Lecture 6: MCS II, LeChatelier

1 Where we are

- Last week, we proved the one-dimensional analog of Topkis' Theorem,

  and then set up the multi-dimensional version and gave the result.

- Let $g : X \times T \rightarrow \mathbb{R}$, with $X \subseteq \mathbb{R}^m$ a product set and $T \subseteq \mathbb{R}^n$,

  and let $x^*(t) = \arg\max_{x \in X} g(x,t)$.

- Suppose g is supermodular in X, meaning $g(x \vee x') + g(x \wedge x') \geq g(x) + g(x')$

  ($x \vee x'$ ("x meet x'") is the componentwise max $(\max\{x_1, x_1'\}, \max\{x_2, x_2'\}, \ldots)$,

  $x \wedge x'$ ("x join x'") is the componentwise min $(\min\{x_1, x_1'\}, \min\{x_2, x_2'\}, \ldots)$)

- Or equivalently, g is SPM if it has increasing differences in x_i and x_j for each pair (i, j),

  holding the other x_k and t constant.

- And suppose g has increasing differences in (X, T),

  meaning increasing differences in x_i and t_j for each i and j,

  holding the other variables constant.

- Then x^* is increasing in t

- When x^* is single-valued, this is in the usual sense –

  each component x_i^* of the solution is weakly increasing in each parameter t_j

- When x^* is not single-valued, this means that when $t' \geq t$,

  if $x \in x^*(t)$ and $x' \in x^*(t')$,

  then $x \vee x' \in x^*(t')$ ("x meet x'") and $x \wedge x' \in x^*(t)$ ("x join x'")

- Today, I want to build a little more intuition for Topkis' Theorem,
  mention some related empirical work,
  and hopefully wrap up with a cool application

- **Before we start... any questions?**

2 Let's go back to two-input production

- We motivated Topkis' Theorem with a single-output firm that uses capital and labor,
  whose profit maximization problem is

$$\max_{k,\ell} \{pf(k,\ell) - w\ell - rk\}$$

- We started with the case where $\frac{\partial^2 f}{\partial k \partial \ell} \geq 0$

- What I typically do to solve problems like this
  is look at the partial derivative of the objective function with respect to each choice variable

- In this case,

$$\frac{\partial g}{\partial k} \quad = \quad p\frac{\partial f}{\partial k} - r$$

  is increasing in ℓ, increasing in p, and increasing in $-r$

$$\frac{\partial g}{\partial \ell} \quad = \quad p\frac{\partial f}{\partial \ell} - w$$

  is increasing in k, increasing in p, and increasing in $-w$

- So we can say that g is supermodular – it has increasing differences in k and ℓ –
  and it has increasing differences in the choice variables (k, ℓ) and the parameters $(p, -r, -w)$

- And that's what we need for Topkis to give us the result –
  both capital and labor go up, at least in the meet-join sense,
  if p rises, r falls, or w falls

- Next. we'll continue with the same example, but when $\frac{\partial^2 f}{\partial k \partial \ell} < 0$

- Topkis' Theorem only applies when the objective function is supermodular, and we know $\frac{\partial^2 g}{\partial k \partial \ell} = p \frac{\partial^2 f}{\partial k \partial \ell}$, so what do we do when this is negative?

- In the one dimensional case, when we had an interaction between x and a parameter that went in the "wrong direction," we just flipped the sign of the parameter

- Here, we can do the same thing – we can flip the sign of one of the choice variables

- Formally, think of defining a new variable $\hat{k} = -k$, and think of the firm's problem as

$$\max_{\ell \geq 0, \hat{k} \leq 0} \{ p f(-\hat{k}, \ell) - w\ell + r\hat{k} \}$$

- Now,

$$\frac{\partial^2 g}{\partial \ell \partial \hat{k}} \quad = \quad -p \frac{\partial^2 f}{\partial \ell \partial k} \quad \geq \quad 0$$

  so if we think of the firm as choosing $(\ell, \hat{k}) = (\ell, -k)$, the problem is now supermodular

