Answers to Empirical IO I: Problem Set 2

Part I

Consider the following setup.[a] Utility for good j in market l by consumer i is given by:

$$u_{ijl} = \delta_j - \alpha p_j + \nu_{ijl}$$

The motivation is that there are potentially many geographic markets L with different preferences for different goods (or categories of goods). For example, snow boots might be very popular in Alaska while flip flops are very popular in Florida.[b]

The main innovation is to introduce a random effect $\eta_{jl} \sim N(0, \sigma_g^2)$ into the nested logit model:[c]

$$\nu_{ijl} = \eta_{jl} + \underbrace{\zeta_{ic} + \varepsilon_{ijl}}_{GEV(\lambda)}$$

The idea is to allow for different locations l to obtain different utility from different products j. Assume (again) that the outside good exists alone in its own nest. We can write the share of product j nationwide as $\pi_j = \sum_{l=1}^{L} \omega_l \pi_{jl}$ as the share of people ω_l who live in location l and the choice probability of j within l: π_{jl}. We can decompose $\pi_{jl} = \pi_{lc} \cdot \pi_{jl|c}$ as the "product of two logits".

Define the logit denominator as:

$$D_{lc} = \sum_{j \in c} \exp\left[\frac{\delta_j - \alpha p_j + \eta_{jl}}{1 - \lambda} \right]$$

---

[a]If you get stuck take a look at Quan and Williams (Rand 2018)

[b]Unfortunately the most unusually popular product in the New York-New Jersey metro area turns out to be Ugg Boots!

[c]This exchanges $\lambda \to 1 - \lambda$ when compared to the expressions in the Train textbook. The lecture notes use both the λ and $1 - \lambda$ formulation of nested logit.

For a market l with observed η_{jl}, the within category c observable part of utility as

$$V_{jl} = \delta_j - \alpha p_j + \eta_{jl}.$$

Under nested logit with $1 - \lambda$ notation for the within cat scale, the category share of product j is

$$\pi_{jl|c} = \frac{\exp\left(\dfrac{V_{jl}}{1 - \lambda}\right)}{D_{lc}} = \frac{\exp\left(\dfrac{\delta_j - \alpha p_j + \eta_{jl}}{1 - \lambda}\right)}{\displaystyle\sum_{k \in c} \exp\left(\dfrac{\delta_k - \alpha p_k + \eta_{kl}}{1 - \lambda}\right)},$$

D_{lc} is just the inclusive value that aggregates up all the options in that category's nest.

For market l, let

$$D_{lc} = \sum_{j \in c} \exp\left(\frac{\delta_j - \alpha p_j + \eta_{jl}}{1 - \lambda}\right).$$

At the upper nest, consumers choose between each category c and the outside option. The observed component of utility for nest c is proportional to its inclusive value

$$V_{lc}^{\text{upper}} = (1 - \lambda) \ln D_{lc}.$$

We can set the outside option's utility to zero so that $\exp(V_{l0}^{\text{upper}}) = 1$, ensuring all probabilities sum to 1 by construction.

Then, the category share π_{lc} and the outside-good share π_{l0} are

$$\pi_{lc} = \frac{D_{lc}^{1-\lambda}}{1 + \sum_c D_{lc}^{1-\lambda}}, \qquad \pi_{l0} = \frac{1}{1 + \sum_c D_{lc}^{1-\lambda}}.$$

From 2) above we have $1 + \sum_{c'} D_{lc'}^{1-\lambda} = 1/\pi_{l0}$. Hence

$$\pi_{lc} = D_{lc}^{1-\lambda} \pi_{l0} \quad \Rightarrow \quad D_{lc}^{1-\lambda} = \frac{\pi_{lc}}{\pi_{l0}} \quad \Rightarrow \quad D_{lc} = \left(\frac{\pi_{lc}}{\pi_{l0}}\right)^{\frac{1}{1-\lambda}}$$

> 4) Now write $\pi_j = \sum \omega_l \cdot \pi_{lg} \cdot \pi_{lj|g}$ as an expression which depends only on: $\pi_j, \pi_{l0}, \pi_{lc}, \eta_{jl}, \lambda$ and features of j but not on features of other products.

