Last updated: September 26, 2017

PhD Micro (Part 1)
More on Consumer Theory

Motaz Al-Chanati

September 26, 2017

1 Integrability

Let's recall how we finished the last section. We had two claims:

Demand and Utility	*Properties*
$x(p, w)$ can be generated by a utility function	(1) $x(p, w)$ is homogenous of degree 0
	(2) Walras' Law is satisfied
	(3) Slutsky matrix is negative semi-definite

We've already proved the $\Rightarrow$ direction, which means that those three properties give us testable implications for consumer theory. Now we want to go the other way $\Leftarrow$, i.e. if we observe the three properties, does that mean $x(p, w)$ can be generated by a utility function? This is called integrability, and the answer is yes.

To prove this we need to use *Frobenius' Theorem*, which says that a system of partial differential equations (PDEs) of the form:

$$\frac{\partial f(p)}{\partial p_k} = g_k(p)$$

has a solution if and only if:

$$\frac{\partial g_l(p)}{\partial p_k} = \frac{\partial g_k(p)}{\partial p_l}, \forall l, k$$

This condition is was referred to in class as *compatability*. Shephard's Lemma has already given us the system of PDEs in exactly the desired form:

$$\frac{\partial e(p, u)}{\partial p_k} = h_k(p, u) = x_k(p, e(p, u))$$

This means it has a solution if and only if:

$$\frac{dx_l(p, e(p, u))}{dp_k} = \frac{dx_k(p, e(p, u))}{dp_l}, \forall l, k$$

$$\implies \frac{\partial x_l(p, w)}{\partial p_k} + \frac{\partial x_l(p, w)}{\partial w} x_k(p, w) = \frac{\partial x_k(p, w)}{\partial p_l} + \frac{\partial x_k(p, w)}{\partial w} x_l(p, w), \forall l, k$$

But this condition is exactly satisfied because the Slutsky matrix is symmetric. So the PDE system has a solution, which will be the expenditure function. How do we find the utility function from this? We just use the relationships we found before:

$$e(p,u) \xrightarrow{\text{Invert}} v(p,w) \xrightarrow{\text{Duality}} u(x)$$

And so we're done! Let's work through an example to see why this is useful.

Example: Suppose that demand for each good is $x_i(p,w) = \frac{\alpha_i w}{p_i}$, where $\sum_i \alpha_i = 1$ and $w > 1$. Let's check each of the conditions:

1. Homogeneous of degree 0: $x_i(\lambda p, \lambda w) = \frac{\alpha_i \lambda w}{\lambda p_i} = \frac{\alpha_i w}{p_i} = x_i(p,w)$
2. Walras' Law: $p \cdot x(p,w) = \sum_i p_i x_i(p,w) = \sum_i \alpha_i w = w \times \sum_i \alpha_i = w$
3. Slutsky Matrix: $\frac{\partial x_l(p,w)}{\partial p_k} = 0$ if $l \neq k$ and $-\frac{\alpha_l w}{p_l^2}$ if $l = k$. $\frac{\partial x_l(p,w)}{\partial w} = \frac{\alpha_l}{p_l}, \forall l$. Therefore, $\frac{\partial x_l(p,w)}{\partial p_k} + \frac{\partial x_l(p,w)}{\partial w} x_k(p,w) = \frac{\alpha_l}{p_l}\frac{\alpha_k}{p_k}$ if $l \neq k$ and $\frac{\alpha_l(\alpha_l - w)}{p_l^2}$ if $l = k$. To check whether it is negative semi-definite, just follow the steps as in Problem Set 2 Question 1 (MWG Ex 3.G.3).

So everything is satisfied and we can now set up our system of PDEs and solve it:

$$\begin{aligned}
\frac{\partial e(p,u)}{\partial p_k} &= x_k(p, e(p,u)) = \frac{\alpha_k e(p,u)}{p_k} \\
\frac{1}{e(p,u)}\frac{\partial e(p,u)}{\partial p_k} &= \frac{\alpha_k}{p_k} \\
\frac{\partial \log(e(p,u))}{\partial p_k} &= \alpha_k \frac{\partial \log p_k}{\partial p_k} && \left[\tfrac{d \log f(x)}{dx} = \tfrac{d \log f(x)}{df(x)}\tfrac{df(x)}{dx} = \tfrac{1}{f(x)}\tfrac{df(x)}{dx}\right] \\
\log(e(p,u)) &= \alpha_k \log p_k + C_k && [\text{Integrate with respect to } p_k] \\
e(p,u) &= p_k^{\alpha_k} e^{C_k}
\end{aligned}$$

Where C_k is an integrating constant that does not depend on p_k, i.e. it can depend on $p_j, \forall j \neq k$ and u (remember we are in the EMP world so prices and the utility threshold are our parameters). We have this system of K equations where the LHS is always $e(p,u)$ so in order to make all the RHS equal too, we need $e^{C_k} = \left(\prod_{j\neq k} p_j^{\alpha_j}\right) \times c(u)$, where $c(u)$ is some function of u.[1] This gives us:

$$e(p,u) = c(u)\prod_k p_k^{\alpha_k}$$

Remember that one of our properties was that $e(p,u)$ was strictly increasing in u, so we need to ensure that $c(u)$ is also strictly increasing in u. For example, we can choose $c(u) = u$ and we then have the expenditure function be $e(p,u) = u\prod_k p_k^{\alpha_k}$.

We're not done just yet. We want to find a utility function, so we have to take two more steps. The first is to 'invert' to get the indirect utility function. Since we're going from EMP world to UMP world, we need to get rid of u and bring in w, i.e. we set $e(p,u) = w$ and $u = v(p,w)$. Solving gives us:

$$v(p,w) = \frac{w}{\prod_k p_k^{\alpha_k}}$$

Finally, we to get the utility function $u(x)$ from the indirect utility function $v(p,w)$ we know that we can use the dual of the UMP: we minimize the indirect utility function with respect

[1] If this doesn't seem to clear to you, try a simple example. Suppose you had $a = b \times f_1(b)$ and $a = d \times f_2(b)$. This must mean that $f_1(d) = d$ and $f_2(b) = b$. We are essentially using the same logic here.

to the budget constraint and the optimal objective will be the utility function. However, our lives are easier here because we already know the Marshallian demand. Inverting the demand function gives us (not-so-surprisingly) the inverse demand function $p(x)$, which is the solution to this minimization problem. In this case $p_i(x, w) = \frac{\alpha_i w}{x_i}$. So we can plug this into the objective function $v(p, w)$ to get the value function $u(x) = \prod_k \left(\frac{x_k}{\alpha_k}\right)^{\alpha_k}$. So this utility function generates the Marshallian demand we started with, $x_i(p, w) = \frac{\alpha_i w}{p_i}$. Notice that since the UMP solution set is invariant to strictly increasing transformations of the utility function, then $\tilde{u}(x) = \prod_k x_k^{\alpha_k}$ is another utility that generates the same Marshallian demand. But you would have already guessed this because this is just our usual Cobb-Douglas formulation.

