Lecture 3: Production in the Differentiable Case

1 Where we are

- Tuesday, we thought about the empirical implications of our model of profit maximization, and considered the following questions
  (and found the following answers):

  1. *When are production or profit data consistent with our model?*
     When $Y^I \subseteq Y^O$, or when the Weak Axiom of Profit Maximization holds, or
     $$p \cdot y \quad \geq \quad p \cdot y'$$
     for any two observations (p, y) and (p', y') in the data
  2. *Given some set of observations, what do we know about Y?*
     We found the "smallest" set Y^I and "largest" set Y^O that Y could be,
     and said that any set "in between" would rationalize the data
  3. *With all possible data, would we learn Y?*
     If Y is convex, then yes, $Y = Y^O$;
     and if Y is convex and $y(p) = Y^*(p)$, then $Y = Y^I_{FD}$ as well
     (But if Y is not convex, then $Y^I_{FD} \subsetneq Y^O$ and we don't know Y exactly)

- The last two results – that when Y is convex,
  it can be recovered exactly from Y^O (which comes from the profit function),
  or from Y^I_{FD} (which comes from the production choices) under an additional assumption –
  follow closely from the Separating Hyperplane Theorem

- As "bonus material," I went ahead and posted proofs of the two Separating Hyperplanes results on Canvas, although you don't need to know them,

and if you want to follow the proof by Weitzman, it may make more sense in about a week,

after we introduce the cost function

- The results say that *if we already know* that Y is convex (or are willing to assume it), then we learn Y exactly from just the profit function (since Y^O is computed from π), or from production at every price even if we only see a subset of the optimal plans
- So if we know Y is convex, it's recoverable

- On the other hand, if Y is not convex, we can only learn that with an additional assumption – that we get to observe *all* of $Y^*(p)$ for each p, not just an element of it – and even then, we won't get to fully recover Y exactly
- The reason is because when Y has a non-convexity, the "non-convex part" consists of points that would never be optimal anyway, so we'd never see the firm choose them anyway, so we can't tell whether they're feasible or not because even if they were the firm would never choose them
- In fact, for any set of observations that are rationalizable at all, there's a convex production set that would rationalize them – it's Y^O
- (I mentioned last time that Y^O is always convex; and we saw the result that data are rationalized by any set Y containing Y^I and contained in Y^O, so Y^O itself would rationalize the data)
- So any rationalizable data, is consistent with a production set that's convex, even if the firm's true production set isn't convex
- (The only way we can tell whether the production set is convex, is if we make the additional assumption that we see observe every optimal production plan, because then if we see the firm choose two plans at the same price vector p, but not choose the points on the line between them, we know those points are "missing" from Y, and so Y is not convex)

2 General properties of π and Y^*

- Today, we go back to our profit function,

$$\pi(p) \quad = \quad \max_{y \in Y} p \cdot y$$

and optimal supply correspondence

$$Y^*(p) \quad = \quad \arg\max_{y \in Y} p \cdot y$$

and consider their properties as mappings from prices to outcomes

(That is, we think about π as a function mapping each price vector to a real number,

and Y^* as a correspondence mapping each price vector to a subset of $\mathbb{R}^k$)

- First, a few general properties of the profit function and supply correspondence,

which I believe were discussed in discussion section last week

1. π is homogeneous of degree 1,

   and Y^* is homogeneous of degree 0

   (for any constant $\lambda > 0$, $\pi(\lambda p) = \lambda \pi(p)$ and $Y^*(\lambda p) = Y^*(p)$)

2. π is convex

3. If Y is convex, then for any given p, $Y^*(p)$ is convex;

   if Y is strictly convex, $p \neq 0$, and $Y^*(p)$ is nonempty, then $Y^*(p)$ is single-valued.