- What about increasing differences in X and T?
- Well, $\frac{\partial g}{\partial \ell} = p\frac{\partial f}{\partial \ell}(-\hat{k}, \ell) - w$ is increasing in p and $-w$
- And $\frac{\partial g}{\partial \hat{k}} = -p\frac{\partial f}{\partial k} + r$ is *decreasing* in p and increasing in r
- So if we need to flip the sign of k to make the problem supermodular,
  we can't get increasing differences in both choice variables and p
- So what can we do?
- Ignore p!
- If we still think of the firm as choosing ℓ and $-k$,
  the objective function has increasing differences in $X = (\ell, -k)$ and $T = (-w, r)$,
  and we just can't say anything about p

- So here, Topkis' Theorem says that if r goes up,
  then ℓ and $-k$ both go up;
  or if the price of capital goes up,
  the firm uses less capital but more labor
- (Since $\frac{\partial f}{\partial \ell}$ is decreasing in k,
  when the price of labor goes up and the firm uses less capital,
  that increases the productivity of labor, so the firm hires more;
  since $\frac{\partial f}{\partial k}$ is decreasing in ℓ,
  hiring more labor depresses the productivity of capital,
  so the firm demands even less capital,
  and so on)
- And if the wage w goes up, then $-w$ goes down,
  so ℓ and $-k$ go down,
  so the firm uses less labor and more capital

- Of course, $\frac{\partial^2 f}{\partial k \partial \ell} < 0$ is basically saying that when k goes down,
  the marginal product of labor goes up –
  capital and labor are substitutes in production
- So when capital gets more expensive,
  the firm uses less capital,
  but that makes labor more productive,
  so then the firm uses more labor
- (This in turn makes capital less productive,
  reinforcing the firm's choice to use less capital –
  all the "feedback loops" keep pushing in the same direction)

- What about p?
  Well, when p goes up, we know the firm will want to produce more;
  but without knowing more about f,
  we can't tell whether it will use more capital and less labor,
  more labor and less capital,
  or more of both
- So we can just ignore p

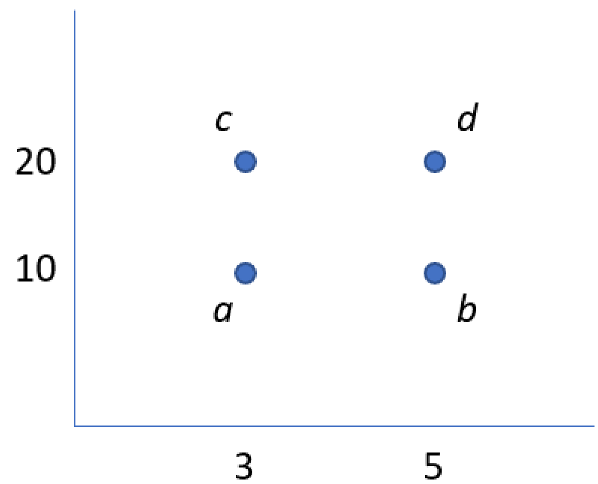
- **Point of emphasis:** we're free to ignore any *parameter* we want
  (i.e., we can look at comparative statics for r and w without considering p at all);
  we **cannot** "ignore" any of the firm's choice variables
- (We'll see what this means in a bit)

3 Why Is Topkis' Theorem True?

- I'm not going to give the formal proof –

  I'll post a video of it as "bonus material" –

  but I do want to give some very very quick intuition for why it holds



- The result is that as a parameter t increases, the set of solutions to a problem increases, provided the problem is supermodular and has ID in the choice variables and t

- So let's consider a problem in two dimensions,

$$\max_{x,y} g(x, y, t)$$

  where g is supermodular in $X = (x, y)$ and has increasing differences in (x, y) and t