2 Labor Supply

2.1 Application

The story here is that the agent has to choose between two goods: consumption and leisure. The trade-off comes from the fact consuming leisure means reducing your labor, which in turn reduces your income. This is essentially going to be a UMP so let's set up all our components.

For this section, we will use slightly different notation, so just be careful. To reduce confusion, I will keep it consistent with the class notation. We have a consumption good C (think of this as a composite good) and leisure L (i.e. not working). C is measured in units but L is measured in hours. The agent enjoys both of these goods (they increase utility), so we write the utility function as $u(C, L)$.

Of course, goods come at a cost to the agent, so there is a price of C, p_C, and a price of L, p_L. For simplicity, we can normalize $p_C = 1$. The cost of leisure is really just the opportunity cost of not working, i.e. it is the wages the agent has foregone in order to have leisure. So we can set $p_L = w$ where w represents the hourly wage of the agent (careful, we used to denote income by w but now it is essentially a price!).

We've figured out the prices of goods, but to have a budget constraint we need to figure out the 'income' (the RHS of the budget constraint). We assume the agent has two sources of income: a fixed amount (e.g. through transfer or inheritance) and a variable amount determined by hours worked. Let's called the fixed amount y and since the hourly wage is w, the variable amount is wH, where H represents hours worked. Normally, we only restrict our demand to be non-negative but since we are working with hours we should put in an additional restriction: $L + H = T$, where T is the total number of hours available in the day for leisure and labor. Maybe you consider $T = 24$ (so that someone could be working 24/7) or you might consider $T = 16$ (so that you let your agent get 8 hours of sleep per day). In any case, we will just leave it as T. What this means is that the *maximum* income the agent could be getting is $y + wT$, i.e. working all the time. We call this amount the *total potential income*, which we will denote with $Y = y + wT$.

This means our UMP is:

$$\begin{gathered} \max_{C,L} u(C, L) \\ \text{s.t. } C + wL \leq Y \\ C \in \mathbb{R}_+, L \in [0, T] \end{gathered}$$

As we know, we will get a Marshallian demand from the UMP. Notice our parameters are still price ($(1, w)$ or simply w) and income Y. Let's call the Marshallian demand for leisure as $\Lambda(w, Y)$. But note that by definition of Y, wage appears twice in the demand function, i.e. $\Lambda(w, y + wT)$. Let's write demand for leisure in a different way: $L(w, y)$, where it is now a function of wage and the fixed transfer income. By definition $L(w, y) = \Lambda(w, Y)$ because they are both demand for leisure. Our goal is to see how leisure changes when w and y change.

We start with the identity $L(w, y) = \Lambda(w, Y)$. Let's first check the case for y by taking the derivative on both sides with respect to y.

$$\begin{aligned} \frac{\partial L(w,y)}{\partial y} &= \frac{\partial \Lambda(w,Y)}{\partial Y}\frac{\partial Y}{\partial y} \\ &= \frac{\partial \Lambda(w,Y)}{\partial Y} \end{aligned}$$

Now let's do the case for w, noting that we actually need to take a total derivative for Λ:

$$\begin{aligned} \frac{\partial L(w,y)}{\partial w} &= \frac{d\Lambda(w,Y)}{dw} \\ &= \frac{\partial \Lambda(w,Y)}{\partial w} + \frac{\partial \Lambda(w,Y)}{\partial Y}\frac{\partial Y}{\partial w} && \text{[total derivative]} \\ &= \frac{\partial \Lambda(w,Y)}{\partial w} + \frac{\partial \Lambda(w,Y)}{\partial Y}T \\ &= \left(\frac{\partial \tilde{\Lambda}(w,u)}{\partial w} - \frac{\partial \Lambda(w,Y)}{\partial Y}L(w,y)\right) + \frac{\partial \Lambda(w,Y)}{\partial Y}T && \text{[apply Slutsky equation]} \\ &= \frac{\partial \tilde{\Lambda}(w,u)}{\partial w} + \frac{\partial \Lambda(w,Y)}{\partial Y}(T - L(w,y)) \\ &= \frac{\partial \tilde{\Lambda}(w,u)}{\partial w} + \frac{\partial L(w,y)}{\partial y}(T - L(w,y)) && \text{[by first relationship found]} \end{aligned}$$

Where $\tilde{\Lambda}(w, u)$ represents the Hicksian demand for leisure. The two parts we're seeing here are the familiar income and substitution effects. On the one hand, a higher wage means that the agent feels richer and so wants to consume more of all goods (including leisure). This income effect is captured in the second term. Since leisure is a normal good, $\frac{\partial L(w,y)}{\partial y} \geq 0$ and $L \leq T$, then this term is positive. On the other hand, a higher wage means that the price of leisure has also gone up (you are forgoing even more wages when you are not working). This means you would want to increase labor/reduce leisure and use more of your money on buying the consumption good (substitution effect). This is captured by the first term. Since this is the Hicksian demand, we know that it has negative own-substitution terms (remember that w is the price of leisure), so $\frac{\partial \tilde{\Lambda}(w,u)}{\partial w} \leq 0$.