(noting that the question introduces g as group notation, but I'll stick with c as above)

Starting from

$$\pi_j = \sum_l \omega_l\, \pi_{lc}\, \pi_{lj|c}, \qquad \pi_{lj|c} = \frac{\exp\left(\dfrac{\delta_j - \alpha p_j + \eta_{jl}}{1 - \lambda}\right)}{D_{lc}},$$

and using the result from 3) that $D_{lc} = \left(\frac{\pi_{lc}}{\pi_{l0}}\right)^{\frac{1}{1-\lambda}}$, we can obtain

$$\pi_j = \sum_l \omega_l\, \pi_{lc}\, \exp\left(\frac{\delta_j - \alpha p_j + \eta_{jl}}{1 - \lambda}\right) \left(\frac{\pi_{l0}}{\pi_{lc}}\right)^{\frac{1}{1-\lambda}}$$

> 5) Now take logs of π_j and solve for δ_j (this is the "Berry (1994) inversion"). Show that your solution reduces to Berry (1994) inversion when $\eta_{jl} = 0$ for all (j, l).

From 4) we have for each product j:

$$\pi_j = \sum_l \omega_l\, \pi_{lc}\, \exp\left(\frac{\delta_j - \alpha p_j + \eta_{jl}}{1 - \lambda}\right) \left(\frac{\pi_{l0}}{\pi_{lc}}\right)^{\frac{1}{1-\lambda}}$$

$$= \exp\left(\frac{\delta_j - \alpha p_j}{1 - \lambda}\right) \sum_l \omega_l\, \pi_{l0}^{\frac{1}{1-\lambda}}\, \pi_{lc}^{-\frac{\lambda}{1-\lambda}}\, \exp\left(\frac{\eta_{jl}}{1 - \lambda}\right)$$

Solving for δ_j:

$$\frac{\delta_j - \alpha p_j}{1 - \lambda} = \log \pi_j - \log\left(\sum_l \omega_l\, \pi_{l0}^{\frac{1}{1-\lambda}}\, \pi_{lc}^{-\frac{\lambda}{1-\lambda}}\, \exp\left(\frac{\eta_{jl}}{1 - \lambda}\right)\right),$$

so

$$\delta_j = \alpha p_j + (1 - \lambda) \log \pi_j - (1 - \lambda) \log\left(\sum_l \omega_l\, \pi_{l0}^{\frac{1}{1-\lambda}}\, \pi_{lc}^{-\frac{\lambda}{1-\lambda}}\, \exp\left(\frac{\eta_{jl}}{1 - \lambda}\right)\right)$$

If we let $\eta_{jl} = 0$ for all (j, l), the exponential term drops out:

$$\delta_j = \alpha p_j + (1 - \lambda) \log \pi_j - (1 - \lambda) \log\left(\sum_l \omega_l\, \pi_{l0}^{\frac{1}{1-\lambda}}\, \pi_{lc}^{-\frac{\lambda}{1-\lambda}}\right).$$

And in the case of a market with common π_{lc}, π_{l0} across l so $\sum_l \omega_l = 1$:

$$\delta_j = \alpha p_j + (1 - \lambda) \log \pi_j - \log \pi_{l0} + \lambda \log \pi_{lc},$$

as in Berry.

> Some trickery to avoid complicated proofs. Suppose that there are a lot of products per category so that you observe π_{lc}, π_{l0} without any error. Moreover assume that any single product has a negligible effect on π_{lc}. The idea is that you can treat the observed values of π_{lc}, π_{l0} and thus D_{lc} as data that are unaffected by your parameters.