- And let's consider $t \in \{0, 1\}$

- What Topkis says here is that if a point (x, y) is optimal at $t = 0$,

  and a point (x', y') is optimal at $t = 1$,

  then the point $(\max\{x, x'\}, \max\{y, y'\})$ must also be optimal at $t = 1$,

  and point $(\min\{x, x'\}, \min\{y, y'\})$ must also be optimal at $t = 0$

- There are a few possible cases
- **Case 1:** a is optimal at $t = 0$ and d is optimal at $t = 1$
- In that case, we need to show that $a \vee d = d$ is optimal at $t = 1$,
  and $a \wedge d = a$ is optimal at $t = 0$,
  so there's nothing to show – we're done

- **Case 2:** b is optimal at $t = 0$ and c is optimal at $t = 1$
- Here, we need to show that $d = b \vee c$ is also optimal at $t = 1$,
  and $a = b \wedge c$ is also optimal at $t = 0$
- Let's do the first
- If $b = (5, 10)$ is optimal at $t = 0$, this means b is at least as good at a at $t = 0$, or

$$g(5, 10, 0) \quad \geq \quad g(3, 10, 0)$$

  or

$$g(5, 10, 0) - g(3, 10, 0) \quad \geq \quad 0$$

- But supermodularity implies $g(5, y, t) - g(3, y, t)$ is increasing in t, giving

$$g(5, 20, 0) - g(3, 20, 0) \quad \geq \quad g(5, 10, 0) - g(3, 10, 0) \quad \geq \quad 0$$

  and increasing differences implies $g(5, y, t) - g(3, y, t)$ is increasing in t, giving

$$g(5, 20, 1) - g(3, 20, 1) \quad \geq \quad g(5, 20, 0) - g(3, 20, 0) \quad \geq \quad 0$$

- But we started with $c = (3, 20)$ is optimal at $t = 1$; so if $d = (5, 20)$ is at least as good, then d must also be optimal at $t = 1$, which is what we wanted to show
- (We could use analogous steps to show a optimal at $t = 0$)

- By symmetry, it should be obvious that if c were optimal at $t = 0$ and b optimal at $t = 1$, we'd get the same result; so all that's left is...

- **Case 3:** a is optimal at $t = 1$ and d is optimal at $t = 0$
- In that case, we need to show d is also optimal at $t = 1$, and a at $t = 0$
- Again, we'll show the first
- If d is optimal at $t = 0$, it's at least as good as a, so

$$g(5, 20, 0) - g(3, 10, 0) \quad \geq \quad 0$$

- Let's add and subtract $g(3, 20, 0)$, giving

$$[g(5, 20, 0) - g(3, 20, 0)] + [g(3, 20, 0) - g(3, 10, 0)] \quad \geq \quad 0$$

- By increasing differences, $g(5, y, t) - g(3, y, t)$ is increasing in t, and $g(x, 20, t) - g(x, 10, t)$ is increasing in t, so

$$\begin{array}{c} [g(5, 20, 1) - g(3, 20, 1)] + [g(3, 20, 1) - g(3, 10, 1)] \\ \geq \\ [g(5, 20, 0) - g(3, 20, 0)] + [g(3, 20, 0) - g(3, 10, 0)] \\ \geq \\ 0 \end{array}$$

  or

$$g(5, 20, 1) - g(3, 10, 1) \quad \geq \quad 0$$

- But since $a = (3, 10)$ is optimal at $t = 1$, and $d = (5, 20)$ is at least as good, d is also optimal at $t = 1$, which is what we wanted to show
- (Again, analogous arguments show a optimal at $t = 0$)

- So that's the result
- if (x, y) is optimal at t and (x', y') is optimal at higher t', then their meet is optimal at t' and their join is optimal at t
- Or, if the problem has a unique solution at each value of t, then that solution is weakly increasing in every dimension in the parameters

4 Moar Examples!

4.1 Cobb Douglas production

- Let's consider a couple of well-known production functions
- First, think of Cobb-Douglas, perhaps with three inputs

$$f(z) = z_1^\alpha z_2^\beta z_3^\gamma$$

  with $\alpha, \beta, \gamma > 0$

- The firm's problem is

$$\max_{z \geq 0} p z_1^\alpha z_2^\beta z_3^\gamma - w_1 z_1 - w_2 z_2 - w_3 z_3$$