Going back to the original question, how does a change in wage affect labor supply? The answer is ambiguous! This is quite surprising because before it was ambiguous due to inferior goods. But now it is still ambiguous even though leisure is a normal good. The key difference here is that the wage is both the price of leisure and a determinant of total income. What we do know is that if $L(w, y) = T$, then there will be no income effect because the second term will be zero. That means an increase in your wage can only have a substitution effect, i.e. you are already consuming your maximum hours of leisure so you cannot feel any richer and consume *even more* leisure. This gives us the idea of of a *reservation wage*. It is the minimum wage required to induce the worker to work, i.e. it is the wage at which the worker is indifferent between working and not working.

Let's think about this graphically, putting L on the x-axis and C on the y-axis. If you try to draw the budget line in this model, you will notice that it is kinked. The maximum C that can be purchased is $wT + y$ (y-intercept). However, the maximum L that can be consumed is T. This costs wT and leaves y dollars that can be used to consume any level of C in $[0, y]$. Therefore, the x-intercept is T and the budget line is vertical between $(T, 0)$ and (T, y). There must be an indifference curve that passes through this kink and its slope at the kink is the reservation wage (denoted as $\underline{w}$). If the wage changes then the left downward sloping side of the budget line will pivot around the kink point (T, y). If the wage is less than the reservation $w < \underline{w}$, then the consumer achieves the the highest utility at the kink point (try drawing the indifference curves). If however $w > \underline{w}$, then the indifferent curve through the kink is no longer tangent to the budget line (again, try drawing this). This means that the agent can be made strictly better off by working a little. Therefore, at $w = \underline{w}$ they will be indifferent between not working and working a little, giving us the concept of the reservation wage. Mathematically, the slope of the indifference curve (in absolute value) is the marginal rate of substitution, so:

$$\underline{w} = \frac{u_L(y, T)}{u_C(y, T)}$$

2.2 Frisch Demand

Before our problem was just in one period, i.e. we chose how much C and L to consume in this period. But maybe a more reaslistic model is that you're actually thinking more long term, e.g. I'll work hard now (low L) and when I'm older I'll relax more (high L). Now we want to choose a *sequence* of consumption and leisure, i.e. $\{(C_t, L_t)\}_{t=1}^{T}$ where T is the number of periods. [2] We call these types of problems as *intertemporal* problems because we are not only making decisions within a period but across time too. This is a problem that will become very familiar to you through your macro classes. Right now you can just think of it like our standard UMP, which means there are two parts to it: utility and budget constraint.

First, utility is now a function of the entire (C_t, L_t) sequence, i.e. $U(C_1, ..., C_T, L_1, ..., L_T)$. The exact form may vary, but we're going to assume the following structure:

$$U(C_1, ..., C_T, L_1, ..., L_T) = \sum_{t=1}^{T} \frac{u(C_t, L_t)}{(1+\rho)^t}$$

Notice the difference between $U(\cdot)$ and $u(\cdot)$. Big $U(\cdot)$ is intertemporal - it depends on the entire sequence of goods. Little $u(\cdot)$ is the utility you get in one period, i.e. how much happiness you get from consuming that period's allocation of C and L. You'll see a few names for $u(\cdot)$, such as the Bernoulli or instantaneous or momentary utility function, but they all capture the idea that within a period your utility from (C, L) isn't changing. It's essentially the same as the $u(\cdot)$ we had in the simple one-period example.

However, we need a way to combine all these Bernoulli utility functions and the way we do it is by discounting. The agent's discount rate, often denoted as $\beta \in [0, 1]$, represents how much they care about the present or how patient they are. $\beta = 1$ means that \$1 tomorrow is worth the same to you today, while $\beta = 0.5$ means that \$1 tomorrow is worth only half its value to you today. If you iterate on this logic, to get today's value of a utility t periods in the future $(u(C_t, L_t))$, you simply need to multiply it by β^t, i.e. $\beta^t u(C_t, L_t)$. In this situation, we are setting $\beta = \frac{1}{1+\rho}$. This has the interpretation

[2]Again, sorry for re-using letters but I think it would be more confusing if I differed from the class notation. Just remember: T is periods not the total number of hours per day.

that ρ is the time preference. You can think of $U(\cdot)$ as being a weighted sum of the future in-period utilities $u(C_t, L_t)$, where the weights represent how far into the future the utility will be experienced.

The second part of a UMP is the budget constraint. Here, just like the utility function, we need to capture the fact there is an intertemporal aspect. We are going to have a few key assumptions:

1. In each period, agents gets a fixed income y_t and a wage w_t
2. The maximum number of hours able to be worked in each period is Θ
3. There is a constant interest rate r (does not change over time)
4. There are perfect financial markets (agents can borrow and save freely across periods)

This is still a budget constraint, which means that the total amount of money you spend has to be less than or equal to the total amount of money you earn. The key difference is that this restriction holds over the *entire* period $t = 1, ..., T$, but not necessarily *within* a single period t (because now you can borrow and save). To be able to make sure we can compare all the dollar amounts, we use the interest rate to make everything into its present discounted value. For example, d dollars saved today gives you $d(1+r)^t$ dollars t periods into the future, which means that d dollars saved t periods into the future is worth $d\frac{1}{(1+r)^t}$ to you today. We use this idea and make all values comparable to period $t = 0$.

On the 'spending' side, each period t the consumer purchases the consumption good C_t (at price 1, normalized just like before) and leisure (at price w_t). To discount this into period 0, we multiply by $\frac{1}{(1+r)^t}$. We add these all up and that gives us the LHS of the budget constraint: $\sum_{t=1}^{T} \frac{C_t + w_t L_t}{(1+r)^t}$. On the 'income' side, as before your total potential income in period t is $y_t + w_t\Theta$. Discounting and summing gives us a RHS of: $\sum_{t=1}^{T} \frac{y_t + w_t\Theta}{(1+r)^t}$. Putting all of this together gives us the intertemporal budget constraint.

