Lecture 10: Lagrangians (cont'd) and Expenditure Minimization

1 Where are we?

- Last time, we stated the consumer problem,

$$\max u(x) \qquad \text{subject to} \qquad p \cdot x \leq w \qquad \text{and} \qquad x \geq 0$$

  and introduced an auxiliary function, the Lagrangian,

$$\mathcal{L}(x, \lambda, \mu) \quad = \quad u(x) + \lambda(w - p \cdot x) + \mu \cdot x$$

  defined for $x \in \mathbb{R}^k$ and $\lambda, \mu \geq 0$;

- And we showed that if (x^*, λ^*, μ^*) is a saddle point of the Lagrangian –

$$\mathcal{L}(x, \lambda^*, \mu^*) \quad \leq \quad \mathcal{L}(x^*, \lambda^*, \mu^*) \quad \leq \quad \mathcal{L}(x^*, \lambda, \mu)$$

  for every $x \in \mathbb{R}^k$ and every $\lambda, \mu \geq 0$ –

  then x^* is a solution to the Consumer Problem;

  and vice versa if u is differentiable and concave

- Today, we'll see how to actually use this stuff –

  the Kuhn-Tucker Conditions, examples, and onward

- First... any questions?

2 The Kuhn-Tucker Conditions

- So, we know that the solutions to the Consumer Problem pretty much correspond to the saddle points of the Lagrangian
- Next, let's think about what conditions (x^*, λ^*, μ^*) must satisfy to be a saddle point, and therefore for x^* to solve the consumer problem
- First, we need x^* to be a solution to

$$\max_{x \in \mathbb{R}^k} \mathcal{L}(x, \lambda^*, \mu^*) \quad = \quad \max_{x \in \mathbb{R}^k} \{u(x) + \lambda^*(w - p \cdot x) + \mu^* \cdot x\}$$

- We're thinking of $\mathcal{L}$ as defined for all x in $\mathbb{R}^k$, not just $\mathbb{R}^k_+$,

  so x^* must be an interior point;

  if u is differentiable, $\mathcal{L}$ is differentiable w.r.t. x, so we'll need the FOC to hold:

$$\frac{\partial \mathcal{L}}{\partial x_i}(x^*, \lambda^*, \mu^*) \quad = \quad 0 \qquad \longrightarrow \qquad \frac{\partial u}{\partial x_i}(x^*) - \lambda^* p_i + \mu_i^* \quad = \quad 0$$

- (If this was either positive or negative, x^* couldn't be maximizing $\mathcal{L}$)

- And second, we need (λ^*, μ^*) to solve

$$\min_{\lambda, \mu \geq 0} \mathcal{L}(x^*, \lambda, \mu) \quad = \quad \min_{\lambda, \mu \geq 0} \{u(x^*) + \lambda(w - p \cdot x^*) + \mu \cdot x^*\}$$

- This means we need $w - p \cdot x^* \geq 0$ – otherwise, we could always reduce $\mathcal{L}$ by increasing λ – and we need $\lambda^* = 0$ if $w - p \cdot x^* > 0$, otherwise we could reduce $\mathcal{L}$ by reducing λ
- And similarly, we need $x^* \geq 0$ – if $x_i^* < 0$, we could reduce $\mathcal{L}$ by increasing μ_i – and we need $\mu_i^* = 0$ if $x_i^* > 0$, or else we could reduce $\mathcal{L}$ by reducing μ_i
- So basically, for (λ^*, μ^*) to minimize $\mathcal{L}$,

  we need for x^* to be feasible, and we need complementary slackness to hold

- And of course, we defined $\mathcal{L}$ for non-negative values of the Lagrange multipliers, so we can think of $\lambda \geq 0$ and $\mu \geq 0$ as additional constraints
- That gives us a bunch of necessary conditions for (x^*, λ^*, μ^*) to be a saddle point:
  1. $\frac{\partial u}{\partial x_i}(x^*) - \lambda^* p_i + \mu_i^* = 0$ for each i (FOC w.r.t. x)
  2. $p \cdot x^* \leq w$, $x_i^* \geq 0$ for each i (original constraints on x)
  3. $\lambda^* \geq 0$ and $\mu_i^* \geq 0$ for each i (non-negative Lagrange multipliers)
  4. $\lambda^*(w - p \cdot x^*) = 0$ and $\mu_i^* x_i^* = 0$ for each i (complementary slackness)
- These are collectively known as the **Kuhn-Tucker Conditions**
- As expected, these are necessary conditions for a solution to the consumer problem, and also sufficient under some additional assumptions:

  Theorem 1. *Assume u is differentiable and $(p, w) \gg 0$. If $x^* > 0$ solves the Consumer Problem, then there exist $\lambda^* \geq 0$ and $\mu^* \geq 0$ such that (x^*, λ^*, μ^*) satisfy the Kuhn-Tucker conditions.*[1]

  Theorem 2. *Suppose u is differentiable and concave.*[2] *Then if (x^*, λ^*, μ^*) satisfies the Kuhn-Tucker conditions, x^* solves the Consumer Problem.*

- Putting these results together with what we saw last time,
  if we let CP mean "x^* solves the consumer problem,"
  $\mathcal{L}$ mean "(x^*, λ^*, μ^*) is a saddlepoint of $\mathcal{L}$ for some λ^* and μ^*,
  and KT mean "(x^*, λ^*, μ^*) satisfy the Kuhn-Tucker conditions for some λ^* and μ^*,"
  we have that if $(p, w) \gg 0$...
  - $\mathcal{L} \rightarrow CP$ with no further restrictions
  - $CP \rightarrow \mathcal{L}$ if u is differentiable and concave
  - $CP \rightarrow KT$ if u is differentiable
  - $KT \rightarrow CP$ if u is differentiable and concave

[1] For a more general constrained optimization problem, there is one additional restriction which would have to hold at x^*, in order to guarantee the Kuhn-Tucker conditions can be satisfied. This is called *constraint qualification*, and it basically says that whichever of the constraints of the problem bind at x^*, their gradients need to be linearly independent, i.e., there can't be any locally "redundant" constraints. Given the constraints in our problem (budget and non-negativity), this condition automatically holds whenever $p \gg 0$ and $w > 0$, so I don't state it in the theorem; this is just something you need to know if you want to extend these techniques to other problems. See MWG page 956-964 for the more general treatment of constrained optimization.

[2] We can weaken concavity to (i) quasi-concavity and (ii) for any $x, x' \in X$, $u(x') > u(x)$ implies $\nabla u(x) \cdot (x' - x) > 0$, both of which hold when u is concave.

3 What's this good for?

- So, if u is differentiable,

  the Kuhn-Tucker conditions will find all the solutions to the Consumer Problem,

  although if u is not concave, it may find some points that aren't solutions too

- But, even if u is not concave

  (so the KT conditions aren't automatically sufficient),

  the KT conditions will often have a unique solution;

  since the CP must have a solution, and it must satisfy KT,

  if the KT conditions have a unique solution, we know it's the unique solution to the CP

- Let's use this to solve one example

- Let $X = \mathbb{R}^3_+$, $u(x_1, x_2, x_3) = x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3}$, a generic price vector $(p_1, p_2, p_3) \gg 0$, and $w > 0$

- First, it's helpful to try to eliminate some constraints

- Note that if any consumption quantity is 0, then $u = 0$, so this can't be optimal –

  so $x^* \gg 0$, which means (for complementary slackness) $\mu^* = 0$,

  so we can ignore the non-negativity constraints

- (For many utility functions,

  the marginal utility of each good is infinite when you're consuming none of it –

  the first little bit of each good is extremely valuable –

  and in this case, it's never optimal to consume 0 of anything given positive wealth,

  and you can just solve the problem without the nonnegativity constraints)

- Next, rather than $u = x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3}$, let's maximize $\ln u$, since this is another utility function representing the same preferences, so we'll use

$$\tilde{u}(x, y, z) \quad = \quad \alpha_1 \log x_1 + \alpha_2 \log x_2 + \alpha_3 \log x_3$$

- The Lagrangian is therefore

$$\mathcal{L} \quad = \quad \alpha_1 \log x_1 + \alpha_2 \log x_2 + \alpha_3 \log x_3 + \lambda(w - p \cdot x)$$

  and the Kuhn-Tucker FOC are therefore

$$\frac{\partial \mathcal{L}}{\partial x_i} \quad = \quad 0 \qquad \longrightarrow \qquad \frac{\alpha_i}{x_i} - \lambda p_i \quad = \quad 0 \qquad \to \qquad x_i \quad = \quad \frac{\alpha_i}{\lambda p_i}$$

  for each good i

- Next, since u is strictly increasing everywhere, the budget constraint will hold with equality
- (We haven't yet proved Walras' Law – that's what I get for changing the order –

  but since u is strictly increasing in every good, the consumer would never choose not to spend the whole budget[3]

