

Hicksian aggregation across consumers
or (the rebound problem after derivation went wrong)
Felipe Del Canto, 9 de Septiembre de 2019

1. The mixed setting

Consider a two period economy composed by a continuum of agents that consume $n + 1$ goods: an essential good z (i.e., water) and n different goods grouped in the vector \mathbf{x} . Good z is valued at price $q \in \mathbb{R}_{++}$ and the \mathbf{x} goods are valued at prices $\mathbf{p} \in \mathbb{R}_{++}^n$. In what follows and for simplicity of notation, the subindices t for all variables will be omitted unless needed. Each individual in this setting is identified with a pair (y, α) , where $y \in [m, M]$ is her income and $\alpha \in (0, 1)$ determines the form of her pseudo-Cobb-Douglas (PCD)¹ utility function, that is,

$$u_\alpha(\mathbf{x}, z) := u_{(y, \alpha)}(\mathbf{x}, z) = \sum_{j=1}^n x_j^\alpha z^{1-\alpha} = \left(\sum_{j=1}^n x_j^\alpha \right) z^{1-\alpha}.$$

Consumers choose how much of the \mathbf{x} goods to consume in each period according to

$$\begin{aligned} \mathbf{x}(\mathbf{p}, q, y) &:= \arg \max_{\mathbf{x}, z} u_\alpha(\mathbf{x}, z), \\ \text{s.t. } &\mathbf{p}\mathbf{x} + qz = y. \end{aligned} \tag{1}$$

The pairs (y, α) follow a distribution F (with marginal distributions F_y and D_α) over its support $S := [m, M] \times (0, 1)$. In this setting, is it possible to aggregate consumption of the \mathbf{x} goods into a single good X given by:

$$g_\alpha(\mathbf{x}) = \left(\sum_{j=1}^n x_j^\alpha \right)^{1/\alpha}.$$

In that case, we have

$$U(g_\alpha(\mathbf{x}), z) := g_\alpha(\mathbf{x})^\alpha z^{1-\alpha},$$

¹The choice of this name will be clear shortly.

which is the usual Cobb-Douglas utility function. Hence, by defining $\varepsilon := (1 - \alpha)^{-1}$ and

$$P_\alpha(\mathbf{p}) := \left(\sum_{j=1}^n p_j^{1-\varepsilon} \right)^{\frac{1}{1-\varepsilon}},$$

we have that the category demand

$$\begin{aligned} X(P(\mathbf{p}), q, y) &:= \arg \max_{X, z} U(X, z) \\ \text{s.a. } &P_\alpha(\mathbf{p})X + qz = y \end{aligned} \tag{2}$$

satisfies

$$X(P_\alpha(\mathbf{p}), q, y) = g_\alpha(\mathbf{x}(\mathbf{p}, q, y)). \tag{3}$$

Note that the functions g_α and P_α depend on the value of α . Thus, in order to correctly aggregate the goods \mathbf{x} a different aggregator and price index for every person is needed.

Consider now the following situation. An investigator at time $t = 0$ has data available on disaggregate consumption of the \mathbf{x} and z goods, income y and prices (\mathbf{p}, q) . Her objective is to estimate category demands for $t = 1$. For simplicity she assumes that all agents have the same preferences, which means that all differences in consumption are accounted by differences in income. The last assumption also implies that in order to obtain the category demand of each agent and the price index, she must set a parameter $\bar{\alpha}$ which in principle could be chosen

- based on previous evidence (i.e., the income share of category X) or,
- optimally in order to minimize discrepancies between actual and predicted category demands in $t = 0$.

For the second method, the choice of $\bar{\alpha}$ should minimize the expectation of

$$\mathcal{D}(\alpha) = |g_\alpha(\mathbf{x}(\mathbf{p}, q, y)) - g_{\bar{\alpha}}(\mathbf{x}(\mathbf{p}, q, y))|,$$

which by (3) is equivalent to

$$\mathcal{D}(\alpha) = |X(P_\alpha(\mathbf{p}), q, y) - X(P_{\bar{\alpha}}(\mathbf{p}), q, y)|,$$

and this difference can be easily computed by taking advantage of the structure of problem (2),

$$\mathcal{D}(\alpha) = y \left| \frac{\alpha}{P_\alpha(\mathbf{p})} - \frac{\bar{\alpha}}{P_{\bar{\alpha}}(\mathbf{p})} \right|.$$

Thus, $\bar{\alpha}$ should satisfy

$$\bar{\alpha} = \arg \min_{\hat{\alpha} \in [0,1]} \int_S y \left| \frac{\alpha}{P_\alpha(\mathbf{p})} - \frac{\bar{\alpha}}{P_{\bar{\alpha}}(\mathbf{p})} \right| dF(y, \alpha).$$

If, as I assumed in previous settings, α and y are independent random variables, then

$$\bar{\alpha} = \arg \min_{\hat{\alpha} \in [0,1]} \int_0^1 \left| \frac{\alpha}{P_\alpha(\mathbf{p})} - \frac{\bar{\alpha}}{P_{\bar{\alpha}}(\mathbf{p})} \right| dF_\alpha(\alpha). \quad (4)$$

Intuitively, $\bar{\alpha} \in (0, 1)$ because F_α has support over the open interval and choosing a parameter away from that support should not be optimal.² As a final comment, observe that choosing $\bar{\alpha}$ needs knowledge about the distribution F_α , which may not be available.