So we can finally state the UMP for this intertemporal labor supply model:

$$\max_{\{C_t, L_t\}_{t=1,\ldots,T}} \sum_{t=1}^{T} \frac{u(C_t, L_t)}{(1+\rho)^t}$$
$$\text{s.t. } \sum_{t=1}^{T} \frac{C_t + w_t L_t}{(1+r)^t} \leq \sum_{t=1}^{T} \frac{y_t + w_t\Theta}{(1+r)^t}$$

This might look terrifying but just remember that it's essentially the same UMP we're used to seeing. Notice however that below the max is $\{C_t, L_t\}_{t=1,\ldots,T}$, i.e. we don't choose one (C_t, L_t) bundle, we need to choose a whole sequence of them! Let's now setup the Lagrangian:

$$\mathcal{L} = \sum_{t=1}^{T} \frac{u(C_t, L_t)}{(1+\rho)^t} + \lambda \left(\sum_{t=1}^{T} \frac{y_t + w_t\Theta}{(1+r)^t} - \sum_{t=1}^{T} \frac{C_t + w_t L_t}{(1+r)^t} \right)$$

Before you get scared, no we don't need take $2T$ FOCs to cover every C_t and L_t. Since this is all symmetrical, we can take the FOCs with respect to an arbitrary C_t and L_t. Fortunately this is quite straightforward because each C_t and L_t only appear in period t. This gives us:

$$\frac{\partial \mathcal{L}}{\partial C_t} = \frac{1}{(1+\rho)^t} \frac{\partial u(C_t, L_t)}{\partial C_t} - \lambda \frac{1}{(1+r)^t} = 0$$
$$\frac{\partial \mathcal{L}}{\partial L_t} = \frac{1}{(1+\rho)^t} \frac{\partial u(C_t, L_t)}{\partial L_t} - \lambda \frac{w_t}{(1+r)^t} = 0$$

We usually have nice continuous Bernoulli utility functions so we can invert each of the FOCs and get:

$$\frac{\partial u(C_t, L_t)}{\partial C_t} = \lambda \frac{(1+\rho)^t}{(1+r)^t} \implies C_t = g\left(\lambda \frac{(1+\rho)^t}{(1+r)^t}, L_t\right)$$
$$\frac{\partial u(C_t, L_t)}{\partial Lt} = w_t \lambda \frac{(1+\rho)^t}{(1+r)^t} \implies L_t = l\left(w_t \lambda \frac{(1+\rho)^t}{(1+r)^t}, C_t\right)$$

So if we plug the second in the first, we eliminate L_t and get C_t as a function of the parameters (including w_t) as well as the Lagrange multiplier. Similarly for the other way, and so we get:

$$C_t = \Gamma\left(\lambda \frac{(1+\rho)^t}{(1+r)^t}, w_t\right)$$
$$L_t = \Lambda\left(\lambda \frac{(1+\rho)^t}{(1+r)^t}, w_t\right)$$

We call these the Frisch demands. Let's make some remarks about them. First, these functions hold for all periods t, which makes analysis in these intertemporal problems much easier. Second, if you compare these to our usual Marshallian demands you'll see that a big difference is the presence of the Lagrange multiplier λ. Recall that our interpretation of λ is the marginal utility of wealth. You should also notice that the 'prices' of the other goods do not appear in the Frisch demands - the other goods being consumption and leisure in *other* time periods. What's actually happening is that the relationship between time periods is being entirely captured through the endogenously determined λ.

Normally we think of demand as a mapping from the price of a good to quantity desired. Implicitly though we have to be holding something constant. Note the comparison of the different demands:

- Marshallian demand: how much quantity will I buy as price changes, holding income constant?
- Hicksian demand: how much quantity will I buy as price changes, holding utility constant?
- Frisch demand: how much quantity will I buy as price changes, holding marginal utility of wealth constant?

3 Aggregation

So far we have only been thinking about a single agent making decisions about their consumption and expenditure. More generally, we want to see whether our results will also hold when we have a group of people and "add up" (aggregate) their individual demands. The central question in this section is: When does a group behave as a single individual? There are a few ways we can try to answer this question:[3]

1. When can aggregate demand be expressed as a function of aggregate wealth (and prices)?
2. When do aggregate budget shares look like they come from an individual's demand?

The other question we are going to ask is how do we model group behavior? There are two approaches. The unitary approach says that we can model group (or household) behavior as if there is a single individual with a unique utility function. The other approach is the collective model, which instead treats the household's decision process as a black box. We don't care how the household comes up with

[3] In this section we will do a bit more than is covered in MWG, so the best resource is this paper written by Professor Chiappori himself: http://www.columbia.edu/~pc2167/Aggregation.pdf

a decision, because whatever they decide we assume will be a Pareto efficient outcome. In the third part of this section, we will discuss the collective model.

Let's first establish our notation for this section. There are I people in our group (indexed by i), each of which has their own income y_i, and a Marshallian demand of $x_i(p, y_i)$. Remember that this Marshallian demand is a $K \times 1$ vector, where each element is $x_i^k(p, y_i)$ representing person i's demand for good k.[4] Note that while everyone has different income levels, every agent faces the same price vector p. We will denote the $I \times 1$ vector of incomes as $y = (y_1, ..., y_I)$. Our goal will be try to understand the aggregate demand, which we will denote as:

$$X(p, y) = X(p, y_1, ..., y_I) = \sum_{i=1}^{I} x_i(p, y_i)$$

This is again a $K \times 1$ vector, where each element is $X^k(p, y) = \sum_i x_i^k(p, y_i)$.

3.1 Aggregate Demand and Aggregate Wealth

The first question is the most natural. If everyone's individual demand can be expressed as a function the common prices and their individual demand, then can we just write aggregated demand as a function of the common prices and aggregate demand? In other words, when can we say $X(p, y) = X\left(p, \sum_i y_i\right)$? This seems quite intuitive: if I want to know how much of a good a group will buy and I don't care how much each person buys, then surely all I need to know is how much total money the group has (and not how much each person has). Maybe a less appealing way of saying this is that the *distribution* of wealth has no impact on aggregate demand, only the *total* amount of wealth in the group matters.