- Now, is this supermodular?
- Yes! If we look at

$$\frac{\partial g}{\partial z_1} = p \alpha z_1^{\alpha - 1} z_2^\beta z_3^\gamma - w_1$$

  this is increasing in z_2 and z_3

  Similarly, $\frac{\partial g}{\partial z_2}$ is increasing in z_1 and z_3, and $\frac{\partial g}{\partial z_3}$ is increasing in z_1 and z_2

- What about increasing differences in choice variables and parameters?
- Well, $\frac{\partial g}{\partial z_1}$ is increasing in p and $-w_1$, and we're free to say it's weakly increasing in $-w_2$ and $-w_3$
- Same with the others, and so the objective function has increasing differences in the choice variables (z_1, z_2, z_3) and the parameters $(p, -w_1, -w_2, -w_3)$
- We also know, conveniently, that Cobb-Douglas production gives a unique solution
- So if p goes up, the firm demands more of every input;

  and if any input price goes up, the firm demands less of every input!

4.2 CES production

- How about a three-input firm – say, capital, labor, and electricity – whose production function is Constant Elasticity of Substitution, but modified to have decreasing returns

- Let

$$f(k, \ell, e) \quad = \quad (k^\rho + \ell^\rho + e^\rho)^{a/\rho}$$

where $\rho \in (0, 1)$ and $a \in (0, 1)$

(In general, CES allows $\rho < 0$, but I want to stick to the cases of $\rho > 0$)

- The firm's objective function is

$$p(k^\rho + \ell^\rho + e^\rho)^{a/\rho} - w\ell - rk - c_e e$$

- To figure out whether this is supermodular, let's calculate

$$\frac{\partial g}{\partial k} \quad = \quad p\frac{a}{\rho}(k^\rho + \ell^\rho + e^\rho)^{\frac{a}{\rho}-1}\rho k^{\rho-1} - r \quad = \quad pa(k^\rho + \ell^\rho + e^\rho)^{\frac{a}{\rho}-1}k^{\rho-1} - r$$

- Now, is this increasing in ℓ and e?

- If $\frac{a}{\rho} - 1 > 0$, then yes –
so if $a > \rho$, $\frac{\partial g}{\partial k}$ is increasing in ℓ and e, so the firm's problem is supermodular

- (Given symmetric cross-partials,
we don't have to check whether $\frac{\partial g}{\partial \ell}$ and $\frac{\partial g}{\partial e}$ are is increasing in k;
in theory, we do need to check whether $\frac{\partial g}{\partial \ell}$ is increasing in e,
but given how symmetric the problem is,
it should be pretty clear that we have increasing differences all around.)

- So if $a > \rho > 0$, the problem is supermodular
- As usual, it will have increasing differences in $X = (k, \ell, e)$ and $T = (p, -w, -r, -c_e)$
- So we know exactly what will happen –

  all three inputs k, ℓ, and e will move up together if p increases,

  and move down together if any of the input prices goes up
- (This is "increases" in the modified-strong-set-order sense,

  although we could show the firm's problem has a unique solution,

  so in face we get movement in the predicted direction for each input)

But what about if $a < \rho$?

- With $a < \rho$, the exponent $\frac{a}{\rho} - 1 < 0$, and so $\frac{\partial g}{\partial k}$ is *decreasing* in ℓ and e
- With two inputs, we could still make the problem supermodular by flipping the sign of one of the choice variables:

  we could make the problem supermodular in $(k, -\ell)$,

  and show it has increasing differences in $(k, -\ell)$ and $(-r, +w)$,

  so we could get comparative statics on input prices that way

  (although not on the output price p)
- With three inputs, this won't work – let's see why

- If $a < \rho$, then $\frac{\partial g}{\partial k}$ is decreasing in both ℓ and e,

  so to make it supermodular, we would need to consider the choice variables $(k, -\ell, -e)$
- But now let's look at

$$\frac{\partial g}{\partial(-\ell)} \quad = \quad -\frac{\partial g}{\partial \ell} \quad = \quad -p\frac{a}{\rho}\left(k^\rho + \ell^\rho + e^\rho\right)^{\frac{a}{\rho}-1}\rho\ell^{\rho-1} + w$$

- This is increasing in e, hence *decreasing* in $-e$
- So there's no combination of "sign flips" that can make the problem supermodular
- (Basically, ℓ and e are both substitutes for k, but they're also substitutes for each other;

  to make them both substitutes for k, they need to have the same sign,

  but to make them substitutes for each other, they need different signs.)

