

# Haciendo puentes entre la econometría estructural y la evaluación de impacto

Francesco Bogliacino

A dark blue, irregular ink splash or blotch serves as the background for the text. The splash has a textured, painterly appearance with various shades of blue and some white highlights, giving it a dynamic and artistic feel. The text is centered within the splash.

# Causalidad y parametros estructurales

# Causalidad y modelos

- En econometría, el enfoque estándar es interpretar la causalidad dentro de modelos económicos especificados previamente
- Si postulamos

$$Y = F(X_1, X_2, \dots, X_n)$$

- Entonces el impacto causal de  $X_j$  se puede definir como

$$\frac{\partial F}{\partial X_j}$$

# Interpretación

- La definición del efecto causal es independiente de los datos (no estamos diciendo nada sobre observabilidad);
- La causalidad supone una variación controlada = ceteris paribus (successionist causation à la Hume);
- $X_j$  tiene que poderse variar independientemente;
- Las variaciones admisibles determinan la interpretación del efecto causal, por ejemplo la de variar un input en una FDP Leontief
- Si asumimos separabilidad  $Y = \sum \varphi(X_i)$  podemos definir los efectos causales de manera independiente

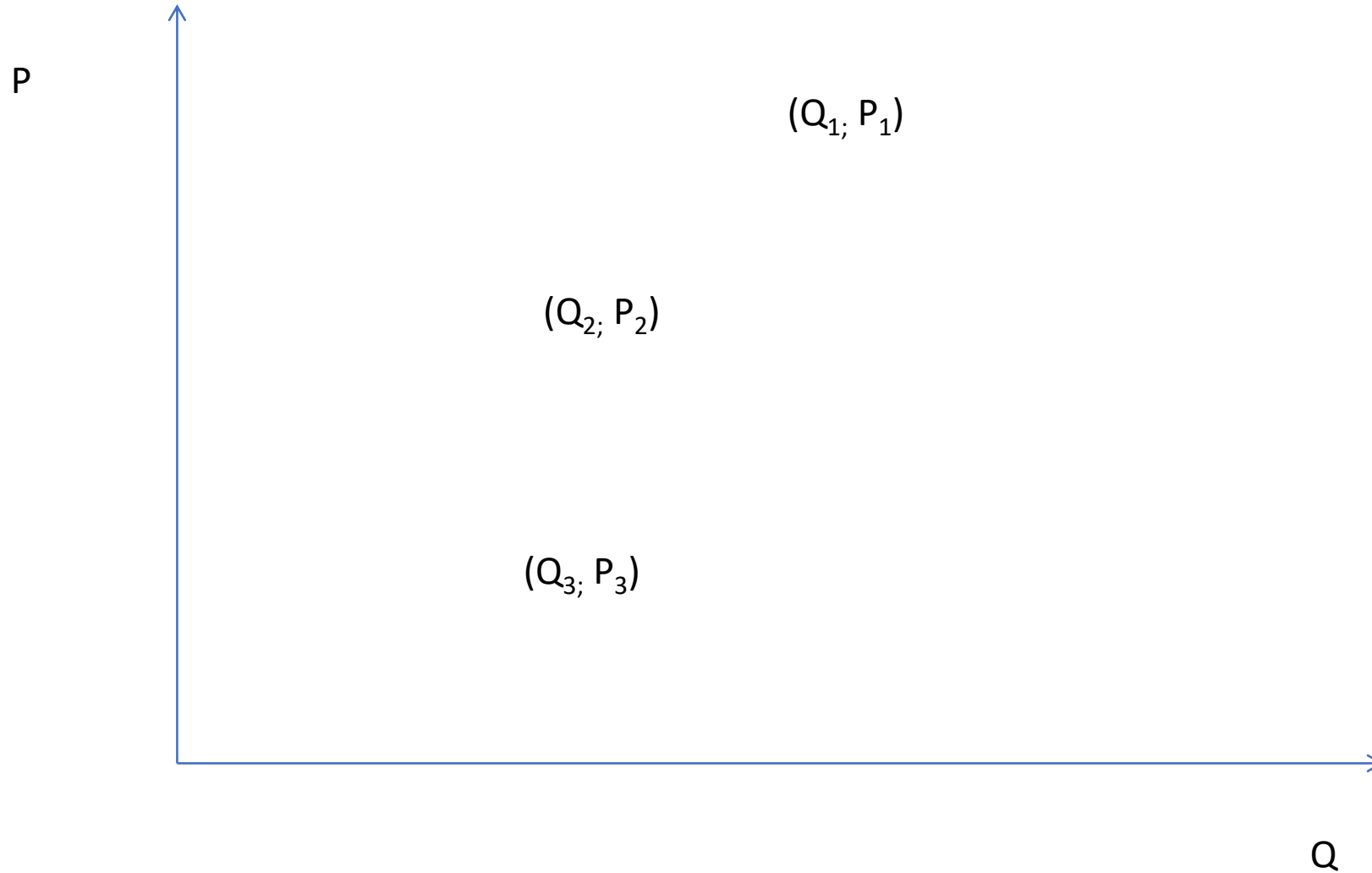
# Rubin Model

- Podemos definir  $X_1 = x_1$  como el tratamiento y
$$Y_{x_1} = F(x_1, X_2, \dots, X_n)$$
- Como el outcome potencial. El modelo de Rubin es un caso particular

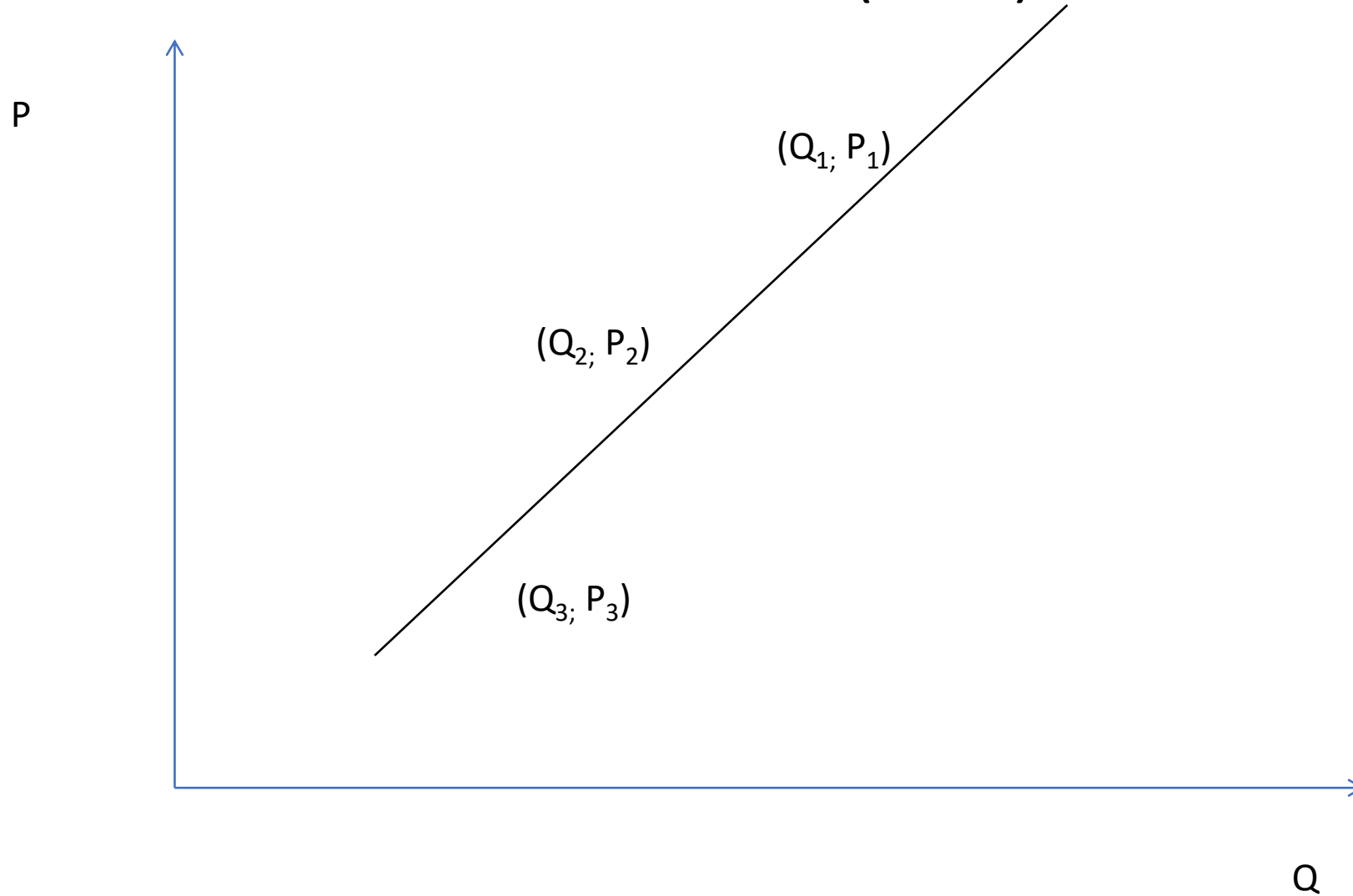
# Parámetros estructurales (1/6)

