian.winkler@economics.ox.ac.uk

DEPARTMENT OF ECONOMICS · UNIVERSITY OF OXFORD · Institute for New Economic Thinking AT THE OXFORD MARTIN SCHOOL

Problem

The industry composition of aggregate output (GDP) changes over time. These changes create a bias in economic indices tasked with tracking aggregate quantity or price. An extrapolated real output index is *more* biased *upwards* if variety in GDP increases. Variety, measured by **entropy**, is quantitatively important.



Basic Concept

Consider a scenario in which nominal GDP Y is accurately measured, in addition to a set of industry prices P_i. Prices are aggregated into one index $\widetilde{P}$ using weights ω_i. Subtracting the price index from the growth rate of GDP would yield a biased, **extrapolated** quantity index $\widetilde{X}$. The key contribution of the present paper is to give this bias a name:

$$\left.\Delta \log Y_t - \Delta \sum_{i=1}^{N} \omega_{i,t} \log P_{i,t} = \right\} \text{Extrapolated: } \widetilde{X}_t$$

$$\underbrace{\Delta \sum_{i=1}^{N} \omega_{i,t} \log Q_{i,t}}_{\text{Real output growth: } \widetilde{Q}_t} + \underbrace{\Delta \sum_{i=1}^{N} \omega_{i,t} \log \left(\frac{Y_t}{Y_{i,t}}\right)}_{\text{Entropy change: } \Delta H(\boldsymbol{\Omega}_t, \mathbf{Y}_t)} .$$

If a change in entropy (variety) is not accounted for, then the extrapolated real output index is **not** equal to index derived from industry-specific quantities.

Motivation

A model producing weights $\boldsymbol{\Omega}_t$ comes at the price of some information loss, to the extent that weights does not match true expenditures.
The Divisia index $\boldsymbol{\Omega}_{D,t} = \mathbf{Y}_t$ fixes the bias to a quantity defined as Shannon entropy: $H(\mathbf{Y}_t)$ [1]. Interpretations of (cross-)entropy include:

- **Log-likelihood**: the information content of observed consumption patterns under our choice of weighting scheme,
- **Dispersion**: the more uniform $Y_{i,t}$, the higher $H(\mathbf{Y}_t)$,
- **Uniform types**: $N^* = \exp[H(\mathbf{Y}_t)]$.

Even though the Divisia index traces along the expenditure path, a preference for variety (increasing entropy) will generate a bias.

References

[1] Shannon, Claude E. (1948), 'A Mathematical Theory of Communication', *The Bell System Technical Journal* **27**, pp. 379–423.

[2] Hulten, C. R. (1997), Comment on "Do Real–Output and Real–Wage Measures Capture Reality? The History of Lighting Suggests Not", in Bresnahan & Gordon (1997), pp. 66—70.

Acknowledgements

This research was supported by Baillie Gifford and the Institute for New Economic Thinking at the Oxford Martin School.

Hulten's Paradox

One illustrative thought experiment demonstrating the importance of a changing output distribution at any level of aggregation is Hulten's Paradox. The ICT revolution in the 1990's sparked a discussion on measuring welfare improvements from the changing quality of products, in addition to their real output. Besides practical difficulties in implementing such 'hedonic' price indices, Charles Hulten famously cautions the use of quality adjustments since they likely overstate actual welfare gains [2]. A passage provides the key insight:

> [...]a person possessing the average disposable income in America today should be willing to accept a massive reduction in spending power – from \$17,200 to the \$90-430 range – in order to avoid being sent back in time to an equivalent status in colonial America. Alternatively, it suggests that the average colonial should prefer living in the America of today, with as little as \$90 per year, to staying put in the late eighteenth century.

The contentious point in Hulten's quote is that a basket of goods for the colonial American evolved very differently from a representative basket of goods consumed today. The \$90 today price a different 'representative' unit than \$90 in colonial America. An individual today would have indeed turned out quite poor in colonial America, if the representative basket of goods was substantially more scarce compared to the basket of a colonial American. I check the bias in real output due to entropy in US KLEMS data, with different indexation schemes: fixing weights to expenditures in i) 1947 and ii) 2014, iii) a Törnqvist average for 1947 and 2014, iv) previous period weights (a 'chained' index), and v) average previous and current period weights (approximating a Divisia index). The 1pp gap in fixed weight indices punishes the use of current preferences to extrapolate historical living standards.

US output index with different industry weights: 1947–2014 (*Data:* `worldklems.net`)

	$\Delta \log Y_t$	$\widetilde{P}_t$	$\widetilde{Q}_t$	$\Delta H(\boldsymbol{\Omega}_t, \mathbf{Y}_t)$
Basket from 1947	6.33	3.17	2.40	0.76
Basket from 2014	6.33	3.53	3.48	-0.68
Difference	0.00	-0.36	-1.08	1.44
Törnqvist average 1947 & 2014	6.33	3.35	2.94	0.04
Chained weights	6.33	3.36	3.15	-0.18
Divisia weights	6.33	3.32	3.01	0.00

Productivity Slowdown: Technology or Allocation?

Productivity requires two aggregate statistics; nominal output and labour, both of which are distributed among N industries. Aggregate productivity growth $\Delta \log q_t = \Delta \log Y_t - \Delta \sum_{i=1}^{N} \omega_{i,t} \log P_{i,t} - \Delta \log L_t$ hinges on the entropy of both nominal output and labour input. This leads to a convenient decomposition for labour productivity between : i) 'technology' (average growth in real output per unit of labour by industry weights) and the allocation of ii) 'demand' (lower expenditures for important industries) and iii) labour (labour input shares equal to industry weights):

$$\Delta \log q_t = \underbrace{\Delta \sum_{i=1}^{N} \omega_{i,t} \log \left(\frac{Q_{i,t}}{L_{i,t}}\right)}_{\text{Technology}} + \overbrace{\underbrace{\Delta D_{KL}(\boldsymbol{\Omega}_t || \mathbf{Y}_t)}_{\text{Demand}} - \underbrace{\Delta D_{KL}(\boldsymbol{\Omega}_t || \mathbf{L}_t)}_{\text{Labour}}}^{\text{Allocation}},$$

where $D_{KL}(x||y)$ is the Kullback-Leibler (KL) divergence of y from x. Demand allocation can contribute positively: consumers benefit from important industries, with larger weights, costing less. Labour allocation contributes negatively: if the amount of labour inputs received by one industry does not match its importance in the weighting scheme, then there is a re-allocation opportunity that can decrease the KL divergence of labour from the weighting scheme towards zero.
Are the allocation terms be significant for the labour productivity slowdown? I use the conventional approach of two-period averages in the industries' nominal value added shares as the indexation scheme. Labour inputs are defined as number of hours worked. To summarise the problem, the first column reports real aggregate labour productivity growth (with a price index derived from the same weights) as an average for years pre- and post-2005, for Germany and the US. These countries experienced different roles for labour allocation, which considerably slowed labour productivity growth in Germany. In the US, better labour allocation improved productivity, thus worsening the slowdown.

Accounting for the slowdown of labour productivity growth (*Data:* `euklems.eu`)

		$\Delta \log q_t$	Technology	Demand	Labour
Germany	1995-2005	1.84	1.49	0.16	0.18
	2006-2017	0.87	0.96	0.02	-0.10
	Slowdown	0.97	0.54	0.15	0.28
	Share	1.00	0.55	0.15	0.29
United States	1997-2005	2.45	2.58	0.22	-0.35
	2006-2017	0.88	1.02	-0.01	-0.13
	Slowdown	1.57	1.56	0.23	-0.22
	Share	1.00	0.99	0.15	-0.14