- (Or to put it another way, if k, ℓ and e are all pairwise substitutes, then when w goes up and the firm responds by cutting ℓ, then this makes the firm want to increase k; but when ℓ goes down, the firm also wants to increase e, which makes it want to *reduce* k
- So there are effects pushing k in both directions, making the net effect ambiguous unless we know more)
- So with more than two inputs, when $a < \rho$, we can't apply Topkis' Theorem, because we can't make the problem supermodular

- This example gives me an opportunity to emphasize an important point: we're always free to ignore any *parameter* we don't want to deal with – like we did with p when it didn't "fit" – but we always have to consider all of the firm's *choice variables* together
- And if the objective function is not supermodular in the choice variables, there's no way we can apply Topkis' Theorem
- So some problems simply won't work with this approach!

- For this particular problem, I'll post a separate note – it turns out we still could use Topkis to get some comparative statics, we just need to be sneaky about it
- But in some settings, we can't use Topkis at all
- Two choice variables might be complements at some levels and substitutes at others – if $\frac{\partial^2 g}{\partial x_1 \partial x_2}$ is sometimes positive and sometimes negative, we won't be able to apply Topkis
- (Or, Topkis might give us comparative statics that only hold locally – we could say what would happen with small changes from a particular starting point, but it would be different from what would happen at a different starting point)

- This makes sense – capital and labor might be substitutes in some ranges,

  as you could replace workers with machines;

  but complements in other ranges,

  since you at least need a few skilled people to maintain the machines

- So as the price of labor goes up, maybe you initially want more capital, to substitute;

  but as labor gets more and more expensive,

  maybe you're better off shutting down, so you don't need any capital at all

- My point is, MCS is a powerful method, but it only works on some problems

5 A little bit on empirics

5.1 Milgrom and Roberts (1990)

- I motivated the MCS approach with the work of Milgrom and Roberts[1]

  on the clustering of adoption of lots of technological advances by U.S. manufacturers –

  more flexible, programmable equipment, smaller production batches,

  shorter production cycles, broader product lines,

  a speeding up of everything, an emphasis on high quality,

  and certain organizational strategies and workforce managment policies

- I'll mention a few other papers that have thought about empirical testing of these types of complementarities

[1]Paul Milgrom and John Roberts (1990), "The Economics of Modern Manufacturing," *American Economic Review* 80.3

5.2 Athey and Stern (1998)

- Susan Athey and Scott Stern, in a working paper from the late 1990s[2]

  propose an econometric framework for testing for complementarities in production choices

- They write, "A major finding of this [recent] literature is that organizational design practices are "clustered":

  the adoption of practices is correlated across firms,

  and some "sets" of practices consistently appear together.

  Economic theory suggests that such clustering might arise if the choices are complements."

- They then mention shortcomings of existing approaches, and go on to say,

  "Our analysis is tailored for cross-sectional applications where many firms face similar production technologies,

  make comparable choices about organizational design,

  but face different costs or benefits to adoption."

- They are "motivated by the policy implications that follow if practices are interrelated in adoption and productivity.

  For example, if a training subsidy affects the adoption of training programs,

  it will also have indirect effects on the adoption and productivity of complementary practices,

  such as a commitment to job security.

  Consequently, optimal subsidies need to account for both direct and indirect effects on organizational design."

- (They don't actually apply the econometric framework they're proposing,

  which may be why the paper has never been published,

  although it does have over 600 Google citations...)