- El enfoque estándar reconoce que hay límites a la variación de una variable porque puede haber codeterminación

# Parámetros estructurales (2/6)

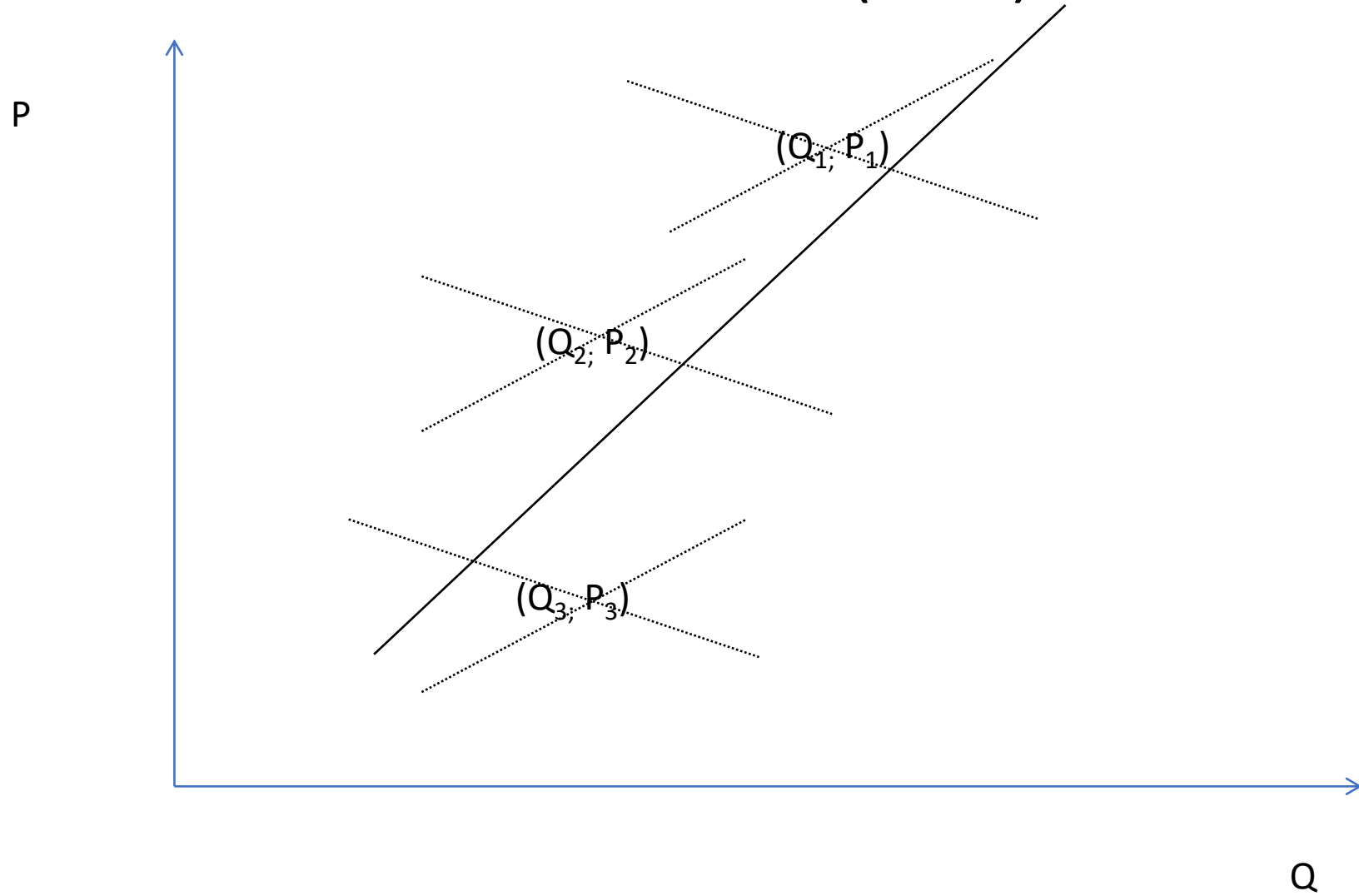


# Parámetros estructurales (3/6)





# Parámetros estructurales (4/6)



## Parámetros estructurales (5/6)

$$Q^d = \alpha_1 P^d + \alpha_2 Z^d + U$$

$$Q^o = \beta_1 P^o + \beta_2 Z^o + V$$

# Parámetros estructurales (6/6)

Forma reducida

$$P = \frac{b_2}{a_1 - b_1} Z^o - \frac{a_2}{a_1 - b_1} Z^d + \frac{U - V}{a_1 - b_1}$$

$$Q = \frac{a_1 b_2}{a_1 - b_1} Z^o - \frac{b_1 a_2}{a_1 - b_1} Z^d + \frac{a_1 U - b_1 V}{a_1 - b_1}$$

# El problema de identificación

- En el contexto demanda y oferta, el problema de identificación es determinar las restricciones que nos permiten variar el precio sin variar la cantidad;
- Si hay variables no observables, hay que controlar su variación para poder identificar

# Resumen

- Cuando hay interdependencia, necesitamos variables que puedan modificar las “causas” sin afectar la relación causa efecto. Estas son restricciones de exclusión;
- El efecto o parámetro causal se define a partir de la variación controlada de una variable en la ecuación estructural o comportamental