- So

$$w \quad = \quad \sum_{j=1}^{3} p_j x_j \quad = \quad \sum_{j=1}^{3} p_j \frac{\alpha_j}{\lambda p_j} \quad = \quad \frac{\alpha_1 + \alpha_2 + \alpha_3}{\lambda}$$

  so $\lambda = \frac{\alpha_1+\alpha_2+\alpha_3}{w}$, and therefore

$$x_i \quad = \quad \frac{\alpha_i}{\lambda p_i} \quad = \quad \frac{\alpha_i}{\alpha_1 + \alpha_2 + \alpha_3} \frac{w}{p_i}$$

  or

$$p_i x_i \quad = \quad \frac{\alpha_i}{\alpha_1 + \alpha_2 + \alpha_3} w$$

- Since this is the only way to satsify the K-T conditions,

  we know this is the (unique) solution to the consumer problem!

- (also note that we didn't need concavity of u –

  we never even specified that $\alpha_1 + \alpha_2 + \alpha_3 \leq 1$!)

[3]Or to be more formal, recall that for any j with $x_j^* > 0$, $\lambda^* = \frac{1}{p_j} \frac{\partial u}{\partial x_j}(x^*) > 0$; and once we know $\lambda^* \neq 0$, then for complimentary slackness to hold, $w - p \cdot x^* = 0$.

3.1 one other thing about Lagrangians

- Now, (x^*, λ^*, μ^*) being a saddle point of $\mathcal{L}$ is really two separate conditions:
  - given λ^* and μ^*, x^* maximizes $\mathcal{L}$
  - and given x^*, the multipliers (λ^*, μ^*) minimize $\mathcal{L}$
- And we've seen that the latter condition is really equivalent to two things:
  feasibility of x^*, and complementary slackness
- So when I learned this stuff in grad school, I didn't learn it as saddle points,
  I learned it as, you solve a constrained optimization problem by solving

$$\max_x \{u(x) + \lambda(w - p \cdot x) + \mu \cdot x\}$$

  and imposing the additional conditions $p \cdot x \leq w$, $x \geq 0$, $\lambda(w - p \cdot x) = 0$, and $\mu \cdot x = 0$
- As we saw, the FOC from maximization over x,
  combined with these other conditions (and non-negativity of the Lagrange multipliers),
  are exactly the Kuhn-Tucker conditions

3.2 end of Lagrangians

- Before we move on, I'll repeat two cautionary points I made earlier

- First of all, while this is a great way to solve the consumer problem when u is differentiable,
  please don't think of it as a cookbook,
  where you get a problem and immediately plug it into the Kuhn-Tucker conditions without thinking about it
- There's a lot to gain by thinking a little about the problem before you start to solve it – whether to expect a corner solution, an interior solution, and other things like that

- And finally, if you use this for other constrained optimization problems beyond CP,
  there's one additional restriction that has to hold – called "constraint qualification" –
  which we didn't go into.[4]

[4] Roughly, constraint qualification requires that all the constraints that bind at the same point have gradients at that point that are not linear combinations of each other – in some sense, there are no locally "redundant" constraints. With the Consumer Problem, this holds automatically, since the gradient of $w - p \cdot x$ is $(-p_1, -p_2, \ldots, -p_k)$ and the gradient of x_i is $(0, \ldots, 0, 1, 0, \ldots, 0)$. But if you were maximizing a function subject to the constraints $x_1 \geq 0$ and $-x_1 + (x_2 - 3)^2 \geq 0$, you'd be in trouble, because at the point $(0, 3)$, the two constraints both bind and have gradients $(1, 0)$ and $(-1, 0)$. See MWG section M.K in the math appendix for more detail.

4 jump back – a few properties of demand

- We defined $x(p, w)$ as the solution to the consumer problem, or Marshallian demand

- **Proposition.** Marshallian demand is homogeneous of degree zero: $x(\lambda p, \lambda w) = x(p, w)$ for any $p \gg 0$, $w > 0$, $\lambda > 0$.
  - Proof: we already showed $B(\lambda p, \lambda w) = B(p, w)$, so $x(\lambda p, \lambda w)$ and $x(p, w)$ are literally the solutions to the same optimization problem.