This last way of presenting it makes it seem very restrictive, and that's indeed the case. How can we get this 'no distributive effects' result to hold? It would have to be that if we transferred Δ dollars from person A to person B, then person A's consumption of each good will decrease by the same amount that person B's consumption of each good will increase, no matter what their initial wealth level was. We need the effects of the income change to 'cancel out', for all goods at for all wealth levels. In a more mathematical way, suppose we changed each person i's wealth by Δ_i such that $\sum_i \Delta_i = 0$ (aggregate wealth is the same), we would need that the changes in consumption for each good to also sum up to zero:

$$\sum_{i=1}^{I} \underbrace{\frac{\partial x_i^k(p, y_i)}{\partial y_i}}_{\substack{\text{How much person } i \text{ changes their consumption} \\ \text{of good } k \text{ because of a \$1 change in income}}} \underbrace{\Delta_i}_{\substack{\text{How much person} \\ i\text{'s income changes}}} = 0, \forall k = 1, ..., K$$

Since this need to hold for any Δ_i that sum to zero, the only way for this to be true is if for any good k:

$$\frac{\partial x_i^k(p, y_i)}{\partial y_i} = \frac{\partial x_j^k(p, y_j)}{\partial y_j}, \forall i, j = 1, ..., I$$

This would imply for any good k that $\frac{\partial x_i^k(p,y_i)}{\partial y_i} = \phi_k, \forall i$, which means that $\sum_i \frac{\partial x_i^k(p,y_i)}{\partial y_i} \Delta_i = \phi_k \sum_i \Delta_i = \phi_k \times 0 = 0$. But this is exactly what our intuition told us already - for every consumer the wealth effect on each good is exactly the same, regardless of what their initial wealth might be.

[4] In MWG you will see this written as x_{ki}. Also note that we're going to keep using y to represent income as in the labor section because we're going to have yet another use for w later on. Again, I will try to stay consistent with the notation in class otherwise it'll be even more confusing when you do your problem sets and exams

Graphically you can think of this as having linear and parallel Engel curves. Engel curves plot demand for a good (on the x-axis) and income (on the y-axis). We tend to see them sloping upwards because most goods are normal goods and so as income increases, demand for the good will increase too. Moreover, for most goods the Engel curve will be convex (it will bend upwards), meaning that you increasing income by a fixed amount will give you successively smaller increases in demand. For luxury goods, those fixed income increases will instead give you successively larger increases in demand. In the special case we're considering, the Engel curves will be linear because we're assuming this constant wealth effect. The slope of the Engle curve for good k is $\partial y_i / \partial x_i^k = 1/\lambda_k$. Since the slope is $1/\lambda_k$ for all agents, this means that they are all linear (constant slope) and parallel (slope does not depend on any agent-specific variables).

Now we know intuitively what we need but what utility function would give us such a result. The necessary and sufficient condition for this is that we need indirect utility functions that have a Gorman form:

Gorman Form
Aggregate demand can be written as a function of aggregate wealth (exact linear aggregation) if and only if the indirect utility functions have taking the Gorman form:

$$v_i(p, y_i) = a_i(p) + b(p)y_i$$

Where $a_i(p)$ varies by consumer and $b(p)$ (the coefficient on wealth) is the same for all consumers.

We won't prove the necessary part, but to see sufficiency we apply Roy's Identity to the indirect utility:

$$\begin{aligned} x_i^k(p, y_i) &= -\frac{\partial v_i(p, y_i)/\partial p_k}{\partial v_i(p, y_i)/\partial y_i} = -\frac{\frac{\partial a_i(p)}{\partial p_k} + \frac{\partial b(p)}{\partial p_k} y_i}{b(p)} \\ \implies \frac{\partial x_i^k(p, y_i)}{\partial y_i} &= \frac{1}{b(p)} \frac{\partial b(p)}{\partial p_k} = \phi_k \end{aligned}$$

And so you can see that the income effect does not depend on i, which is exactly what we were looking for. Let's go back to the central question of a representative agent. If agents in the group have Gorman form utilities, then we know we can represent their aggregate demand using a representative agent whose income is equal to aggregate income.

$$\begin{aligned} X^k(p, y) &= \sum_{i=1}^{I} x_i^k(p, y_i) \\ &= \sum_{i=1}^{I} \frac{-1}{b(p)} \left(\frac{\partial a_i(p)}{\partial p_k} + \frac{\partial b(p)}{\partial p_k} y_i \right) \\ &= -\frac{1}{b(p)} \sum_{i=1}^{I} \left(\frac{\partial a_i(p)}{\partial p_k} \right) - \frac{1}{b(p)} \frac{\partial b(p)}{\partial p_k} \sum_{i=1}^{I} y_i \\ &= X^k \left(p, \textstyle\sum_i y_i\right) \end{aligned}$$

While this is a clean result, it puts a large restriction on the types of preferences allowed.

3.2 Aggregate Demand and Budget Shares

Another way to think about aggregation is that we want aggregate budget shares to mirror their individual analogues. We will define w_i^k as the share of their budget that agent i spends on good k:

$$w_i^k = \frac{p^k x_i^k}{y_i}$$

Let's define the aggregate analogues. The share of total income spent on good k by all consumers will be denoted as $W^k = \sum_i w_i^k$. As before, the total demand of good k will be $X^k = \sum_i x_i^k(p, y_i)$. By definition, this means that:

$$W^k = \frac{p^k X^k}{\sum_i y_i}$$

Now we want to write a formula for budget share w_i that is only a function of price and income (without demand x_i^k), i.e. we are looking for some $w_i^k(p, y_i)$. The question then becomes can we write the aggregate budget share in a similar fashion: $W_i^k(p, y) = W_i^k(p, Y)$, where $Y = \theta(y)$ is some function of $y = (y_1, ..., y_I)$ that may not necessarily be $\sum_i y_i$. The point here is that using total income was very restrictive, but now we are allowing for a much flexible measure of group income Y that could also capture (for example) the distribution of incomes. Going back to the central question, if we can find such a situation then we can just work with a single representative agent that faces prices p and has income Y. We call this situation, exact non-linear aggregation.