  - Si es linear, el efecto estimado es un parámetro estructural
  - Si no es linear, el parámetro estructural es el que nos permite caracterizar la relación causal
- Las exógenas tienen efectos directos e indirectos:
  - Ambos son causales
  - Los directos son estructurales
  - El total es el efecto en forma reducida

# Análisis de las políticas

El problema de las políticas (1/4)

$$Q^d = \alpha_1 P^d (1 + \tau) + \alpha_2 Z^d + U$$

$$Q^o = \beta_1 P^o + \beta_2 Z^o + V$$

# El problema de las políticas (2/4)

- Con una ecuación estructural yo puedo estimar el impacto de una política que nunca se implementó...
- ... a condición que se pueda interpretar en término de alguna variable que está en el modelo
- Si quiero ver el impacto de un impuesto:
  - Necesito los parámetros estructurales;
  - Necesito poder interpretar la policy a partir de algo que está en el modelo



## El problema de las políticas (3/4)

$$Q^* = \frac{\alpha_1(1+\tau)\beta_2}{\alpha_1(1+\tau)-\beta_1} Z^o - \frac{\alpha_1(1+\tau)\beta_2\alpha_2}{\alpha_1(1+\tau)-\beta_1} Z^d + \frac{\alpha_1(1+\tau)}{\alpha_1(1+\tau)-\beta_1} (V-U)$$

$$P^* = \frac{\beta_2}{\alpha_1(1+\tau)-\beta_1} Z^o - \frac{\alpha_2}{\alpha_1(1+\tau)-\beta_1} Z^d + \frac{U-V}{\alpha_1(1+\tau)-\beta_1}$$