- Again, the problem is invariant to the unit of currency

- **Proposition [Walras' Law].** If preferences are locally non-satiated, then for any (p, w) and any $x \in x(p, w)$, $p \cdot x = w$.
  - Suppose not, i.e., suppose $p \cdot x < w$ for some $x \in x(p, w)$
  - If we make ε small enough, then for every y within ε of x, $p \cdot y \leq w$
    * specifically, let $\varepsilon = (w - p \cdot x)/\|p\|$
    * then
    $$p \cdot y \quad = \quad p \cdot (x + (y - x)) \quad = \quad p \cdot x + p \cdot (y - x)$$
    * since $a \cdot b = \|a\|\|b\| \cos\theta \leq \|a\|\|b\|$ (where θ is the angle between vectors a and b), this is
    $$\leq \quad p \cdot x + \|p\|\|y - x\| \quad \leq \quad p \cdot x + \|p\| \frac{w - p \cdot x}{\|p\|} \quad = \quad w$$
    so if $\|y - x\| \leq \varepsilon$, $p \cdot y \leq w$
  - but if preferences are LNS, then there must be some y within ε of x such that $y \succ x$
  - since we just showed that y was affordable, this contradicts x being the maximizer

- so if preferences are LNS, optimizing consumers always spend their entire budget;

  so if prefs are LNS, we can rewrite the consumer problem as

$$\max u(x) \quad s.t. \quad p \cdot x = w$$

- We said that $x(p,w)$ could be a set, i.e., the consumer problem could have multiple solutions
- **Proposition.**
  1. If preferences are convex, then for each p and w, $x(p,w)$ is a convex set
  2. If preferences are strictly convex, then for each p and w, $x(p,w)$ is a single point, i.e., x is a function.

  - Proof. $x(p,w)$ is a convex set if for any x and $x' \in x(p,w)$, $tx + (1-t)x' \in x(p,w)$ as well. So let's show it.
  - If x and x' are both in $B(p,w)$, then $p \cdot x \leq w$ and $p \cdot x' \leq w$, so

    $$p \cdot (tx + (1-t)x') \quad = \quad t(p \cdot x) + (1-t)(p \cdot x') \quad \leq \quad tw + (1-t)w \quad = \quad w$$

    so $tx + (1-t)x'$ is in the budget set as well
  - And if x and x' are in $x(p,w)$, then for any other y in the budget set, $x \succsim y$ and $x' \succsim y$; if preferences are convex, this means $tx+(1-t)x' \succsim y$ as well, so $tx+(1-t)x' \in x(p,w)$

  - for the second part, if preferences are strictly convex, suppose there were two separate points x and x' in $x(p,w)$
  - since $x \succsim x'$ and $x' \succsim x'$, if preferences are strictly convex, for any $t \in (0,1)$, $tx + (1-t)x' \succ x'$
  - as we argued above, $tx+(1-t)x'$ is in the budget set, so this contradicts x' being optimal
- so now we know that Marshallian demand...
  - exists
  - is homogeneous of degree 0
  - exhausts the budget if preferences are LNS
  - is single-valued if preferences are strictly convex
- onward!

5 The Indirect Utility Function

- We stated the consumer problem as

$$\max u(x) \quad \text{subject to} \quad x \in B(p,w)$$

and called the solution $x(p,w)$; now define $v(p,w)$ as the value function, i.e.,

$$v(p,w) \quad = \quad \max u(x) \quad \text{subject to} \quad x \in B(p,w)$$

- by definition, $v(p,w) = u(x)$ for any $x \in x(p,w)$ –

  v is the "utility value" of having wealth w and facing prices p

- note that $v : \mathbb{R}^k_{++} \times \mathbb{R}_+ \to \mathbb{R}$ is a function:

  if the CP has multiple solutions, they all give the same utility, so there's a unique max value

- v is called the *indirect utility function* (because it's a function of p and w, not consumption)

- Some properties of v:

- **Proposition.** Suppose u is a continuous utility function representing a LNS preference relation $\succsim$ on $\mathbb{R}^k_+$. Then $v(p,w)$ is...

  1. homogeneous of degree 0: $v(\lambda p, \lambda w) = v(p,w)$
  2. continuous on $\{(p,w) : p \gg 0, w \geq 0\}$
  3. nonincreasing in p, and strictly increasing in w
  4. quasi-convex, i.e., the set $\{(p,w) : v(p,w) \leq \bar{v}\}$ is a convex set for each $\bar{v}$

- I'll skip the formal proofs of most parts.