> 6) Find the closed form solution to the integral $\int \exp[\eta_{jl}/(1 - \lambda)] f(\eta_{jl}|\sigma^2) d\eta_{jl}$.

$$\int \exp\left(\frac{\eta_{jl}}{1 - \lambda}\right) f(\eta_{jl} \mid \sigma^2) d\eta_{jl} = M_\eta\left(\frac{1}{1 - \lambda}\right),$$

where

$$M_\eta(t) \equiv \int \exp(t\,\eta) f(\eta \mid \sigma^2) d\eta$$

is the moment generating function of η_{jl}

We have that $\eta_{jl} \sim \mathcal{N}(0, \sigma^2)$ independently across (j, l), so from the formula for the mgf of the normal:

$$M_\eta(t) = \exp\left(\tfrac{1}{2}\sigma^2 t^2\right)$$

so

$$\int \exp\left(\frac{\eta_{jl}}{1 - \lambda}\right) f(\eta_{jl} \mid \sigma^2) d\eta_{jl} = \exp\left(\frac{\sigma^2}{2(1 - \lambda)^2}\right).$$

> 7) Plug your solution to (6) into your solutions for (5) and (4). For this to be allowed this requires applying the law of iteration expectations and some technical conditions that we will ignore.

Using the result from (4):

$$\pi_j = \sum_l \omega_l \pi_{lc} \exp\left(\frac{\delta_j - \alpha p_j + \eta_{jl}}{1 - \lambda}\right) \left(\frac{\pi_{l0}}{\pi_{lc}}\right)^{\frac{1}{1 - \lambda}},$$

Factor out the j–specific terms and collect l–specific pieces:

$$\pi_j = \exp\left(\frac{\delta_j - \alpha p_j}{1 - \lambda}\right) \sum_l \omega_l \pi_{l0}^{\frac{1}{1 - \lambda}} \pi_{lc}^{-\frac{\lambda}{1 - \lambda}} \exp\left(\frac{\eta_{jl}}{1 - \lambda}\right).$$

4

Take expectations over η_{jl}, treating $\{\omega_l, \pi_{lc}, \pi_{l0}\}$ as data:

$$\pi_j = \exp\left(\frac{\delta_j - \alpha p_j}{1 - \lambda}\right) \sum_l \omega_l \, \pi_{l0}^{\frac{1}{1-\lambda}} \, \pi_{lc}^{-\frac{\lambda}{1-\lambda}} \, \mathbb{E}\left[\exp\left(\frac{\eta_{jl}}{1 - \lambda}\right)\right].$$

This expectation is the moment generating function of η_{jl}:

$$\pi_j = \exp\left(\frac{\delta_j - \alpha p_j}{1 - \lambda}\right) \left(\sum_l \omega_l \, \pi_{l0}^{\frac{1}{1-\lambda}} \, \pi_{lc}^{-\frac{\lambda}{1-\lambda}}\right) M_\eta\left(\frac{1}{1 - \lambda}\right).$$

Since $\eta_{jl} \sim \mathcal{N}(0, \sigma_g^2)$, as in (6):

$$\pi_j = \exp\left(\frac{\delta_j - \alpha p_j}{1 - \lambda}\right) \left(\sum_l \omega_l \, \pi_{l0}^{\frac{1}{1-\lambda}} \, \pi_{lc}^{-\frac{\lambda}{1-\lambda}}\right) \exp\left(\frac{\sigma_g^2}{2(1 - \lambda)^2}\right).$$

For the same reason we also have:

$$\mathbb{E}\left[\exp\left(\frac{\eta_{jl}}{1 - \lambda}\right)\right] = \exp\left(\frac{\sigma_g^2}{2(1 - \lambda)^2}\right),$$

Substituting this result into the previous expression gives:

$$\pi_j = \exp\left(\frac{\delta_j - \alpha p_j}{1 - \lambda}\right) \exp\left(\frac{\sigma_g^2}{2(1 - \lambda)^2}\right) \sum_l \omega_l \, \pi_{l0}^{\frac{1}{1-\lambda}} \, \pi_{lc}^{-\frac{\lambda}{1-\lambda}}.$$

If we define the weighted market term:

$$S \; = \; \sum_l \omega_l \, \pi_{l0}^{\frac{1}{1-\lambda}} \, \pi_{lc}^{-\frac{\lambda}{1-\lambda}}.$$

We can write

$$\pi_j = \exp\left(\frac{\delta_j - \alpha p_j}{1 - \lambda}\right) \exp\left(\frac{\sigma_g^2}{2(1 - \lambda)^2}\right) S$$

Taking logs and solving for δ_j Berry-style gives

$$\frac{\delta_j - \alpha p_j}{1 - \lambda} = \log \pi_j - \log S - \frac{\sigma_g^2}{2(1 - \lambda)^2},$$

and hence

$$\delta_j = \alpha p_j + (1 - \lambda) \log \pi_j - (1 - \lambda) \log S - \frac{\sigma_g^2}{2(1 - \lambda)}.$$

The final term $-\frac{\sigma_g^2}{2(1-\lambda)}$ is the contribution of the random–effects element.