- (Note that convex means two different things here depending on context:

$\pi : \mathbb{R}^k_+ - \{0\} \to \mathbb{R}$ is a *function*,

so saying it's convex means $\pi(tp + (1-t)p') \leq t\pi(p) + (1-t)\pi(p')$ for any p, p' and any $t \in (0, 1)$;

while $Y^*(p)$ and Y are sets, so when they're convex,

it means that for any two points in the set, the line between them is too)

- I'll skip the formal proof of the first result[1] –

  the idea is, if you double all prices, the firm's optimal production plan stays the same,

  but the firm now makes twice as much money

- For the second, to show π is a convex function of prices, we need to show

$$\pi(tp + (1-t)p') \quad \leq \quad t\pi(p) + (1-t)\pi(p')$$

  for any two price vectors p' and p and any constant $t \in (0, 1)$

- So let $p^t = tp + (1-t)p'$, and pick $y^t \in Y^*(p^t)$, so that $\pi(p^t) = p^t \cdot y^t$

  Since $y^t \in Y$, $\pi(p) \geq p \cdot y^t$ and $\pi(p') \geq p' \cdot y^t$

  So $t\pi(p) + (1-t)\pi(p') \geq tp \cdot y^t + (1-t)p' \cdot y^t = p^t \cdot y^t = \pi(p^t)$

  (But the intuition is that the max of a bunch of linear functions is convex –

  show it with a picture!)

- For the third result, I'll again skip the proof,

  because I don't think you get a lot of insight from it,

  just a bit of practice manipulating the definition of a convex set[2]

- But these results are nice, in part,

  because they give a natural condition under which optimal supply is single-valued –

  and therefore Y^* is a function, rather than a correspondence

- This allows us to prove some stronger results about π and Y^*

- And if we're willing to go one step further –

  and assume that Y^* is not just a function, but also differentiable –

  we can prove even more interesting stuff

[1] Formally, the proof is $\pi(\lambda p) = \max_{y \in Y} (\lambda p) \cdot y = \lambda \max_{y \in Y} p \cdot y = \lambda \pi(p)$ for the first part, and $Y^*(\lambda p) = \{y \in Y : (\lambda p) \cdot y = \pi(\lambda p)\} = \{y \in Y : (\lambda p) \cdot y = \lambda \pi(p)\} = \{y \in Y : p \cdot y = \pi(p)\} = Y^*(p)$ for the second.

[2] Suppose Y is convex; fix a price vector p, pick two points y and y' in $Y^*(p)$, and let $y^t = ty + (1-t)y'$ be a point on the line between them. Since y and y' are both in Y and Y is convex, $y^t \in Y$. If y and y' are both in $Y^*(p)$, they're equally profitable at prices p, with $p \cdot y = p \cdot y' = \pi(p)$. We can then calculate

$$p \cdot y^t \quad = \quad p \cdot \big(ty + (1-t)y'\big) \quad = \quad t(p \cdot y) + (1-t)(p \cdot y') \quad = \quad t\pi(p) + (1-t)\pi(p) \quad = \quad \pi(p)$$

to see that y^t is as profitable as y and y', and is therefore also in $Y^*(p)$.

If Y is strictly convex, the same thing holds, except if $y \neq y'$, then y^t is in the *interior* of Y. This means there's an ϵ-ball around y^t which is contained in Y, which means there are other points in Y which would be strictly more profitable than y^t, hence give strictly higher profits than $\pi(p)$ at prices p. Since this is impossible, it rules out having multiple points $y \neq y'$ in $Y^*(p)$ when Y is strictly convex.

- So that's where we're headed next
- For the rest of today, we'll assume Y^* is single-valued (and often differentiable),
  and for convenience, we'll use the notation $y(p)$ for $Y^*(p)$
  now that $Y^*(p)$ is a single point rather than a set

- Before we get to more results,
  we're going to establish one cool general mathematical result
  that we'll use in one of the proofs (and a bunch more times going forward)

3 The Envelope Theorem

- Consider an unconstrained optimization problem

$$\max_x f(x,t)$$

where x is the choice variable and t is an exogenous parameter, and define

$$V(t) \quad = \quad \max_{x \in X} f(x,t)$$

as the max

- (In our production model, t is the vector of prices,

  x is the choice of production plan, and V is the profit function $\pi(p)$;