# El problema de las políticas (4/4)

- No es necesario estimar TODO el modelo estructural, puede ser relevante medir únicamente la elasticidad de la demanda si nos interesa el excedente del consumidor
- A veces combinación de parámetros estructurales es suficiente (Marshallian maximum principle)

# Crítica de Lucas

$$Y = \alpha E[Y] + \beta X + U$$

- Si no incluimos las expectativas (racionales) y sus determinantes los parámetros no son policy invariant, porque modificarán el proceso de formación de las expectativas y a través de este el parámetro causal
- Si las políticas económicas hacen variar variables que antes eran fijas (ej. expectativas en el debate sobre la curva de Philips) no podemos hacer predicciones



“La historia nos cuenta como  
terminó el recorrido” (cit)

# Límites del programa de la Cowles

- Fue un éxito teórico
- Fue un fracaso empírico:
  - Los modelos eran subidentificados respecto a la teoría económica;
  - Los supuestos de identificación pocos plausibles;
  - Los parámetros inestables

# Respuestas (1)

- VAR:

$$Z_t = A_1 Z_{t-1} + A_2 Z_{t-2} + \cdots + A_q Z_{t-q} + \varepsilon_t$$

- Se puede transformar en un modelo estructural
  - Identificación usa el impacto de una “innovación” sobre la senda temporal de las variables de interés
- Aun en forma estructural no capturan un verdadero modelo macro dinámico
- La identificación vía series de tiempo no es menos arbitrarias

# Respuestas (2)

- Modelos estructurales más robustos, nace la microeconometría
  - Se toma en serio el problema de autoselección y de endogeneidad
  - Tenemos mucho más datos
- Problemas:
  - Instrumentos débiles
  - Problemas computacionales
  - Modelos que no tienen en consideración las políticas en vigor

# Estrategias recientes: calibración

- Avances en modelos recursivos, en teoría de juegos en EEG muestran modelos muy complejos que no se pueden formular como VAR;
- Uso de modelos complicados con calibración de parámetros a partir de micro-datos (non necesariamente compatibles con hp del modelo);
- Fundación empírica es débil, contrafactuales son muy criticados



# Estrategias recientes (Sensitivity analysis)

- $E[W \mid D=1]$  es la distribución de los sueldos de los que trabajan y están inscritos al sindicato
- $P(D=1 \mid Z)$  es el modelo de participación que depende de  $Z$ , y que obviamente nos permite determinar  $P(D=0 \mid Z)$
- Necesitamos saber  $E[W \mid D=0]$

$$E[W \mid Z] = E[W \mid D=1, Z] P(D=1 \mid Z) + E[W \mid D=0, Z] P(D=0 \mid Z)$$

Si  $E[W \mid D=0, Z] \in [w_l, w_u]$  puedo calcular espectro de valores

Trabajo similar es el non-paramétrico, que permite distinguir entre las implicaciones que dependen de las distribuciones y las de las formas funcionales

# Estrategias recientes: experimentos naturales

- Búsqueda de plausibles fuentes de variación exógena para poder usar el estimar de IV, por transparencia y replicabilidad;
- No hay énfasis en modelos estructurales, más bien simple intuición económica;
- Problema es la dificultad de comparar entre estudios;
- Énfasis en explicación más que predicción



Haciendo puentes

# Estructural vs Program Evaluation

- Enfoque à lo Cowles Foundation:
  - Causalidad es *ceteris paribus* dentro de un modelo (variación hipotética de un factor detrás de un resultado);
  - Coeficientes estructurales:
    - Causa de los efectos
    - Acumulabilidad
  - ¿Credibilidad? Problema de identificación
- Enfoque evaluación de programa
  - Efectos de las causas;
  - transparentes y más creíbles
  - Difícil de interpretar, solo intuición económica

# Un ingrediente clave

- El problema de la elección:
  - ¿por qué personas con características similares toman decisiones diferentes?
  - ¿Qué motiva las decisiones?
  - ¿cuál es la información disponible a la hora de tomar las decisiones?