The more general Gorman form for a Marshallian demand function is as follows:

$$x_i^k(p, y_i) = \sum_{j=1}^{J} a_i^{kj}(p) b_i^{kj}(y_i)$$

Where $J \in \{1, 2, 3\}$. One example of this is the PIGLOG form, which we saw in class. This has the following form:

$$x_i^k(p, y_i) = a_i^{k1}(p) \overbrace{y_i}^{b_i^{k1}(y_i)} + \underbrace{a^{k2}(p)}_{\text{Independent of } i} \overbrace{y_i \log y_i}^{b_i^{k2}(y_i)}$$

The budget share for an individual is therefore:

$$w_i^k = \frac{p^k x_i^k(p, y_i)}{y_i} = p^k a_i^{k1}(p) + p^k a^{k2}(p) \log y_i = w_i^k(p, y_i)$$

Now we'll define $a_i^k(p) = p^k a_i^{k1}(p)$ and $b^k(p) = p^k a^{k2}(p)$. This gives us the formulation we saw in class:

$$w_i^k(p, y_i) = a_i^k(p) + b^k(p) \log y_i$$

Note that the difference between this and the Gorman form we saw in the first part is that the common coefficient is on *log* of income and that the dependent variable is now budget share rather than indirect utility. Our goal here is to try to make the analogue version mirror this form, so we want to have $W^k(p, Y) = a^k(p) + b^k(p) \log Y$. Let's check what the aggregated form gives us, noting that, by definition,

$w_i^k y_i = p^k x_i^k$, which means that we can write $p^k X^k = \sum_i p^k x_i^k = \sum_i w_i^k y_i$

$$\begin{aligned} W^k &= \frac{p^k X^k}{\sum_i y_i} \\ &= \frac{\sum_i w_i^k y_i}{\sum_i y_i} \\ &= \frac{\sum_i \left(a_i^k(p) + b^k(p)\log y_i\right) y_i}{\sum_i y_i} \\ &= \underbrace{\frac{\sum_i a_i^k(p) y_i}{\sum_i y_i}}_{\substack{=a^k(p) \textbf{ if} \\ a_i^k(p)=a^k(p),\forall i}} + b^k(p) \underbrace{\frac{\sum_i y_i \log y_i}{\sum_i y_i}}_{=\log Y} \end{aligned}$$

So this gives an *implicit* definition for Y such that $\log Y = \frac{\sum_i y_i \log y_i}{\sum_i y_i}$. As you can see Y is not total income! But since we have this similar form, that means there exists a representative agent and has a utility function of a similar form as the individual agents. Notice that we had to assume that $a_i^k(p) = a^k(p), \forall i$ for this to work (just as you did in your problem set). When we make this assumption, the individual demands for the consumers all become exactly the same, i.e. they all have the same utility and preferences. Let's call their utility function $u(\cdot)$ - now we can use this exact same utility function $u(\cdot)$ for the representative agent. If you solved this agent's UMP using the parameters (p, Y) (i.e. you make their income this non-linear statistic Y) then you would find their that their budget shares would be exactly $W^k(p, Y)$. The identical preferences assumption can be relaxed a little, but you still need them to have very similar preferences. The idea here is that group behaves as a single agent not in terms of demand, but in terms of budget shares.

3.3 Modeling Group Behavior

Now we're going to change gears and think about how to model group behavior. As discussed in the introduction, we are going to be focusing on the collective model. We're not going to be trying to find a representative agent any more. We're going to start with an assumption that the household always comes to an efficient allocation (this will be defined later). From there we want to see whether the standard results still hold, in particular with the Slutsky matrix.

For this section let's have $I = 2$ (imagine this is a household consisting of two people), but this is easy enough to generalize. Household demand is just the sum of individual demand: $x(p, y) = x_1(p, y) + x_2(p, y)$. However, we only observe the total household demand and not the individual demands. Implicitly this means that there is only private consumption: they divide up the goods in some way and then consume their share (but we don't see how they make this division).

Let's model this behavior as if the group has its own UMP to solve, which looks like:

$$\begin{gathered} \max_{(x_1, x_2)} u_1(x_1) \\ \text{s.t. } u_2(x_2) \geq \bar{u}_2 \\ \text{and } p \cdot (x_1 + x_2) \leq y \end{gathered}$$

Now you might think that it is odd that we're solving a group problem by just maximizing person 1's utility. However, it doesn't really matter that we're maximizing person 1 because we have to also make sure that person 2 gets ‘enough’ utility, $\bar{u}_2$. Our goal here is to find a Pareto optimal allocation: some division of the total bundle x such that you can't make one person better off without making someone else worse off. Pareto optimality is usually neither unique nor fair. Take $x_1 = x$ and $x_2 = 0$; this is Pareto optimal but definitely not fair (similarly for the reverse allocation). So our various choices of person 2's threshold $\bar{u}_2$ will trace out all the possible Pareto optimal allocations because we know that in optimality the constraint will be binding, i.e. $u_2(x_2) = \bar{u}_2$. You'll see this formulation of a problem quite a lot when we do general equilibrium in the second half of the semester.

This UMP has two constraints. We can write the Lagrangian, noting that each constraint requires its own Lagrange multiplier:

$$\mathcal{L} = u_1(x_1) + \mu\left(u_2(x_2) - \bar{u}_2\right) + \lambda\left(y - p \cdot (x_1 + x_2)\right)$$

You can think of $u_1(x_1) + \mu u_2(x_2)$ as a weighted total utility, where person 2's weight is μ and person 1's weight is normalized to 1. We can use the multipliers to represent power dynamics in the household where the weights represent decision making power. You'll also see these Lagrange multipliers referred to as *Pareto weights*. Be careful with how you interpret these multipliers though - they are endogenously determined and are a function of prices and income. If μ is constant and not a function of parameters, then the UMP becomes a standard problem and the household behaves like a single agent.

Let's focus on the more interesting case where μ is not constant. Does we still get the usual properties we would expect? Let's think of demand in two ways. The ‘observable’ household demand is $x(p, y)$, which is the analogue to the Marshallian demand. We can also think about one that also depends on the Lagrange multipliers (like the Frisch demand), denoted as $\mathcal{X}(p, y, \mu(p, y))$. Obviously we must have $x(p, y) = \mathcal{X}(p, y, \mu(p, y))$ because they are both the solution to the same problem. $\mathcal{X}$ is not observable because μ is not observable. But, holding μ fixed, it is a standard demand function and will have a symmetric negative semidefinite Slutsky matrix. Since we only observe $x(p, y)$, however, will it also satisfy those properties?