8) [Hard] Discuss how this relates to Ackerberg and Rysman (Rand 2005).

Ackerberg and Rysman (2005) show what happens when we don't make the simplifying assumption that we can treat (π_{lc}, π_{l0}) as data and therefore replace η_{jl} with its expectation. In reality, there will be sampling error and simulation bias that we have to deal with. We should really keep η_{jl} inside the simulated choice probabilities, which will make a difference for consistency through Jensen's inequality.

Part II

Now consider a very particular nested logit model:[a]

$$u_{ij} = \delta_j - \alpha p_j + \epsilon_{ij}$$

We assume that ϵ_{ij} is distributed GEV with some nesting parameter σ like the Cardell (1997) paper or the Berry (1994) paper. Assume that all products are in the same nest, so that the outside good is the only product in the second nest.

---

[a]If you get stuck look at Ackerberg and Rysman (Rand 2005) or Kroft, Laliberte, Leal-Vizcaino and Notowidigdo (WP 2017).

9) [Easy] Derive a simplified form of the nested logit choice probabilities and the expected (ex-ante) consumer surplus in dollars. [Just write them from Berry (1994)].

Let $V_j = \delta_j - \alpha p_j$ and write the set of inside products as $\mathcal{J}$. Define

$$D = \sum_{k \in \mathcal{J}} \exp\left(\frac{V_k}{1 - \sigma}\right).$$

With all inside products in one nest and the outside good alone in the second nest, the nested logit probabilities are:

$$s_{j|\text{in}} = \frac{\exp\left(V_j/(1-\sigma)\right)}{D}, \qquad s_{\text{in}} = \frac{D^{1-\sigma}}{1 + D^{1-\sigma}}, \qquad s_0 = \frac{1}{1 + D^{1-\sigma}}.$$

Hence the product choice probability is

$$s_j = s_{j|\text{in}} \cdot s_{\text{in}} = \frac{\exp\left(V_j/(1-\sigma)\right)}{D} \cdot \frac{D^{1-\sigma}}{1 + D^{1-\sigma}}.$$

Expected consumer surplus (level) is the log-sum at the upper level divided by the marginal utility of income α:

$$CS = \frac{1}{\alpha} \ln(1 + D^{1-\sigma}),$$

since the two-nest generator yields an upper-level inclusive value $\ln(1 + D^{1-\sigma})$. This looks like standard multinomial logit when $\sigma \to 0$.

> 10) [Easy] Now assume that all products are symmetric. That is they have the same $\delta_j = \delta$ and $p_j = p$. Derive the nested logit choice probabilities and the expected (ex-ante) consumer surplus in dollars.

With symmetry: $V_j = \delta_j - \alpha p_j \equiv V$ for all $j = 1, \ldots, J$. Let

$$D = \sum_{k=1}^{J} \exp\left(\frac{V}{1 - \sigma}\right) = J \exp\left(\frac{V}{1 - \sigma}\right).$$

Then

$$s_{j|\text{in}} = \frac{\exp(V/(1 - \sigma))}{D} = \frac{1}{J}, \qquad s_{\text{in}} = \frac{D^{1-\sigma}}{1 + D^{1-\sigma}}, \qquad s_0 = \frac{1}{1 + D^{1-\sigma}}.$$

Using $D^{1-\sigma} = (J \exp(V/(1 - \sigma)))^{1-\sigma} = J^{1-\sigma} \exp(V)$,

$$s_{\text{in}} = \frac{J^{1-\sigma} \exp(V)}{1 + J^{1-\sigma} \exp(V)}, \qquad s_0 = \frac{1}{1 + J^{1-\sigma} \exp(V)}.$$