# (Neyman-)Rubin model...

- Modelo de outcome potencial alternativo:

$$\begin{array}{c} Y^0 \\ Y^1 \end{array}$$

- Formulación en término de switching regression

$$Y(i) = D(i)(Y^1) + (1 - D(i))Y^0, \quad D(i) = \begin{cases} 0 \\ 1 \end{cases}$$

... es un caso particular del Modelo de Roy

- Modelo de outcome potencial alternativo:

$$\begin{matrix} Y^0 \\ Y^1 \end{matrix}$$

- En los problemas económicos hay alguien que hace una elección (que puede no ser el interesado, e.g. educación)

$$D(i) = \mathbf{1}(Y^1 - Y^0 > 0)$$

... es un caso particular del Roy extendido

- Costos de transición:

$$D(i) = \mathbf{1}(Y^1 - Y^0 - C) > 0$$

- El outcome es:

$$Y(i) = D(i)(Y^1) + (1 - D(i))Y^0$$

- Si hay incertidumbre:

$$D(i) = \mathbf{1}(E[Y^1 - Y^0 - C|\mathfrak{I}] > 0)$$

Cuidado que el conjunto de información del decisor y del econometrista no necesariamente coinciden



# Parámetros de interés

- El efecto tratamiento a nivel individual es:  $Y^1 - Y^0$
- No se puede medir por el *problema de la evaluación*: observo individuo solo en un estado
- $ATE = E[Y^1 - Y^0]$
- $TT = E[Y^1 - Y^0 | D = 1]$
- $TUT = E[Y^1 - Y^0 | D = 0]$
- Criterio voto:  $Pr(E[Y^1 - Y^0 - C | \mathfrak{Z}] > 0)$
- Beneficiados:  $Pr(Y^1 - Y^0 - C > 0)$
- $MRMI = E[Y^1 - Y^0 | E[Y^1 - Y^0 - C | \mathfrak{Z}] = 0]$

# Ejemplo de especificación

$$Y^1 = \mu_1(X) + U^1$$

$$Y^0 = \mu_0(X) + U^0$$

$$C = \mu_C(Z) + U^C$$

Inclusión de variables tiene justificación teórica. Distinción entre observables y no observables es clave. Instrumentos son las variables dentro de  $Z$  que no están en  $X$  ( $Z$  incluye  $X$  típicamente);

$$\begin{aligned} D &= \mathbf{1}(E[Y^1 - Y^0 - C | \mathfrak{I}] > 0) = \\ &= \mathbf{1}(E[\mu_1(X) - \mu_0(X) - \mu_C(Z) + U^1 - U^0 - U^C | \mathfrak{I}] > 0) = \\ &= \mathbf{1}(\mu_D(Z) > V) \end{aligned}$$

# Switching regression

$$Y = Y^0 + D(Y^1 - Y^0) = \mu_0 + D(\mu_1 - \mu_0 + U^1 - U^0) + U^0 = \alpha + \beta D + \varepsilon$$

- Si el efecto es igual para todos  $U^1 = U^0$  no hay problema de estimación;
- Si  $U^1 - U^0$  no está correlacionado con  $D$ , cualquier instrumento sirve para identificar el efecto promedio. En este caso, los agentes no tienen info sobre su ganancia idiosincrática;
- Si  $U^1 - U^0$  está en parte en  $\mathfrak{Z}$  entonces el instrumento no sirve para identificar el efecto promedio, y diferentes instrumentos dan diferentes estimadores

# LATE (1/3)

- Nos da la ganancia para los inducidos a cambiar de estado de acuerdo a un instrumento;
  - Es un estimador de variables instrumentales;
  - ¿Qué es el LATE?
- 
- Noten que los que  $D=1$  por el instrumento no tienen porque ser los que cambian de estado por la policy, y diferentes instrumentos inducen diferentes muestra de  $D=1$