Start by taking the derivative with respect to price:

$$\begin{aligned}
\frac{\partial x(p, y)}{\partial p_k} &= \frac{d\mathcal{X}(p, y, \mu)}{dp_k} && [x(p, y) = \mathcal{X}(p, y, \mu)] \\
&= \frac{\partial \mathcal{X}(p, y, \mu)}{\partial p_k} + \frac{\partial \mathcal{X}(p, y, \mu)}{\partial \mu}\frac{\partial \mu}{\partial p_k} && [\text{total derivative}]
\end{aligned}$$

A similar process for income would give us:

$$\frac{\partial x(p, y)}{\partial y} = \frac{\partial \mathcal{X}(p, y, \mu)}{\partial y} + \frac{\partial \mathcal{X}(p, y, \mu)}{\partial \mu}\frac{\partial \mu}{\partial y}$$

Let's put it together to check the Slutsky matrix of x. Let $S_x[l, k]$ represent the l-k^{th} element of x's Slutsky matrix:

$$\begin{aligned}
S_x[l, k] &= \frac{\partial x_l(p, y)}{\partial p_k} + \frac{\partial x_l(p, y)}{\partial y} x_k(p, y) \\
&= \left[\frac{\partial \mathcal{X}(p, y, \mu)}{\partial p_k} + \frac{\partial \mathcal{X}(p, y, \mu)}{\partial \mu}\frac{\partial \mu}{\partial p_k}\right] + \left[\frac{\partial \mathcal{X}(p, y, \mu)}{\partial y} + \frac{\partial \mathcal{X}(p, y, \mu)}{\partial \mu}\frac{\partial \mu}{\partial y}\right]\mathcal{X}_k(p, y, \mu) \\
&= \underbrace{\left(\frac{\partial \mathcal{X}(p, y, \mu)}{\partial p_k} + \frac{\partial \mathcal{X}(p, y, \mu)}{\partial y}\mathcal{X}(p, y, \mu)\right)}_{S_{\mathcal{X}}[l,k]} + \frac{\partial \mathcal{X}(p, y, \mu)}{\partial \mu}\left(\frac{\partial \mu}{\partial p_k} + \frac{\partial \mu}{\partial y}\mathcal{X}_k(p, y, \mu)\right)
\end{aligned}$$

Where $S_{\mathcal{X}}$ is the Slutsky matrix for $\mathcal{X}$. We know that $S_{\mathcal{X}}$ is symmetric and negative semidefinite, but since S_x is $S_{\mathcal{X}}$ plus 'something else', we can't be sure that S_x is going to be as well. You can read more about this in Section 4 of Professor Chiappori's paper, but I think this general idea will be all that you need to know.

4 Welfare

In this section we want to understand how a change in prices affects consumers. At face value, we know that if the consumer's utility increases than this price change is a 'good' thing. But we would like to go further: how can we quantify how much the agents has gained or lost through this change? How can we compare across different situations when utility doesn't really have any meaningful units?

Let's go over notation again. We will go back to the notation we used in the first consumer theory lecture notes and be consistent with MWG. Assume that an agent has income w, indirect utility $v(p, w)$, and an expenditure function $e(p, u)$. Throughout we will consider a price change from p^0 to p^1, where $p^i = (p_1^i, ..., p_K^i)$.

As we started off by saying, if we observe that $v(p^1, u) > v(p^0, u)$ then we know that the agent is 'happier' with this price change. But what does a gain of $v(p^1, u) - v(p^0, u)$ even mean? One way we can quantify this is to use a *money metric* indirect utility function. These are indirect utility functions whose units can be interpreted as dollars. Remember from our last notes we know that $e(p, v(p, w)) = w$ because w represents exactly the minimum amount of wealth required to achieve utility $v(p, w)$ given prices p. Now we take any arbitrary price vector $p^* \in \mathbb{R}_{++}$ and change the price parameter so that the expenditure function is now: $e(p^*, v(p, w))$. The interpretation is the same as before: this is the minimum amount of wealth required to achieve utility $v(p, w)$ given prices p^*. Note that p^* is fixed so this value is entirely a function of (p, w) via the indirect utility function. Actually because $e(p, u)$ is strictly increasing in u, this is just a monotonic transformation of the indirect utility function - which means that it is an indirect function itself. As a quick verification, let's look at its monotonicity:

- As prices p increase, $v(p, w)$ weakly decreases. Since $e(p, u)$ is strictly increasing in u, then $e(p^*, v(p, w))$ will weakly decrease if $v(p, w)$ weakly decreases. Therefore: $e(p^*, v(p, w))$ is weakly decreasing in p
- As wealth w increases, $v(p, w)$ strictly increases. Since $e(p, u)$ is strictly increasing in u, then $e(p^*, v(p, w))$ will strictly increase if $v(p, w)$ strictly increases. Therefore: $e(p^*, v(p, w))$ is strictly increasing in w

Here's another way to interpret this. Say you live in a world where you have wealth w and face prices p. You optimize and achieve a utility of $v(p, w)$. Now someone comes along and tells you that they can take you to a world where you will face prices p^* and offer you wealth w^*. At what level of w^* will you be indifferent between staying in your current world and moving to the new world? Well that's exactly where $w^* = e(p^*, v(p, w))$, because we know that when you optimize in that new world your utility will be exactly $v(p, w)$ (binding constraint of the EMP). So now we can make the same comparison of indirect utility: $e(p^*, v(p^1, w)) - e(p^*, v(p^0, w))$. The difference now is that this welfare change has a meaningful representation in dollars (income changes).

Note that a key part is that the fixed price p^* provides a reference point for our comparison. in fact our interpretation of the dollar value of $e(p^*, v(p, w))$ depends entirely on p^*. Comparing $e(p^*, v(p^0, w))$ and $e(p^{**}, v(p^1, w))$ for some other price $p^{**} \neq p^*$ wouldn't make any sense. It would be like saying I would be willing to move to Germany for 100 Euros and I would be willing to move to Mexico for 1,000 pesos. If you didn't have an exchange rate to compare the two currencies, how could you possibly tell which one is more?

Our choice of p^* was completely arbitrary but remember that we are considering a change from p^0 to p^1. So what if we let $p^* = p^0$ or $p^* = p^1$? Does this have any meaningful interpretation? Yes! If we do this we get two very useful concepts:

Equivalent Variation

$$EV(p^0, p^1, w) = e(p^0, v(p^1, w)) - e(p^0, v(p^0, w)) = e(p^0, v(p^1, w)) - w$$

Compensating Variation

$$CV(p^0, p^1, w) = e(p^1, v(p^1, w)) - e(p^1, v(p^0, w)) = w - e(p^1, v(p^0, w))$$

Let's interpret this using the exact same logic as when we had an arbitrary p^*. For our story, we're going to imagine the government announcing or enacting a policy, while their rival opposition party is going to propose an alternative policy. You as a voter need to decide what you prefer.