Hence each product's probability is

$$s_j = s_{j|\text{in}} s_{\text{in}} = \frac{1}{J} \cdot \frac{J^{1-\sigma} \exp(V)}{1 + J^{1-\sigma} \exp(V)} = \frac{J^{-\sigma} \exp(V)}{1 + J^{1-\sigma} \exp(V)}.$$

Expected consumer surplus is the upper-level log-sum divided by α:

$$CS = \frac{1}{\alpha} \ln(1 + D^{1-\sigma}) = \frac{1}{\alpha} \ln(1 + J^{1-\sigma} \exp(V)).$$

> 11) Derive $Q(p, J)$ the total market demand for J goods each with price p. The result for $Q(p, J)$ should not include any summations. Also derive the ex-ante consumer surplus expression without a summation.

Total market demand (the inside share) is

$$Q(p, J) = \frac{D^{1-\sigma}}{1 + D^{1-\sigma}} = \frac{J^{1-\sigma} \exp(\delta - \alpha p)}{1 + J^{1-\sigma} \exp(\delta - \alpha p)}.$$

The (ex-ante) consumer surplus (level, in dollars) is the upper-level log-sum divided by α:

$$CS(p, J) = \frac{1}{\alpha} \ln\left(1 + D^{1-\sigma}\right) = \frac{1}{\alpha} \ln\left(1 + J^{1-\sigma} \exp(\delta - \alpha p)\right).$$

12) Derive the inverse demand $P(Q, J)$ by solving $Q(p, J)$ for p rather than Q. Does an increase in variety cause a parallel shift in the inverse demand curve or not?

Start from 11)

$$Q(p, J) = \frac{J^{1-\sigma} \exp(\delta - \alpha p)}{1 + J^{1-\sigma} \exp(\delta - \alpha p)}.$$

Let $y \equiv J^{1-\sigma} \exp(\delta - \alpha p)$. Then $Q = \frac{y}{1+y} \Rightarrow y = \frac{Q}{1-Q}$, so

$$\exp(\delta - \alpha p) = \frac{Q}{(1-Q) J^{1-\sigma}}.$$

Taking logs and solving for p,

$$\delta - \alpha p = \ln\left(\frac{Q}{1-Q}\right) - (1-\sigma) \ln J \quad \Rightarrow \quad P(Q, J) = \frac{\delta}{\alpha} - \frac{1}{\alpha} \ln\left(\frac{Q}{1-Q}\right) + \frac{1-\sigma}{\alpha} \ln J.$$

Effect of variety. J only enters through the constant term $\frac{1-\sigma}{\alpha} \ln J$. So an increase in variety does shift the inverse demand in parallel (i.e. vertically), with the slope unchanged. Plug in values J_1 and J_2:

$$P(Q, J_2) - P(Q, J_1) = \frac{1-\sigma}{\alpha} \ln\left(\frac{J_2}{J_1}\right),$$

which is all constant.

13) Let's think about what happens when we add an additional product $J \to J+1$ in this market. One is that it affects the overall price index p or P and the other is that it causes a shift/rotation in the (inverse) demand curve. Show that $\frac{dCS}{dJ} = -Q\frac{dp}{dj} + c \cdot q_{j+1}$ where c is a constant that depends on the parameters of the model and q_{j+1} is the sales of the new product after introduction. [Hint: Recall that CS is the integral under the demand curve. You will probably want to apply the Leibniz Integral Rule and exchange differentiation and integration.]

CS is the area under the inverse demand curve:

$$CS(J) = \int_0^{Q(J)} P(q, J) \, dq.$$

Differentiate using Leibniz:

$$\frac{d\,CS}{dJ} \;=\; \int_0^Q \frac{\partial P(q,J)}{\partial J}\,dq \;+\; P(Q,J)\frac{dQ}{dJ}.$$

Impose market clearing with observed price $p(J)$ so that $P(Q,J)=p(J)$:

$$\frac{d\,CS}{dJ} \;=\; \underbrace{\int_0^Q \frac{\partial P(q,J)}{\partial J}\,dq}_{\text{variety (shift) effect}} + p(J)\frac{dQ}{dJ}.$$