## LATE (2/3)

- $Z$  es el instrumento, puede ser variado independientemente de  $X$  el efecto de  $Z$  es únicamente a través de  $D$ , que es la decisión de cambiar de estado;
- Supuestos:
  - Independencia de  $Z$  respecto a  $D, Y^0, Y^1$  (condicional a  $X$ )
  - $P(D|Z)$  es alguna función no degenerada, condicionadamente a  $X$
  - Monotonicidad: para todos, si consideramos dos valores  $z^1, z^2$ , vale  $D(z^1) \geq D(z^2)$

$$\text{LATE} = E[Y^1 - Y^0 | D(z^1) = 1, D(z^2) = 0]$$

# LATE (3/3)

- LATE no nos dice QUIEN cambia de estado. Con el mismo instrumento pero dos valores diferentes de  $z$  no podemos decir que el impacto sea el mismo;
- No nos permite hacer inferencia sobre policy en otro contexto o ex ante assessment

# Formalizando LATE dentro del Roy Model (1/2)

$$\begin{aligned}\mu_D(Z) &= E[\mu_1(X) - \mu_0(X) - \mu_C(Z) | \mathfrak{Z}] \\ V &= -E[U^1 - U^0 - U^C | \mathfrak{Z}]\end{aligned}$$

El contrafactual es generado por  $D = \mathbf{1}(\mu_D(Z) > V)$

- Separabilidad es clave para poder garantizar monotonicidad;
- Independencia de  $U^1, U^0, V$  respecto a  $Z$  condicionadamente a  $X$ ;
- La probabilidad de tratamiento  $P(\mu_D(z) > V) = F_V(\mu_D(Z))$  es un propensity score
  - Noten que  $U_D = F_V(V)$  es una variable casual uniforme sobre el soporte  $[0,1]$ , así que el percentil  $p$  es exactamente  $p$ ;
  - $P(Z) = F_V(\mu_D(Z))$  por lo tanto  $D = \mathbf{1}(\mu_D(Z) > V) = \mathbf{1}(F_V(\mu_D(Z)) > F_V(V)) = \mathbf{1}(P(Z) > U_D)$

# Formalizando LATE dentro del Roy Model (2/2)

- El propensity score nos dice quien se siente beneficiado con su información ex ante;

$$LATE(z^1, z^2) = E[Y^1 - Y^0 | D(z^1) = 1, D(z^2) = 0] = E[Y^1 - Y^0 | P(z^1) \leq U_D \leq P(z^2)]$$



# Surplus from treatment and MTE (1/3)

- La ganancia promedio bruta (sin costos) de mover al tratamiento los que tienen  $U_D < p$ , donde  $p$  es el percentil  $p$  del Propensity Score es:

$$E[Y^1 - Y^0 | P(Z) \geq U_D, P(Z) = p]$$

- Que por independencia es igual a

$$E[Y^1 - Y^0 | p \geq U_D]$$

- La ganancia bruta promedio en la población es

$$S(p) = E[Y^1 - Y^0 | p \geq U_D]p$$

# Surplus from treatment and MTE (2/3)

- El valor promedio del outcome se vuelve:

$$\begin{aligned} E[Y|P(Z) = p] &= E[Y^0 + (Y^1 - Y^0)\mathbf{1}(p \geq U_D)] \\ &= E[Y^0] + E[(Y^1 - Y^0)|p \geq U_D] = E[Y^0] + S(p) \end{aligned}$$

Al margen:

$$\frac{\partial E[Y|P(Z) = p]}{\partial p} = \frac{\partial S(p)}{\partial p} = E[(Y^1 - Y^0)|p = U_D] = MTE$$

## Surplus from treatment and MTE (3/3)

$$S(p_2) - S(p_1) = E[Y^1 - Y^0 | p_1 \leq U_D \leq p_2](p_2 - p_1)$$

Y desde la definición de LATE:

$$LATE(p_1, p_2) = \frac{S(p_2) - S(p_1)}{p_2 - p_1}$$

Y

$$E[Y | U_D \leq u_{D,k}] = E[Y^0] + \sum LATE(u_{D,j} - u_{d,j-1})(u_{D,j} - u_{d,j-1})$$