  I use the envelope theorem constantly in models of auctions,

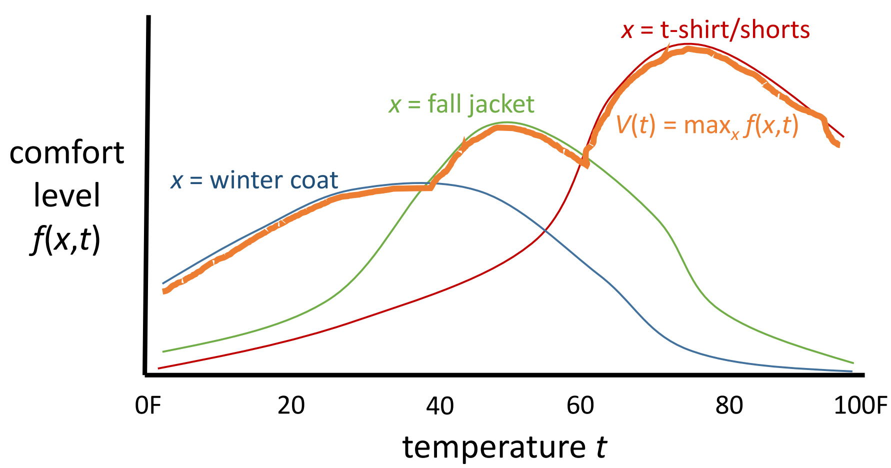
  where t might be how highly you value an object (exogenous) and x is how you bid.)

- We'll want t to be continuous, but x can be continuous or discrete – here's an example I like

- t is the temperature outside, x is three choices of how you're dressed –

  in a winter coat, a fall jacket, or a t-shirt and shorts –

  and $f(x,t)$ is how comfortable you are



- Now, the point of the Envelope Theorem is to relate the derivative of the "value function" V (which is a function only of t) to the objective function f (which is a function of x and t)

- And the punchline is, they relate quite strongly

Theorem (The Envelope Theorem). Let $V(t) = \max_{x \in X} f(x,t)$, where t is a continuous-valued parameter, and let $x^*(t)$ be the set of maximizers. If $V'(t)$ exists, then

$$V'(t) = \frac{\partial f}{\partial t}(\hat{x}, t)$$

for any $\hat{x} \in x^*(t)$.

- The key idea behind the proof is that if we look at the objective function at a particular choice $\hat{x}$, so it's just a function of t,

  and we look at the value function V as a function of t,

  the value function is always weakly higher, but they're equal at values of t where $\hat{x}$ is optimal
- (look at it on previous picture)
- Knowing that...
  - Fix $\hat{t}$, and fix $\hat{x} \in x^*(\hat{t})$
  - Consider a new function $g(t) = V(t) - f(\hat{x}, t)$
  - This is weakly positive for all t, and 0 at $t = \hat{t}$
  - Which means at $\hat{t}$, it's either flat, or achieves a local minimum
  - Either way, $g'(\hat{t}) = 0$, which means $V'(\hat{t}) = \partial f / \partial t(\hat{x}, \hat{t})$

- (You can also get a bit of cheap intuition from thinking about the special case where $x^*(t)$ is single-valued and differentiable
- In that case, $V(t) = f(x^*(t), t)$, and if we differentiate with respect to t, we get

$$V'(t) = \frac{\partial f}{\partial x}(x^*(t), t)(x^*)'(t) + \frac{\partial f}{\partial t}(x^*(t), t)$$

  but because $x^*(t)$ is the maximizer, $\frac{\partial f}{\partial x} = 0$ at $(x^*(t), t)$,
  so the first term is zero, giving $V' = \frac{\partial f}{\partial t}(x^*(t), t)$
- Or, if you're already at the optimal x and then t changes slightly and you adjust,

  the change in t gives you a first-order change in payoff,

  but because you're close to the optimum, the change in x only gives a second-order change,

  so V' reflects only the "direct effect" of the change in t
- The theorem is just saying this all holds generally –

  even if x^* isn't single-valued and smooth –

  whenever V is differentiable)

Two more related results:

Theorem. Let $V(t) = \max_x f(x,t)$, where t is a continuous-valued parameter, and let $x^*(t)$ be the set of maximizers.

1. (Integral Version of the Envelope Theorem.) If $\frac{\partial f}{\partial t}(x,t)$ exists for all x and t, and is bounded above by an integrable function for all x and almost all $t \in [a,b]$, then for any selection $\hat{x}(s)$ from $x^*(s)$,
$$V(b) \quad = \quad V(a) + \int_a^b \frac{\partial f}{\partial t}(\hat{x}(s), s) ds$$

2. (Some results about when V' exists.) If X is compact and both f and $\frac{\partial f}{\partial t}$ are continuous, then V is differentiable at t if $x^*(t)$ is a singleton, V is concave, or t is a maximizer of V

- We already know, when V is differentiable, what its derivate must be
- The second part here gives conditions for when V is differentiable
- And the first part gives conditions for when, even if V is not differentiable everywhere, it's differentiable almost everywhere and absolutely continuous, so it's still the integral of its derivative wherever that exists.