Add and subtract $Q\frac{dp}{dJ}$, and use the identity $\frac{dQ}{dJ} = \left(\frac{\partial P}{\partial Q}\right)^{-1}\left(\frac{dp}{dJ} - \frac{\partial P}{\partial J}\right)$ (which follows by totally differentiating $P(Q,J)=p(J)$) to obtain

$$\frac{d\,CS}{dJ} \;=\; -Q\,\frac{dp}{dJ} \;+\; \underbrace{\left[\int_0^Q \frac{\partial P(q,J)}{\partial J}\,dq \;-\; p(J)\left(\frac{\partial P}{\partial Q}\right)^{-1}\frac{\partial P}{\partial J}\right]}_{c\,\cdot\,q_{j+1}}.$$

Interpretation. The first term is the price-index channel: if equilibrium price p falls as variety rises, consumers gain $-Q\,dp/dJ$. The second term is the *direct variety* channel: the new product generates extra area under demand equal to a constant c times its sales q_{j+1}.

With $P(q,J) = \frac{\delta}{\alpha} - \frac{1}{\alpha}\ln\!\left(\frac{q}{1-q}\right) + \frac{1-\sigma}{\alpha}\ln J$, we have $\partial P/\partial J = \frac{1-\sigma}{\alpha J}$ (independent of q). Hence

$$\int_0^Q \frac{\partial P}{\partial J}\,dq \;=\; \frac{1-\sigma}{\alpha J}\,Q,$$

so the variety term is linear in Q. Writing this as $c\cdot q_{j+1}$ just defines $c = \frac{1-\sigma}{\alpha J}\frac{Q}{q_{j+1}}$ (which depends on parameters via J,σ,α and the equilibrium shares). Under symmetry the new product's share is

$$q_{j+1} \;=\; \frac{(J+1)^{-\sigma}\,\exp(\delta-\alpha p)}{1+(J+1)^{\,1-\sigma}\,\exp(\delta-\alpha p)}.$$

The welfare effect splits into a price index (parallel shift) term and a direct new-variety term proportional to the new product's sales.

---

14) What does this tell us about what is important to observe in order to measure the welfare impact of new product variety?

To measure the welfare impact of new product variety, we need to observe both prices and quantities: particularly how total demand Q and equilibrium price p respond when J changes. Without data on prices (the price index channel) and new product sales q_{j+1} (the variety channel), we have no way to decompose the welfare gain.

Part III

People often claim that in order to understand welfare in the mixed logit model, we only need to look at the share out of the outside good. Show me why this is (almost) true and how it relates to the willingness-to-pay (WTP) measure of:

> Town, Robert and Gregory Vistnes, "Hospital competition in HMO networks," *Journal of Health Economics*, 2001, 20(5), 733–753.
>
> Capps, Cory, David Dranove, and Mark Satterthwaite, "Competition and Market Power in Option Demand Markets," *RAND Journal of Economics*, September 2003, 34(4), 737–763.

$$WTP_i(j) = E[\max_{k \in \mathcal{J}} u_{ij}] - E[\max_{k \in \mathcal{J} \setminus j} u_{ik}]$$

Normalise the outside option's deterministic utility to zero. For a mixed logit with Type I EV errors, an individual's expected surplus is

$$CS_i \;=\; \frac{1}{\alpha_i} \ln\Big(1 + \sum_{k \in \mathcal{J}} \exp\big(V_{ik}\big)\Big) \;=\; -\frac{1}{\alpha_i} \ln s_{0i}, \quad s_{0i} \;=\; \frac{1}{1 + \sum_{k \in \mathcal{J}} \exp(V_{ik})}.$$

Hence changes in welfare depend only on changes in the outside-good probability at the individual level:

$$\Delta CS_i \;=\; -\frac{1}{\alpha_i} \Delta \ln s_{0i}.$$

Aggregating, $CS = \int CS_i \, dF$ implies $CS = -\frac{1}{\alpha} \mathbb{E}[\ln s_{0i}]$ when $\alpha_i \equiv \alpha$. So in this sense, we do only need the outside share: welfare levels and changes are pinned down by s_{0i}.

But by Jensen's inequality, $\mathbb{E}[\ln s_{0i}] \neq \ln \mathbb{E}[s_{0i}]$. So we can't recover welfare exactly from the aggregate outside share alone without modeling heterogeneity. In practice, this might mean that differences in individual substitution patterns and income sensitivities matter for welfare, even if aggregate outside shares look similar across markets.