For the *equivalent variation* imagine you're in a world with prices p^0 and income w (this is World 0). The government announces that it is about to change prices to p^1 (this is World 1). You work out that since your income doesn't change, under those prices you will achieve an optimal utility of $v(p^1, w)$. The opposition party then announces that it would instead keep prices at p^0 but there will be an income transfer such that your income will be w^* (this is World 0B). What w^* will make you indifferent between the two proposals: World 1 (p^1, w) and World 0B (p^0, w^*)? Of course, it's when $w^* = e(p^0, v(p^1, w))$. But you know that in the original World 0 (p^0, w) you had income w. So if $w^* > w$, then the price change is equivalent to getting an income boost of $w^* - w > 0$. This amount is the EV and it represents how much the agent is willing to pay to get the new price level. A symmetric argument holds if $w^* < w$ and the agent dislikes the price change so much that you could generate the same welfare effect by taking money away from them. So another way of expressing EV is $v(p^0, w + EV) = v(p^1, w)$. Graphically you can think of this as starting at the pre-change budget line and considering by how much you need to change income so that the budget line becomes tangent with the post-change indifference curve.

For the *compensating variation* imagine you *used to* live in a world with prices p^0 and income w (this is World 0). However, the government has changed the prices to p^1 (this is World 1). Back in the good ol' days, your utility was $v(p^0, w)$ but now it is $v(p^1, w)$. The opposition party then announces that it will get you back to your old utility $v(p^0, w)$ without changing prices, but there will be an income transfer such that your income will be w^* (this is World 1B). What w^* will they need to make their claim true? Of course, it's when $w^* = e(p^1, v(p^0, w))$. But you know that in the current World 1 (p^1, w) you have income w. So if $w^* > w$, then to make you as happy as before, the government would need to be giving you money. Which means that you didn't like the price change and its effect was equivalent to getting an income change of $w - w^* < 0$. This amount is the CV and it represents how much the government would have to pay the agent to bring them back to their old utility. Similarly, you can think of it as the *negative* of the income change that the agent would need. A symmetric argument holds if $w^* < w$ and the agent likes the price change so much that you could create the replicate the welfare change by taking money away from them. So another way of expressing CV is $v(p^0, w) = v(p^1, w - CV)$. Graphically you can think of this as starting at the post-change budget line and considering by how much you need to change income so that the budget line becomes tangent with the pre-change indifference curve

To re-iterate, the point is to be able to quantify the change in consumer welfare between World 0 (p^0, w) and World 1 (p^1, w). One approach is to find a World 0B such that the utility of World 0B is the same as World 1. That means the change from World 0 to World 0B would have the same effect on consumer welfare as the change from World 0 to World 1. Then we can just compare the change in wealth, at the pre-change prices p^0 between World 0B and World 0 (equivalent variation). The other approach is to find a World 1B such that the utility of World 1B is the same as World 0. That means the change from World 1 to World 1B would have the same effect on consumer welfare as the change from World 1 to World 0. Then we can just compare the change in wealth, at the post-change prices p^1 between World 1 and World 1B (compensating variation). Remember that these two wealth changes however are different because we are using starting prices as a reference point. In general, EV and CV are not going to be equal. You can try drawing it graphically to see you why. One special case where they do coincide is for quasilinear utilities.

4.1 Quasilinear Utility

Quasilinear utility takes the form of $u(x) = x_1 + \tilde{u}(x_2, ..., x_K)$, where x_1 is a good called the numeraire. The point of quasilinear utility is the the numeraire enters linearly and in the same way for all consumers, which means that its marginal utility is constant and the same for all consumers. It's easiest to think of the numeraire as money - we all benefit from \$1 in the same way and it provides an easy point of comparison for all other goods. The other special part of quasilinear utility is that there are no wealth effects on the non-numeraire goods $x_2, ..., x_K$. What that means is that if $u(x_1, x_2, ..., x_K) \geq u(y_1, y_2, ..., y_K)$ then $u(x_1 + \Delta, x_2, ..., x_K) \geq u(y_1 + \Delta, y_2, ...y_K)$. For example, if I prefer a bundle of pizza and \$5 to a bundle of cake and \$2, then I'll also prefer pizza and \$10 to cake and \$7.

Let's see why by setting up the UMP. Note that it is usually helpful to normalize to price of the numeraire to 1.

$$\begin{aligned} &\max_{x \in X} u(x) = x_1 + \tilde{u}(x_2, ..., x_K) \\ &\text{s.t. } x_1 + p_2 x_2 + ...p_K x_K \leq w \end{aligned}$$

$$\begin{gathered} \therefore \mathcal{L} = x_1 + \tilde{u}(x_2, ..., x_K) + \lambda \left(w - x_1 - p_2 x_2 - ...p_K x_K\right) \\ \text{FOCs:} \\ 1 - \lambda = 0 \text{ for } k = 1 \implies \lambda = 1 \\ \frac{\partial \tilde{u}(x_2, ..., x_K)}{\partial v_k} - \lambda p_k = 0 \text{ for } k > 1 \implies \frac{\partial \tilde{u}(x_2, ..., x_K)}{\partial v_k} = p_k \end{gathered}$$

In this case we see that the marginal utility of wealth is 1, which means that when we invert the second FOC, we can find the demand for $x_2, ..., x_K$ independent of income. So if we see a price change and the bundle of goods changes, we know with quasilinear utility that the change in demand for a non-numeraire good is entirely because of the substitution effect and not the income effect. Remember that the Marshallian demand captured income and substitution effect while the Hicksian demand only capture the substitution effect. This means that in the quasilinear case, the Marshallian demand and Hicksian Demand for a non-numeraire good are exactly the same. Moreover, quasi-linear utilities have parallel indifference curves (along the numeraire). Using our graphic intuition for EV and CV, since the indifference curves are parallel, we will need the same amount of income to shift between them at any point along the curve.