- For more on the Envelope Theorem, see Milgrom and Segal (2002), "Envelope Theorems for Arbitrary Choice Sets," *Econometrica* 70.2

4 now let's use it!

- Like I said, we're going to consider the case where Y^* is single-valued and differentiable, so for that case, we'll simplify the notation to $y(p)$
- That is, we'll let

$$y(p) \quad = \quad \arg\max_{y \in Y} p \cdot y$$

  and consider the properties of $y(\cdot)$ as a function of p

- The first result:

Hotelling's Lemma. Let Y be closed and satisfy free disposal. If $y(\cdot)$ is single-valued in a neighborhood of p, then π is differentiable at p and

$$\frac{\partial \pi}{\partial p_i}(p) \quad = \quad y_i(p)$$

- Proof is the Envelope Theorem: we know that

$$\pi(p) \quad = \quad \max(p \cdot y)$$

  where p is a parameter and $y \in Y$ is the choice variable

- The envelope theorem then says that

$$\frac{\partial \pi}{\partial p_i} \quad = \quad \frac{\partial}{\partial p_i}(p_1 y_1 + p_2 y_2 + \cdots + p_k y_k)\Big|_{y=y^*(p)} \quad = \quad y_i|_{y=y^*(p)} \quad = \quad y_i(p)$$

- (Yeah, that's the whole proof –
  that's why the envelope theorem is so awesome sometimes!)

- Next, we establish a group of results that hold if $y(\cdot)$ is not only single-valued, but also differentable
- When $y(\cdot)$ is differentiable, define its Jacobian matrix

$$D_p y(p) = \begin{bmatrix} \frac{\partial y_1}{\partial p_1} & \frac{\partial y_1}{\partial p_2} & \cdots & \frac{\partial y_1}{\partial p_k} \\ \frac{\partial y_2}{\partial p_1} & \frac{\partial y_2}{\partial p_2} & \cdots & \frac{\partial y_2}{\partial p_k} \\ \vdots & \vdots & & \vdots \\ \frac{\partial y_k}{\partial p_1} & \frac{\partial y_k}{\partial p_2} & \cdots & \frac{\partial y_k}{\partial p_k} \end{bmatrix}$$

Proposition. Let Y be closed and satisfy free disposal. If $y(\cdot)$ is single-valued and differentiable, then for any p, $D_p y(p)$ is symmetric and positive semidefinite, and $[D_p y(p)]\, p = 0$.

- Symmetry is immediate from Hotelling's Lemma and the symmetry of mixed partials:

$$\frac{\partial y_i}{\partial p_j} = \frac{\partial}{\partial p_j}\left(\frac{\partial \pi}{\partial p_i}\right) = \frac{\partial}{\partial p_i}\left(\frac{\partial \pi}{\partial p_j}\right) = \frac{\partial y_j}{\partial p_i}$$

- We could argue that $D_p y$ must be positive semidefinite because it's the Hessian matrix (matrix of mixed partials) of π, which is convex;

  but we can also show it straight from the Law of Supply
- We need to show $z[D_p y]z' \geq 0$ for any vector z
- Let's make z small, and call it $dp = [dp_1\ dp_2\ \ldots\ dp_k]$
- If we calculate

$$[D_p y]z' = \begin{bmatrix} \vdots \\ \frac{\partial y_i}{\partial p_1} dp_1 + \frac{\partial y_i}{\partial p_2} dp_2 + \ldots + \frac{\partial y_i}{\partial p_k} dp_k \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ dy_i \\ \vdots \end{bmatrix}$$

- So $z[D_p y]z' = dp \cdot dy$, which we know is weakly positive from the Law of Supply

- For the last result, we need to show $[D_p y]p = 0$
- There are lots of different ways to prove this[3]
- I'll skip the technical and give the geometric-intuitive
- Since we know $[D_p y]$ is symmetric, we can replace it with its transpose:

  $[D_p y]^T p$ is a column matrix whose i^{th} element is

$$\left(\frac{\partial y_1}{\partial p_i}, \frac{\partial y_2}{\partial p_i}, \ldots, \frac{\partial y_k}{\partial p_i} \right) \cdot (p_1, p_2, \ldots, p_k)$$

- The first vector is the change in y in response to a small change in some price
- Now, we know that whatever Y looks like,

  it's always optimal to be on the boundary, not the interior
- So as p_i changes, we're going to be moving along that boundary in some way
- And so this is saying, a move along the boundary of Y is perpendicular to the price vector;

  or really, optimal y is where the boundary of Y is perpendicular to p
- But, like, obviously, amirite?