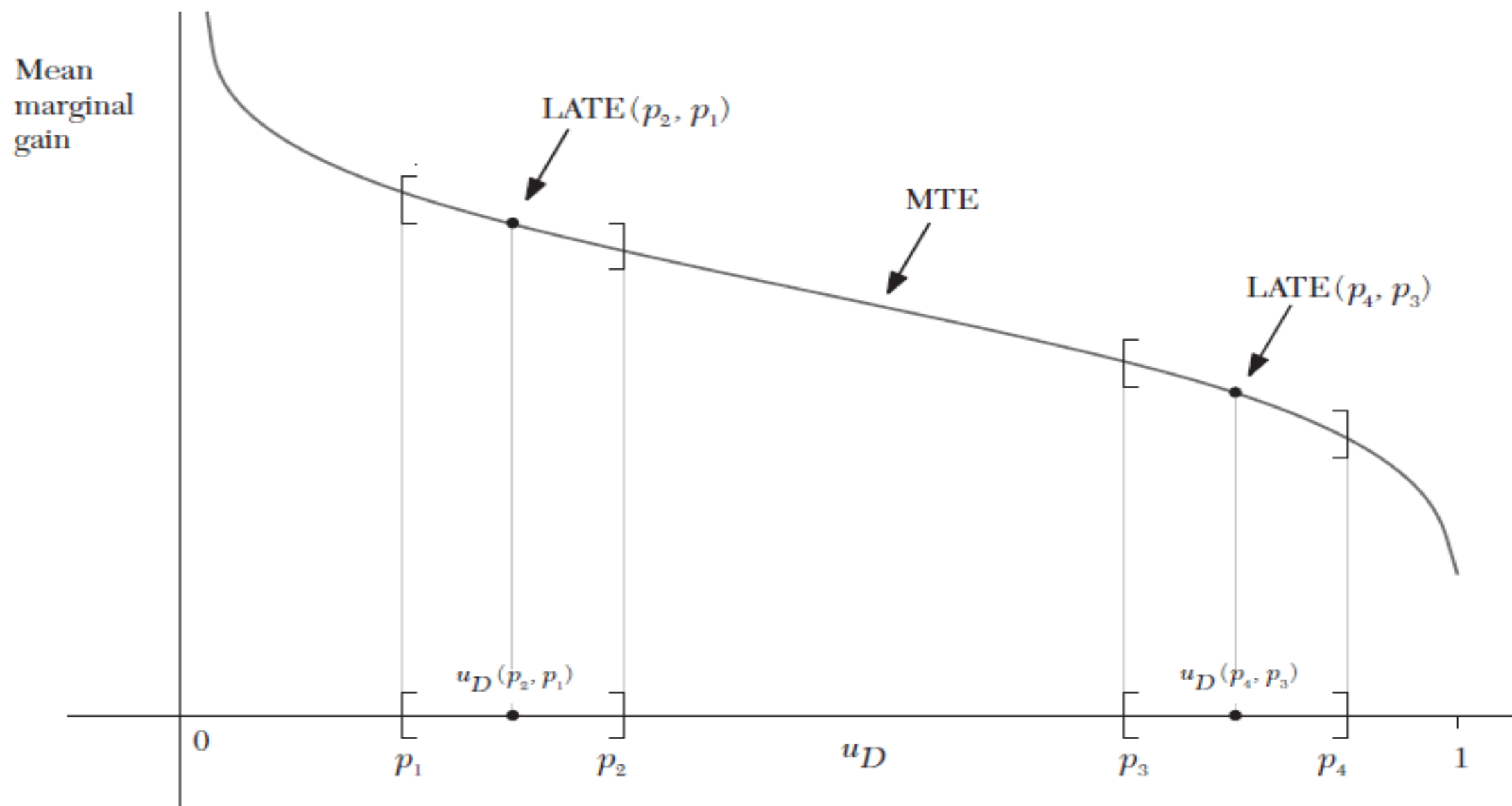


Figure 2. MTE as a Function of  $u_D$ : What Sections of the MTE Different Values of the Instruments and Different Instruments Approximate

# Noten

- MTE (y el LATE redefinido de esta manera ) es invariante respecto a  $Z$  y a cualquier cosa (incluyendo policy) que pase por modificar  $Z$ , es un parámetro estructural;
- MTE y el LATE redefinido de esta manera sirven para hacer inferencia sobre extensión de policy y ex ante assessment
- Cualquier efecto tratamiento es un promedio ponderado de MTE