[2]Susan Athey and Scott Stern (1998), "An Empirical Framework for Testing Theories about Complementarity in Organizational Design," NBER working paper #6600, https://www.nber.org/papers/w6600.pdf

5.3 Bresnahan, Brynjolfsson and Hitt (2002)

- Bresnahan et al.[3] look to explain what's known as "Skill-Biased Technical Change" – the shift in demand toward more-skilled workers, away from less-skilled workers

- They note cynically, "...[SBTC] also tends to be something of a residual concept, whose operational meaning is often "labor demand shifts with invisible causes" – it's often given as an *explanation* for what happened, when it's really just a *description* for something when we don't know the cause – but they seek to better understand its causes

- They note, "Firms do not simply plug in computers or telecommunications equipment and achieve service quality or efficiency gains.

  Instead they go through a process of organizational redesign and make substantial changes to their product and service mix.

  ... That is, IT is embedded in a cluster of related innovations, notably organizational changes and product innovation.

  These three complementary innovations – a) increased use of IT, b) changes in organization practices, and c) changes in products and services – *taken together* are the SBTC that calls for a higher-skilled labor mix."

- They look at data from 300 large U.S. firms in the late 1980s and early 1990s, and conducted surveys of senior HR managers to better understand organization practices; they give a variety of empirical results, but broadly do find evidence of complementarity among the use of IT, a variable they call "work organization," and human capital levels within a firm

[3] Timothy Bresnahan, Erik Brynjolfsson and Lorin Hitt (2002), "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence," *Quarterly Journal of Economics* 117.1

5.4 Bloom and Van Reenen (2007)

- Bloom and Van Reenen[4] study management practices across firms and countries,

  seeking to explain part of the huge variation in productivity,

  even within an industry and within a country

- Their primary focus isn't complementarities –

  they focus more on documenting the range of different practices,

  and note that there are just a lot of well-run firms and a lot of poorly-run firms:

  "Most notably, we see a large number of firms that appear to be extremely badly managed,

  with ineffective monitoring, targets and incentives."

- The biggest predictors they find:

  higher levels of competition in a market are associated with better management practices,

  and family firms where the oldest male child is the CEO tend to be terrible

- They do, however, mention the Athey and Stern paper,

  and test for complementarity in one particular area:

  they hypothesize, and find some evidence for,

  complementarity between good human-capital management practices and high-skill environments (which they proxy for with workers having above-average salaries, or more workers having college degrees)

[4]Nicholas Bloom and John Van Reenen (2007), "Measuring and Explaining Management Practices Across Firms and Countries," *Quarterly Journal of Economics* 72.4

5.5 Others

- Caroli and Van Reenen (2001)[5] similarly find evidence of complementarity between certain organizational changes and worker skills
- They find that certain organizational changes reduce the demand for unskilled workers;

  are negatively associated with bigger wage differentials for higher-skilled workers

  (because these changes are more likely when higher-skilled workers are available cheaply);

  and lead to greater productivity increases in establishments with larger initial skill endowments
- Cassiman and Veugelers (2006)[6] find that internal R&D and external knowledge acquisition are complementary, but the degree of complementarity depends on other factors

5.6 Why am I talking about this?

- Monotone Comparative Statics and Topkis' Theorem are very powerful, general results,

  which can apply to a wide range of problems
- That said, they came about largely to try to explain particular types of complementarities between different changes in firm behavior,

  so I thought it was worth mentioning a bit of the empirical literature looking at those complementarities
- (And as I noted on the first day:

  empirical observation can guide the development of theory,

  and theory can both tell us what to look for in data and help us develop tools to "look for it" better)

- Up next: one other cool application of Topkis' Theorem:

  LeChatelier's Principle

[5] Eve Caroli and John Van Reenen (2001), "Skill-Biased Organizational Change? Evidence from a Panel of British and French Establishments," *Quarterly Journal of Economics* 116.4

[6] Bruno Cassiman and Reinhilde Veugelers (2006), "In Search of Complementarity in Innovation Strategy: Internal R&D and External Knowledge Acquisition," *Management Science* 52.1