[3]Here are three. (1) $[D_p y]p$ is a column vector whose i^{th} element is $p_1 \frac{\partial y_i}{\partial p_1} + p_2 \frac{\partial y_i}{\partial p_2} + \cdots + p_k \frac{\partial y_i}{\partial p_k}$. We know $y(\cdot)$ is homogeneous of degree 0, meaning $y_i(\lambda p) = y_i(p)$; if we differentiate with respect to λ, we get

$$p_1 \frac{\partial y_i}{\partial p_1}(\lambda p) + p_2 \frac{\partial y_i}{\partial p_2}(\lambda p) + \cdots + p_k \frac{\partial y_i}{\partial p_k}(\lambda p) \quad = \quad 0$$

which must hold at $\lambda = 1$, giving the result. (2) Alternatively, since we've already shown that $[D_p y]$ is symmetric, we could instead show that $[D_p y]^T p = 0$; this is a column matrix whose i^{th} element is $p_1 \frac{\partial y_1}{\partial p_i} + p_2 \frac{\partial y_2}{\partial p_i} + \cdots + p_k \frac{\partial y_k}{\partial p_i}$. If we write $\pi(p) = p \cdot y(p)$ and differentiate with respect to p_i, we get

$$\frac{\partial \pi}{\partial p_i} \quad = \quad y_i + p_1 \frac{\partial y_1}{\partial p_i} + \ldots + p_k \frac{\partial y_k}{\partial p_i}$$

Noting that $\frac{\partial \pi}{\partial p_i}(p) = y_i(p)$ by Hotelling's Lemma, the remaining terms must sum to zero. (3) Finally, another way to show $[D_p y]^T p = 0$ is to note that note that if $y(p)$ is optimal at p, it must be the solution to $\max_{p'}(p \cdot y(p'))$, since $y(p')$ for any p' is another feasible production plan the firm could have chosen instead; the first-order conditions of this problem with respect to p'_i is

$$p_1 \frac{\partial y_1}{\partial p_i}(p') + p_2 \frac{\partial y_2}{\partial p_i}(p') + \ldots + p_k \frac{\partial y_k}{\partial p_i}(p')$$

which must be equal to 0 at the optimum, $p = p'$.

5 Rationalizability in the differentiable case

- Tuesday, we figured out when data – a finite series of observations of prices and firm production and/or profit – is rationalizable
- Specifically, we found that data is rationalizable when $Y^I \subseteq Y^O$,

  which is equivalent to when the Weak Axiom of Revealed Preference holds:

  that $p \cdot y \geq p \cdot y'$ for any two observations (p, y) and (p', y')
- But checking whether $Y^I \subseteq Y^O$ this way is tedious –

  it means checking every choice $y(p)$ against every other choice $y(p')$ at each price p –

  seems fine for small data sets, but pretty annoying if you observe a lot of p
- We've shown that if a firm is maximizing profits, then π and Y^* have certain nice properties – maybe we can use therse to come up a "local test" of rationalizability

- Suppose we had observations for every strictly positive price vector $p \gg 0$, or $P = \mathbb{R}^k_{++}$
- (Actually, any **open** convex subset of $\mathbb{R}^k_+$ would work, but I like $\mathbb{R}^k_{++}$)
- Further, suppose the supply correspondence $y(p)$ we observe is single-valued,

  and the profit function we observe is differentiable
- We just proved some properties that π and $y(\cdot)$ must satisfy;

  it turns out, these are also sufficient

Poposition. Start with a supply function $y : P \to \mathbb{R}^k$ and a differentiable profit function $\pi : P \to \mathbb{R}$ defind on an open convex set $P \in \mathbb{R}^k_+$ and satisfying the adding-up condition. y and π are jointly rationalizable if and only if the following hold:

1. $y_i(p) = \frac{\partial \pi}{\partial p_i}(p)$ for every i and every $p \in P$ (Hotelling)
2. π is convex

- We already know that if the firm is maximizing profits, these must be true
- Now we're saying that if we observe y and π but don't know Y, these are sufficient to know that y and π are consistent with profit maximization according to some Y

6 skipped this part – I believe it'll be covered in section, but if not, we'll talk about it next week

- Before we prove this, let's consider how this extends to the cases where we only observe y *or* π, not both:

- **Proposition.** A differentiable supply function $y : P \to \mathbb{R}^k$ (P open and convex in $\mathbb{R}^k_+$) is rationalizable if and only if it's homogeneous of degree 0 and $D_p y(p)$ is symmetric and positive semidefinite.

- (To prove this, we would let $\pi(p) = p \cdot y(p)$, so adding-up would automatically hold;

  we'd then confirm that Hotelling holds and that π is convex, and invoke the big result.)

- **Proposition.** A differentiable profit function $\pi : P \to \mathbb{R}$ (P open and convex in $\mathbb{R}^k_+$) is rationalizable if and only if it's homogeneous of degree 1 and convex.

- (We already have convexity; we'd define $y(p)$ by letting $y_i = \frac{\partial \pi}{\partial p_i}$, so Hotelling holds;

  we'd then need to prove that the adding-up condition holds, which it does, and then invoke the big result.)

- (Something cool here: we know the profit function must be homogeneous of degree 1,

  and the optimal supply correspondence must be homogeneous of degree 0

- If we observe both π and y, though, we don't have to check – these conditions are redundant

- Because if π and y satisfy both Hotelling's lemma and the adding-up condition, they *have to* be homogeneous of the right degree

- But if we only observe y *or* π, we need to check that it's homogeneous of the right degree)

- So now, let's prove the big result: that if we observe functions π and y,

  they're jointly rationalizable if they satisfy adding-up and Hotelling and π is convex.

7 Proof of the "Big Result"

- Given functions $y : P \to \mathbb{R}^k$ and $\pi : P \to \mathbb{R}$ satisfying adding-up
  we know from last time that the data is rationalizable if and only if $Y^I \subseteq Y^O$
- We'll show that if π is convex and Hotelling's Lemma holds,
  then $Y^I \subseteq Y^O$
- To show $Y^I \subseteq Y^O$, we can pick a point $y^* \in Y^I$ and show it's in Y^O
- So pick a point $y^* \in Y^I$, and let p^* be the prices at which it was chosen
- To show $y^* \in Y^O$, we want to show that

$$p \cdot y^* \leq \pi(p)$$

  for every $p \in P$
- With the point y^* fixed, define a new function

$$L(p) \quad \equiv \quad \pi(p) - p \cdot y^*$$

  as the "loss" from selecting y^* instead of the optimal production plan
- Showing $y^* \in Y^O$ is the same as showing $L(p) \geq 0$ for every p
- And we know, since y^* was chosen at p^*, that $L(p^*) = 0$
- So $y^* \in Y^O$ is equivalent to showing that $L(p)$ achieves a minimum at $p = p^*$

- Now, we assumed π was differentiable,
  and $p \cdot y^*$ is just a linear function of p,
  so $L(\cdot)$ is differentiable;
  and if π is convex in p, since $p \cdot y^*$ is linear,
  $L(\cdot)$ is convex
- And since P is an open set, p^* is interior
- So L achieving a minimum at an interior point $p = p^*$,
  since L is differentiable and convex,
  is equivalent to the first-order condition holding at p^*

- What's the FOC?

$$\frac{\partial L}{\partial p_i} \quad = \quad \frac{\partial \pi}{\partial p_i} - y_i^*$$

and at $p = p^*$, recalling y^* was chosen at p^*, this is

$$\frac{\partial \pi}{\partial p_i}(p^*) - y_i(p^*)$$

- Hotelling's Lemma is exactly that this is equal to zero!
- So Hotelling tells us the FOC for L holds at $p = p^*$;

  convexity of π gives us convexity of L, which tells us the FOC guarantees a minimum;

  and L being minimized at $p = p^*$ guarantees that $Y^I \subseteq Y^O$,

  which means that y and π are jointly rationalizable