



THE WORLD BANK



# المسار الفني

## الجلسة الرابعة

### المتغيرات المساعدة

كريستل م. ج. فرميرش

عمان 2009

Human Development  
Network

Latin American and  
the Caribbean Region

Finance, Private Sector,  
and Infrastructure  
Department

World Bank Institute  
Evaluation Group

## مثال لنبدأ به....

- لنفترض أننا نود تقييم برنامج تطوعي للتدريب الوظيفي
  - أي شخص عاطل عن العمل مؤهل للاشتراك
  - بعض الناس اختار التسجيل ("مجموعة المعالجة")
  - البعض الآخر اختار عدم التسجيل (مجموعة المقارنة")
  
- من الطرق البسيطة (وإن لم تكن الجيدة) لتقييم البرنامج:
  - مقارنة وضع مجموعة المعالجة قبل وبعد
  - مقارنة وضع مجموعتي المعالجة والمقارنة بعد التدخل
  - مقارنة وضع مجموعتي المعالجة والمقارنة قبل وبعد

# البرنامج التطوعي للتدريب الوظيفي

لنفترض أننا نقرر مقارنة النواتج الخاصة بالمشاركين مع النواتج الخاصة بغير المشاركين:  
(معادلة)

$$Y = a + B1D + B2x + \varepsilon$$

$D=1$  إذا شارك الشخص في تدريب

$D=0$  إذا لم يشارك الشخص في تدريب

$X=$  متغيرات الضبط (خارجية وقابلة للملاحظة)

**لماذا لا يعمل هذا بصورة سليمة؟ مشكلتان:**

➤ المتغيرات التي حذفناها (لعدة أسباب) ولكنها مهمة

➤ قرار المشاركة في التدريب من الداخل

# المسألة رقم 1: المتغيرات المحذوفة

- حتى في أي نموذج محكم سنفقد
- سمات "منسية": لم نعرف أنها مهمة
  - سمات معقدة للغاية بصورة يصعب قياسها

أمثلة:

- تنوع الموهبة ومستويات التحفيز
- مستويات مختلفة من المعرفة
- اختلاف تكاليف المشاركة
- تنوع مستوى الوصول إلى الخدمات

النموذج التام الصحيح:

$$y = a + y_1D + y_2x + y_3M + n$$

النموذج الذي نستخدمه:

$$Y = a + B_1D + B_2x + \varepsilon$$

## المسألة رقم 2: القرار الداخلي بالمشاركة

المشاركة هي متغير خاص بالقرارات؟ فهي تأتي من الداخل!

(أي يعتمد على المشاركين أنفسهم.)

$$y = \alpha + \beta_1 D + \beta_2 x + \varepsilon$$

$$D = \pi + \pi_2 M + \xi$$

$$\Rightarrow y = \alpha + \beta_1 (\pi + \pi_2 M + \xi) + \beta_2 x + \varepsilon$$

$$\Rightarrow y = \alpha + \beta_1 \pi + \beta_2 x + \beta_1 \pi_2 M + \beta_1 \xi + \varepsilon$$

إذن: في الحالتين: تنقصنا M، التي نحتاج إليها لتقدير النموذج الصحيح، ولكن لا يمكننا قياسها بالصورة الصحيحة

$$y = \alpha + \gamma_1 T + \gamma_2 x + \gamma_3 D + \eta$$

□ النموذج الصحيح هو:

$$y = \alpha + \beta_1 T + \beta_2 x + \varepsilon$$

□ النموذج المبسط:

□ هب أننا قدرنا أثر المعالجة  $\gamma_1$  ب  $\beta_{1,OLS}$

□ إذا كانت  $M$  متلازمة مع  $D$ ، ولم ندخل  $M$  في النموذج المبسط، فإن مقدر الحد على  $D$  سيلتقط جزءاً من أثر  $M$ . وسيحدث ذلك إلى الحد الذي تصبح فيه  $M$  و  $D$  متلازمتين.

□ وهكذا: فإن مقدر ال OLS وهو  $\beta_{1,OL}$  سيسجل أثر الخصائص الأخرى ( $M$ ) فضلاً عن أثر المعالجة.

□ وهذا يعني أن هناك اختلافاً بين  $\gamma_1$  و  $E(\beta_{1,OLS})$

القيمة المتوقعة لمقدر  $\beta_1$  ليست  $\gamma_1$ ، وهي الأثر الحقيقي للمعالجة

$\beta_{1,OLS}$  هي مقدر منحرف لأثر المعالجة  $\gamma_1$ .

$$y = \alpha + \gamma_1 T + \gamma_2 x + \gamma_3 D + \eta$$

$$y = \alpha + \beta_1 T + \beta_2 x + \varepsilon$$

النموذج الصحيح هو:

النموذج المبسط:



□ هذا يعني أن هناك اختلافاً بين  $\gamma_1$  و  $E(\beta_{1,OLS})$   
القيمة المتوقعة لمقدّر  $\beta_1$  ليست  $\gamma_1$ ، وهي الأثر الحقيقي للمعالجة  
 $\beta_{1,OLS}$  هي مقدر منحرف لأثر المعالجة  $\gamma_1$ .