6 A Cool Application of MCS: The LeChatelier Principle

- I mentioned last week that it's natural to be skeptical of Decreasing Returns to Scale – why can't a firm just clone a factory and build another identical one?
- And yet we often need to assume DRS to get interior solutions to profit maximization, and positive but finite firm profits
- One solution is to imagine that there are some inputs that can't be adjusted quickly
- A firm might have constant returns to scale when you count all inputs,
  but some inputs are fixed in the short-term,
  and with those fixed, there are decreasing returns over the rest
- One example is Cobb-Douglas production, say

$$f(z) = z_1^\alpha z_2^\beta z_3^\gamma$$

  where output is CRS if the exponents sum to 1, and DRS if they sum to less than 1
- Now, suppose the exponents sum to 1, but that z_1 is fixed in the short term
- Then in the short term, z_1 is like a parameter,
  and when the firm thinks about setting z_2 and z_3, it sees itself as having decreasing returns
- Later on, z_1 will adjust to some new, higher level,
  but then for the short term, we'll again have DRS over (z_2, z_3) until we can adjust it again

- Of course, I'm being very vague about what's "short term" and what's "long term"
- But at least this is one reasonable story for why we might have a technology that's constant-returns-to-scale in the long run –
  the firm could always clone its entire operation to double inputs and double outputs –
  and yet still leads to a unique interior solution to the firm's problem,
  because on a shorter-term horizon there are decreasing returns

- Once we're assuming that one or more inputs may be fixed in the short term and only adjustable in the long term, this leads us to a very cool result

- Basically, we can use Topkis' Theorem to show that under pretty general conditions,

  when the economic environment changes,

  the long-run change in a firm's responses is bigger than the short-run change

- This is known as the LeChatelier Principle

- LeChatelier was actually a chemist – other people imported his ideas into economics[7]

  (and in fact, if I understand it correctly, the result in chemistry is more or less the opposite)

- I'll formalize the result for two inputs, but it still holds for more,

  just as long as the firm's problem is supermodular (with, perhaps, some sign flips)

- **Proposition.** Suppose a firm has two inputs, capital and labor, with capital fixed in the short-run but labor free to adjust quickly. If either

  1. output is supermodular in (k, ℓ), *or*
  2. output is supermodular in $(k, -\ell)$

  then if the price of labor w goes up, the long-term decrease in labor demand (when capital is allowed to adjust) is larger than the short-term decrease (while capital is fixed).

- Paul Samuelson proved that this result holds for "small changes,"

  but without the need for supermodularity;

  Milgrom and Roberts (1996 AER) showed it in more or less the form I'm showing

- The intuition is that at first, when w goes up, the firm reduces ℓ, but the level of capital is stuck

- In the long run, when capital can be adjusted, the reduction in ℓ leads to a change in capital – either up or down –

  which in turn *amplifies* the firm's change in ℓ

- (If k and ℓ are complements, then in the long run, capital is reduced, which reduces labor further;

  if k and ℓ are substitutes, then in the long run, capital is increased, but this again reduces labor further)

[7]See Paul Milgrom and John Roberts (1996), "The LeChatelier Principle," *American Economic Reivew* 86.1, and several works by Paul Samuelson that are cited there.

- Let's show it more formally
- For simplicity, let's suppose the firm's problem has a unique solution;

  at initial prices (p, r, w), we'll call it (k_0, ℓ_0)
- Then suppose the wage goes up from w to w',

  and at prices (p, r, w') when capital has not yet adjusted,

  we'll label the firm's optimal plan (k_0, ℓ_1);

  and after the firm can adjust capital as well, we'll call it (k_2, ℓ_2)
- The result we want is that $\ell_2 \leq \ell_1 \leq \ell_0$

- First, let's suppose f is supermodular in k, ℓ
- In the short run, the firm's capital level is fixed at k_0, so the firm solves

$$\max_{\ell} \{pf(k_0, \ell) - rk_0 - w\ell\}$$

  This is a one-dimensional problem with increasing differences in ℓ and $-w$,

  so of course when w goes up, ℓ goes down, or $\ell_1 \leq \ell_0$
- In the long run, the firm solves

$$\max_{k,\ell} \{pf(k, \ell) - rk - w\ell\}$$

  This is supermodular in (k, ℓ), with increasing differences in (k, ℓ) and $-w$,

  so if w goes up, k and ℓ both go down –

  so $k_2 \leq k_0$ and $\ell_2 \leq \ell_0$
- But the result we want is that $\ell_2 \leq \ell_1$ –

  that the long-run adjustment is bigger than the short-run
- Why is that?