2. World dynamics

In the second period ($t = 1$), prices change to

$$\mathbf{p}_1 = (1 + \pi)\mathbf{p}_0,$$

where π is the inflation between periods. The random variable y_2 relates to y_1 according to the following process

$$y_2 - y_1 = \mu + \varepsilon,$$

where $\mu > 0$ is a constant and $\varepsilon \sim \mathcal{N}_{[-\mu, \mu]}(0, \sigma^2)$.³ Hence, y_2 has support over $[m, M + 2\mu]$ and $E[y_2] = E[y_1] + \mu$. Preferences also change in such a way that $\mathbb{E}[\alpha_1] \neq \mathbb{E}[\alpha_0]$.⁴

²A formal argument for this claim would be nice.

³ $\mathcal{N}_{[a,b]}(\mu, \sigma^2)$ is the truncated normal distribution in $[a, b]$ with mean μ and variance σ^2

⁴I don't know how to model a reasonable change in preferences.

3. Relying on the past

Let's consider the case where $\bar{\alpha}$ is chosen based on the income share of category X , that is $\bar{\alpha} = \mathbb{E}[\alpha_0]$. If the investigator is not aware of income dynamics and has to estimate inflation by $\bar{\pi}$ (i.e., the central bank inflation goal), then her best estimation is

$$X(P_{\bar{\alpha}}((1 + \bar{\pi})\mathbf{p}, (1 + \bar{\pi})q, y_1)) = y_1 \frac{\bar{\alpha}}{P_{\bar{\alpha}}((1 + \bar{\pi})\mathbf{p})} = y_1 \frac{\bar{\alpha}}{(1 + \bar{\pi})P_{\bar{\alpha}}(\mathbf{p})},$$

where for simplicity of notation $\alpha = \alpha_1$ and $\mathbf{p} = \mathbf{p}_0$. Thus, the monetized estimation error for each person is

$$\begin{aligned} \mathcal{D}(y_1, \alpha) &= P_{\alpha}((1 + \pi)\mathbf{p}) \left| X(P_{\bar{\alpha}}((1 + \bar{\pi})\mathbf{p}, (1 + \bar{\pi})q, y_1)) - X(P_{\alpha}((1 + \pi)\mathbf{p}, (1 + \pi)q, y_2)) \right| \\ &= \left| y_1 \bar{\alpha} \frac{(1 + \pi)P_{\alpha}(\mathbf{p})}{(1 + \bar{\pi})P_{\bar{\alpha}}(\mathbf{p})} - y_2 \alpha \right|. \end{aligned} \tag{5}$$

To measure the total (monetized) estimation error one should integrate \mathcal{D} . I'm not entirely sure how this should be done.

In the scenario where the investigator takes into account the income dynamics, then her best estimation is just

$$X(P_{\bar{\alpha}}((1 + \bar{\pi})\mathbf{p}, (1 + \bar{\pi})q, y_2)) = y_1 \frac{\bar{\alpha}}{P_{\bar{\alpha}}((1 + \bar{\pi})\mathbf{p})} = y_1 \frac{\bar{\alpha}}{(1 + \bar{\pi})P_{\bar{\alpha}}(\mathbf{p})},$$

and the estimation error in this case comes only from the difference in the inflation rate and the choice of $\bar{\alpha}$

$$\begin{aligned} \mathcal{D}(y_2, \alpha) &= P_{\alpha}((1 + \pi)\mathbf{p}) \left| X(P_{\bar{\alpha}}((1 + \bar{\pi})\mathbf{p}, (1 + \bar{\pi})q, y_2)) - X(P_{\alpha}((1 + \pi)\mathbf{p}, (1 + \pi)q, y_2)) \right| \\ &= y_2 \left| \bar{\alpha} \frac{(1 + \pi)P_{\alpha}(\mathbf{p})}{(1 + \bar{\pi})P_{\bar{\alpha}}(\mathbf{p})} - \alpha \right|. \end{aligned}$$

In this case the expectation is clearly taken in terms of the distribution of y_2 and α . Although it is not that easy to compute and bound.

4. Optimizing $\bar{\alpha}$

Assume now $\bar{\alpha}$ is obtained as in (4). Observe that when the income dynamics are taken into account, if $\pi = \bar{\pi}$ then the (expected) estimation error is the smallest possible by construction.⁵ When less dynamics are considered and when $\bar{\pi} \neq \pi$, the estimation error grows but again it depends on $\mathcal{D}(y_1, \alpha)$ as obtained in (5). A valid question I don't have an answer for is: Should I incorporate $\bar{\pi}$ in problem (4)? Or in other words, should I assume $\bar{\pi} = \pi$ to simplify?

⁵Is it zero?