  – Part 1 is the usual – if you scale up prices and wealth, B is unchanged, so you're solving the same optimization problem.

  – Part 2, I'll skip the formal proof, but the intuition is this. If (p,w) is close to (p',w'), then the budget sets are close to being the same set; so if u is continuous, the utility you can achieve on one set is close to the utility you can achieve on the other.

  – Part 3, if p increases, the budget set shrinks, so you're maximizing the same function over a smaller set, so the value goes weakly down. If w goes up, then relative to what used to be your optimal point, you can afford strictly more, and can therefore afford any nearby bundle; by LNS, there's a nearby bundle that's strictly better, so you can afford at least one bundle better than your old optimum.

- Skipped in class, but here's the formal proof for part 4, that v is quasi-convex:
  - We need to show that for any $\bar{v}$, the set $\{(p, w) : v(p, w) \leq \bar{v}\}$ is a convex set
  - Or, if (p, w) and (p', w') are both in that set,
    so is $t(p, w) + (1 - t)(p', w')$, which we'll abbreviate (p^t, w^t)
  - so we need to show that if $v(p, w) \leq \bar{v}$ and $v(p', w') \leq \bar{v}$, then $v(p^t, w^t) \leq \bar{v}$ as well

  - To show this, note that for any $x \in B(p^t, w^t)$,
    x must also be in at least one of $B(p, w)$ or $B(p', w')$
  - If it weren't, then we'd know that both $p \cdot x > w$ and $p' \cdot x > w'$, which would mean

$$t(p \cdot x) + (1 - t)(p' \cdot x) > tw + (1 - t)w' \quad \longrightarrow \quad p^t \cdot x > w^t$$

    which would give a contradiction
  - So if $v(p, w)$ and $v(p', w')$ are both below $\bar{v}$,
    then any bundle with $u(x) > \bar{v}$ must be outside of both $B(p, w)$ and $B(p', w')$
    which we just said means it's outside of $B(p^t, w^t)$
  - so there's no way to achieve utility more than $\bar{v}$ in $B(p^t, w^t)$, so $v(p^t, w^t) \leq \bar{v}$.

6 Using the Lagrangian to prove one more cool property of v

- We noted last time that a saddle point is the solution to the min-max problem

$$\min_{\lambda,\mu\geq 0}\max_x \mathcal{L}(x,\lambda,\mu)$$

- Further, since we know complementary slackness holds at a saddle point,

  the value of that problem is

$$\min_{\lambda,\mu}\max_x \mathcal{L}(x,\lambda,\mu) \quad = \quad \mathcal{L}(x^*,\lambda^*,\mu^*) \quad = \quad u(x^*)+\lambda^*(w-p\cdot x^*)+\mu^*\cdot x^* \quad = \quad u(x^*) \quad = \quad v(p,w)$$

  since x^* is the solution to the consume rproblem

- So we can write

$$v(p,w) \quad = \quad \min_{\lambda,\mu\geq 0}\max_x \mathcal{L}(x,\lambda,\mu)$$

- Which gives us an elegant way to prove a cool result:

  Theorem (Roy's Identity). *Suppose v is differentiable at $(p,w) \gg 0$, and $\partial v/\partial w > 0$. Then $x(p,w)$ is a singleton and*

$$x_i(p,w) \quad = \quad -\frac{\partial v}{\partial p_i}(p,w) \Bigg/ \frac{\partial v}{\partial w}(p,w)$$

- Originally, we defined $v(p,w)$ as $\max u(x)$ subject to $p \cdot x \leq w$ and $x \geq 0$,

  which is a constrained problem – and in particular,

  a change in a parameter p or w changes the constraints, not the objective function,

  so it was hard to see how changes in parameters would change the outcome

- But now, the Lagrangian lets us move the constraints into the objective function,

  which gives us a way to apply the Envelope Theorem to a constrained problem

- Starting with

$$v(p,w) = \min_{\lambda,\mu\geq 0}\max_x \{u(x)+\lambda(w-p\cdot x)+\mu\cdot x\}$$

and define

$$\Phi(\lambda,\mu,p,w) = \max_x \{u(x)+\lambda(w-p\cdot x)+\mu\cdot x\}$$

as the result of the inner maximization problem, so that

$$v(p,w) = \min_{\lambda,\mu\geq 0}\Phi(\lambda,\mu,p,w)$$

- By the envelope theorem,

$$\frac{\partial v}{\partial w} = \left.\frac{\partial \Phi}{\partial w}\right|_{\lambda=\lambda^*,\mu=\mu^*}$$