- Well, if (k_2, ℓ_2) solve the firm's problem at the new prices,

  then ℓ_2 must be the optimal level of labor at the new prices, given $k = k_2$

- And we know ℓ_1 is optimal at the same prices, given $k = k_0$

- So consider the one-dimensional problem of choosing ℓ, fixing prices at the new level,

  but now treating k as a parameter:

$$\max_{\ell} \{pf(k, \ell) - rk - w'\ell\}$$

- This has increasing differences in the choice variable ℓ and the parameter k,

  so if k goes down, ℓ goes down –

  so when k drops from k_0 to k_2, ℓ goes down, giving $\ell_2 \leq \ell_1$

- The intuition is that initially, when w goes up, the firm lowers ℓ;

  this reduces the marginal product of capital, so the firm would like to reduce k,

  but it can't immediately

- In the longer run, when the firm is able to reduce k, it does;

  this lowers the marginal product of labor, which causes it to reduce labor more

- (In a sense, the whole point of supermodularity is to ensure that all the "feedback effects" go in the same direction and reinforce each other – lowering capital makes you want to lower labor, that makes you want to lower capital, that makes you want to labor, and so on – and all the indirect effects go the same way if there are more than two choice variables)

- What's cool is, if capital and labor are instead substitutes – $\partial^2 f / \partial k \partial \ell \leq 0$ –

  the proof is almost identical, we just have to flip the sign of k throughout

- Now when w goes up initially, ℓ goes down with k fixed;

  but since f is supermodular in $(-k, \ell)$, once k is free to adjust, k now goes up, which pushes ℓ down even more

- Samuelson proved this holds generally – short-term changes are smaller than long-term changes – for **small** changes

- Basically, "small" means that price changes, and input responses, are small enough that the sign of the cross-partials stay the same, so what we just did goes through

- Milgrom and Roberts showed it holds for large changes, as long as the firm's problem is supermodular –

  and of course, you can make it supermodular by flipping the signs of some inputs if needed

- So, for example, if the firm had four inputs,

  and production was supermodular in $(z_1, z_2, -z_3, -z_4)$,

  and z_1 was fixed in the short-run,

  then an increase in w_2 would lead to a decrease in z_2 and increases in z_3 and z_4;

  and then z_2 would fall further, and z_3 and z_4 would go up more,

  once z_1 was free to adjust, since z_1 would fall in response to the other changes

- **However,** if two inputs are sometimes complements but sometimes substitutes,
  and changes are "big", then the result can fail to hold
- (This is the example I didn't get to in lecture)
- For example, consider a single-output firm,
  with two inputs, capital and labor;
  and suppose Y contains just three points, $(0, 0, 0)$, $(10, 0, -2)$, and $(10, -1, -1)$
- Fix output price p at 1 throughout
- Initial input prices are $(r, w) = (3, 2)$, so labor is cheaper than capital, and $Y^* = (10, 0, -2)$
- Now let input prices change to $(r, w) = (3, 6)$
- With capital fixed at 0 in the short run, the firm would need 2 units of labor to produce, and is better off shutting down
- But in the long run, the firm could switch to $(10, -1, -1)$ and still make a profit
- So the short-term change is to drop labor from 2 to 0, but in the long run it recovers to 1
- This is because capital and labor are sometimes complements and sometimes substitutes –
  so $f_{k\ell}$ is neither always positive nor always negative
- If they were always complements or always substitutes –
  or in a continuous problem with sufficiently small changes –
  the long-term change is always larger than the short-term.