For the WTP measure, removing option j changes individual surplus by

$$WTP_i(j) \;=\; \frac{1}{\alpha_i}\Big[\ln\Big(1 + \sum_{k \in \mathcal{J}} \exp(V_{ik})\Big) - \ln\Big(1 + \sum_{k \in \mathcal{J} \setminus \{j\}} \exp(V_{ik})\Big)\Big] \;=\; \frac{1}{\alpha_i} \ln\Big(\frac{1}{1 - s_{ij}}\Big),$$

since $1 - s_{ij} = \dfrac{1 + \sum_{k \neq j} \exp(V_{ik})}{1 + \sum_{k} \exp(V_{ik})}$. This is the formula in the referenced papers and in mixed logit it is computed at the individual draw level and then averaged.

Part IV (Optional)

This part is optional. If you are not an economics student, you can skip this part (probably because you didn't cover the envelope theorem in first year).

Consider a representative consumer facing a price vector p endowed with wealth Z. Z is denominated in units of the numeraire good y. To make life easy, you can assume that $u(x)$ is strictly quasi-concave.

$$\max_{x,y} u(x_1, \ldots, x_J) + y \quad \text{s.t.} \quad p \cdot x + tx_1 + y = Z$$

There is also a representative firm which takes prices as given and chooses output

$$\max_x p \cdot x - c(x)$$

The market clears so that $x^d(p) = x^s(p)$. We are interested in how a tax on x_1 might affect total surplus in this market. Recall that total surplus is the sum of the consumer and producer surplus. Also recall that total surplus is about *allocation* and prices are just a transfer from buyers to sellers, so we can re-write this as the planner's problem. Assume that we add a tax t on good 1 and return that tax to consumers.

$$W(t) = \max_x \{u(x) + Z - tx_1 - p(t) \cdot x\} + \max_x \{p(t) \cdot x - c(x)\} + tx_1 = \max_x \{u(x) + Z - tx_1 - c(x)\} + tx_1$$

1.) Show that the *total derivative* $\frac{dW}{dt} = t\frac{dx_1(t)}{dt}$. In other words, the part in the curly brackets cancels out.[a] This requires invoking the envelope theorem, and ignoring the possibility of oligopoly pricing. I have already cancelled out prices (a transfer) but you need to make substitution across other goods cancel.

---
[a]If you get horribly stuck, consider Chetty (2009)'s Annual Review article.

The planner's problem is:

$$W(t) = \max_x \{\, u(x) + Z - t\,x_1 - c(x) \,\} + t\,x_1.$$

Let $x^*(t)$ solve the inner maximisation and define $F(x,t) \equiv u(x) + Z - t\,x_1 - c(x)$, so $W(t) = F(x^*(t), t) + t\,x_1^*(t)$.

Envelope theorem: the total derivative is just the partial derivative of the objective with respect to the parameter, evaluated at the optimum.

So we can differentiate using the total derivative and invoke the envelope theo-

rem:

$$\frac{dW}{dt} = \underbrace{\frac{\partial F}{\partial t}}_{=-x_1^*} + \underbrace{\left(\nabla_x F\big|_{x^*}\right)}_{=0} \cdot \frac{dx^*}{dt} + x_1^* + t\frac{dx_1^*}{dt} = -x_1^* + x_1^* + t\frac{dx_1^*}{dt} = t\frac{dx_1(t)}{dt}.$$

Here $\nabla_x F\big|_{x^*} = 0$ are the first-order conditions for the optimal allocation, so all substitution effects across goods cancel in the total derivative. The prices have already been eliminated as pure transfers in the planner's objective.

We can see the logit has this property in a more concrete way. If you don't get the first part – you can start here (this is a bit easier). Given $u_{ij} = v_{ij} + \varepsilon_{ij}$ where ε_{ij} has a Type I EV (logit) distribution. We can derive individual purchase probabilities:

$$S_{ij} = \frac{\exp v_{ij}}{\sum_k \exp v_{ik}}$$

To make life easier assume that $\frac{\partial v_{ij}}{\partial p_{ij}} = 1$ for all (i,j) or that utility is quasi-linear (and money-metric in price).[a]

---

[a] You can relax this but then you need to carry around $\frac{1}{\alpha_i}$ terms.