- And applying the envelope theorem again,

$$\frac{\partial \Phi}{\partial w} = \left.\frac{\partial \mathcal{L}}{\partial w}\right|_{x=x^*}$$

- So now

$$\frac{\partial v}{\partial w} = \left.\frac{\partial}{\partial w}\left(u(x)+\lambda(w-p\cdot x)+\sum_i \mu_i x_i\right)\right|_{\lambda=\lambda^*,\mu=\mu^*,x=x^*} = \lambda|_{\lambda=\lambda^*} = \lambda^*$$

- This is interesting in its own right – if (x^*,λ^*,μ^*) is a saddlepoint, $dv/dw=\lambda^*$ –

  the value of the multiplier on the budget constraint, is the marginal value of wealth!

  (That is, the Lagrange multiplier tells you the marginal benefit of relaxing that constraint!)

- (This makes sense – for any good where $x_i^*>0$, the KT FOC is $\lambda=\frac{1}{p_i}\frac{\partial u}{\partial x_i}$,

  so at a saddlepoint, λ is the marginal utility you'd get from spending an incremental amount of money on good i

- If you're already at the optimum, this is equalized across all goods,

  and so a marginal increase in w gives you this much additional utility

- But it's nice that the envelope theorem let us show this formally)

- Then we can do the same thing differentiating with respect to price p_i, and find

$$\frac{\partial v}{\partial p_i} = \frac{\partial}{\partial p_i}\left(u(x) + \lambda(w - p \cdot x) + \sum_i \mu_i x_i\right)\Bigg|_{\lambda=\lambda^*,\mu=\mu^*,x=x^*} = -\lambda x_i\big|_{\lambda=\lambda^*,x=x^*} = -\lambda^* x_i^*$$

- (This also makes sense – if p_i goes up incrementally,

  then before you adjust your consumption,

  you're over budget by an amount proportional to x_i^*,

  and scaling down consumption to get back under budget costs you that amount times λ^*)

- So

$$-\frac{\partial v/\partial p_i}{\partial v/\partial w} = -\frac{-\lambda^* x_i^*}{\lambda^*} = x_i^*$$

  which is Roy's Identity

7 The Expenditure Minimization Problem

- Next, we're going to consider a different formulation of the consumer problem, look at its properties, and then look at the link between the two problems
- In producer theory, we considered the general problem of profit maximization, and we also considered the more specialized problem of cost minimization
- Here, we're going to do something very analogous

- We defined the consumer problem

$$\max_{x \geq 0} u(x) \quad \text{subject to} \quad p \cdot x \leq w$$

with $v(p, w)$ its value and $x(p, w)$ (Marshallian demand) its solution

- The Expenditure Minimization Problem – or dual consumer problem – is defined as

$$\min_{x \geq 0} p \cdot x \quad \text{subject to} \quad u(x) \geq u$$

- That is, rather than asking how much utility we can buy for wealth w, we ask how cheaply we can achieve a target utility level u

- We call $e(p, u)$ the min – the smallest amount of money that can achieve utility u – this is the *expenditure function*

  And $h(p, u)$, the solution to the problem,

  we'll call Hicksian demand, or compensated demand

7.1 Why is this problem interesting?

- Our original consumer problem seems like a reasonable proxy for how people might actually make decisions
- The expenditure minimization problem seems more contrived –
  less related to a problem people really solve
- So why do we like it?

- Three reasons
- First, Marshallian Demand faces two effects when a price changes –
  substitution and wealth effects
  (If p_1 goes down, you might buy less of another good and more of good 1
  But in addition, if p_1 goes down, this makes you effectively richer –
  which can change your demand in any direction)
  Hicksian demand eliminates the wealth effect –
  so it gives a way to kind of decompose price effects into their separate components

- Second, because Hicksian demand doesn't have wealth effects,
  we can get stronger, testable predictions
  And importantly, we can infer Hicksian demand from Marshallian demand –
  so if we observe data from the regular consumer problem,
  this gives us a way to test the predictions of the model

- Third, we can't really do welfare analysis using indirect utility functions,
  because utility has no natural scale and people aren't comparable
  But the Expenditure Function is in dollars –
  we can measure welfare changes by how much money it would take to compensate people,
  and then aggregate across people in a more valid way

- (It also creates a nice parallel with what we did in producer theory)