---

2) Write the individual surplus function $S_i = E[\max_j u_{ij}]$ and the total welfare function $W(t)$ (for all consumers and the representative firm) where $\pi(p) = p \cdot S - c(S)$ (where S is the stacked vector of S_j and $S_j = \sum_i n_i S_{ij}$, the sum of individual demands).

For each individual i, expected surplus under the logit model is

$$S_i = \mathbb{E}\left[\max_j (v_{ij} + \varepsilon_{ij})\right] = \gamma + \ln\left(1 + \sum_j \exp(v_{ij})\right),$$

where γ is Euler's constant. The individual choice probability is

$$S_{ij} = \frac{\exp(v_{ij})}{1 + \sum_k \exp(v_{ik})}.$$

Aggregating across consumers, total demand for product j is

$$S_j = \sum_i n_i S_{ij},$$

and the stacked vector of these demands can be denoted $S = (S_1, \ldots, S_J)^\top$. Total welfare, including both consumer and producer surplus, is therefore

$$W(t) = \sum_i n_i S_i\big(v_i(p(t))\big) + \pi(p(t)), \qquad \pi(p) = p \cdot S - c(S).$$

Substituting gives

$$W(t) \;=\; \sum_i n_i \ln\!\Big(1+\sum_j \exp(v_{ij}(p(t)))\Big) \;+\; p(t)\cdot S(p(t)) \;-\; c\big(S(p(t))\big).$$

3) Use the envelope theorem to show that $\frac{dW(t)}{dt} = t\frac{d\,S_1(t)}{dt_1}$ for a tax t_1 on good 1 that is returned lump sum to consumers.

Let the tax t_1 enter the deterministic utility of good 1 as $v_{i1}(t_1) = v_{i1}^0 - t_1$, and individual surplus under logit is $S_i(t_1) = \mathbb{E}[\max_j\{v_{ij}(t_1) + \varepsilon_{ij}\}] = \gamma + \ln\big(1 + \sum_j \exp(v_{ij}(t_1))\big)$. Aggregate consumer surplus is $CS(t_1) = \sum_i n_i S_i(t_1)$, so

$$\frac{d\,CS(t_1)}{dt_1} = \sum_i n_i \frac{\partial S_i}{\partial v_{i1}}\frac{\partial v_{i1}}{\partial t_1} = \sum_i n_i\, S_{i1}(t_1)\cdot(-1) = -\,S_1(t_1),$$

where $S_{i1}(t_1)$ is the logit choice probability and $S_1(t_1) = \sum_i n_i S_{i1}(t_1)$.
Total welfare (consumers + firm) with lump-sum rebate of the tax revenue is

$$W(t_1) = CS(t_1) \;+\; \pi\big(p(t_1)\big) \;+\; t_1\, S_1(t_1), \qquad \pi(p) = p\cdot S - c(S).$$

Differentiate and use the envelope theorem for the firm (the profit term's direct effect drops out at the optimum):

$$\frac{dW(t_1)}{dt_1} = \frac{d\,CS(t_1)}{dt_1} + \underbrace{\frac{d\,\pi(p(t_1))}{dt_1}}_{\text{envelope}\,=\,0} + \frac{d}{dt_1}\big[t_1 S_1(t_1)\big] = -S_1(t_1) + S_1(t_1) + t_1\,\frac{dS_1(t_1)}{dt_1}.$$

Hence,

$$\frac{dW(t_1)}{dt_1} \;=\; t_1\,\frac{dS_1(t_1)}{dt_1}.$$

4) What does this tell us about the welfare effects of taxation in this model? What would we want to observe in the data? What is the most important feature of our model?

This result shows that in the model, the welfare effects of taxation come only through how the tax changes quantities rather than through prices themselves.

To measure these effects empirically, we'd want to observe how demand for the taxed good responds to the tax (the slope of $S_1(t_1)$).

The key feature is quasi-linearity, so there are no income effects and taxes only distort quantities, not overall values.