□ لماذا حدث ذلك؟

■ أحد الشروط الأساسية ل OLS كي يكون BLUE قد انتهك:

□ بمعنى آخر  $E(\beta_{1,OLS}) \neq \gamma_1$  (مقدّر منحرف)

□ بل والأسوأ من ذلك  $E(\beta_{1,OLS}) \neq \gamma_1$  (مقدّر غير متنسق)

# ماذا نستطيع أن نفعل لكي نحل هذه المشكلة؟

$$y = \alpha + \gamma_1 T + \gamma_2 x + \gamma_3 D + \eta$$

$$y = \alpha + \beta_1 T + \beta_2 x + \varepsilon$$

- حاول تفكيك الارتباط بين T و  $\varepsilon$
- إ عزل التغير في D الذي لا يتلازم مع  $\varepsilon$  من خلال المتغير المحذوف M
- يمكننا عمل ذلك باستخدام متغير مساعد



# الفكرة الأساسية وراء المتغير المساعد

$$y = \alpha + \gamma_1 D + \gamma_2 x + \gamma_3 M + \eta$$

$$y = \alpha + \beta_1 D + \beta_2 x + \varepsilon$$

المشكلة الأساسية هي أن المترابطين (D , M)  $0 \neq$  □

إبحث عن متغير Z يلبي كلا الشرطين: □

1. يكون مترابطاً مع D: (D و Z)  $0 \neq$

--- تكون Z و D مترابطين، أو تتنبأ Z بجزء من D

2. لا يكون مرتبطاً بـ  $\varepsilon$ : ترابط ( Z و  $\varepsilon$  ) = 0

---في ذاتها، ليس لـ Z تأثير على y. والطريقة الوحيدة لكي تستطيع التأثير على

y هي لأنها تؤثر على D. فكل تأثيرات Z على y تمر من خلال D.

□ أمثلة على Z في حالة البرنامج التطوعي للتدريب الوظيفي؟

## التربيعات الأقل ذات المرحلتين (2SLS)

تذكر النموذج الأصلي مع العنصر الداخلي (D)

$$y = a + B_1D + B_2x + E$$

**خطوة 1:** قم بعمل انحدار للمتغير الداخلي (D) على المتغير المساعد (Z) وغيره من المتغيرات الداخلية  
انظر المعادلة الواردة في خطوة 1 أعلاه

- احتسب القيمة المتنبأ بها لـ (D) لكل ملاحظة: D
- حيث إن (z) و (x) لا يرتبطان بـ (E)، فإنهما لا يرتبطان بـ (D)
- ستحتاج متغير مساعد واحد لكل منحدر داخلي متوقع

Remember the original model with endogenous  $D$ :

$$y = \alpha + \beta_1 D + \beta_2 x + \varepsilon$$

**Step 1:** Regress the endogenous variable  $D$  on the instrumental variable(s)  $Z$  and other exogenous variables

$$D = \delta_0 + \delta_1 x + \theta_1 Z + \tau$$

- ▷ Calculate the predicted value of  $D$  for each observation:  $\hat{D}$
- ▷ Since  $Z$  and  $x$  are not correlated with  $\varepsilon$ , neither will be  $\hat{D}$ .
- ▷ You will need one instrumental variable for each potentially endogenous regressor

## التربيعات الأقل ذات المرحلتين (2SLS)

- خطوة 2:** قم بعمل انحدار لـ  $(y)$  على المتغير المتنبأ به  $(D)$  وغيره من المتغيرات الخارجية الأخرى (انظر المعادلة أعلاه)
- ملاحظة: الأخطاء المعيارية للمرحلة الثانية من طريقة التربيعات الأقل المعتادة (OLS) يجب تصحيحها لأن  $(D)$  ليست منحدرًا ثابتًا.
  - عملياً: استخدم أمر (STATA ivreg) الذي يقوم بالخطوتين في مرة واحدة ويظهر الأخطاء المعيارية الصحيحة.
  - بديهيًا: باستخدام  $(z)$  مكان  $(D)$ ، فإننا نحرر  $(D)$  من ترابطها مع  $(\varepsilon)$ .
  - يمكن أن يتبين (في ظروف محددة) أن (IV) تعطي مقدراً متوافقاً لـ  $(yI)$  [نظرية العينة الكبيرة]

# استخدامات المتغيرات المساعدة

- التزامن:  $X$  و  $Y$  تسببان بعضهما البعض
  - الأداة  $X$
- المتغيرات المحذوفة:  $X$  تلتقط أثر المتغيرات الأخرى التي حُذفت
  - الأداة  $X$  مع متغير غير مترابط مع المتغير(ات) المحذوفة
- خطأ القياس:  $X$  لم يتم قياسها بدقة
  - الأداة  $X$

# أين نجد متغيرات مساعدة؟

□ البحث عن واحدة – بقوة!

□ إنشاء واحدة مع معلومات

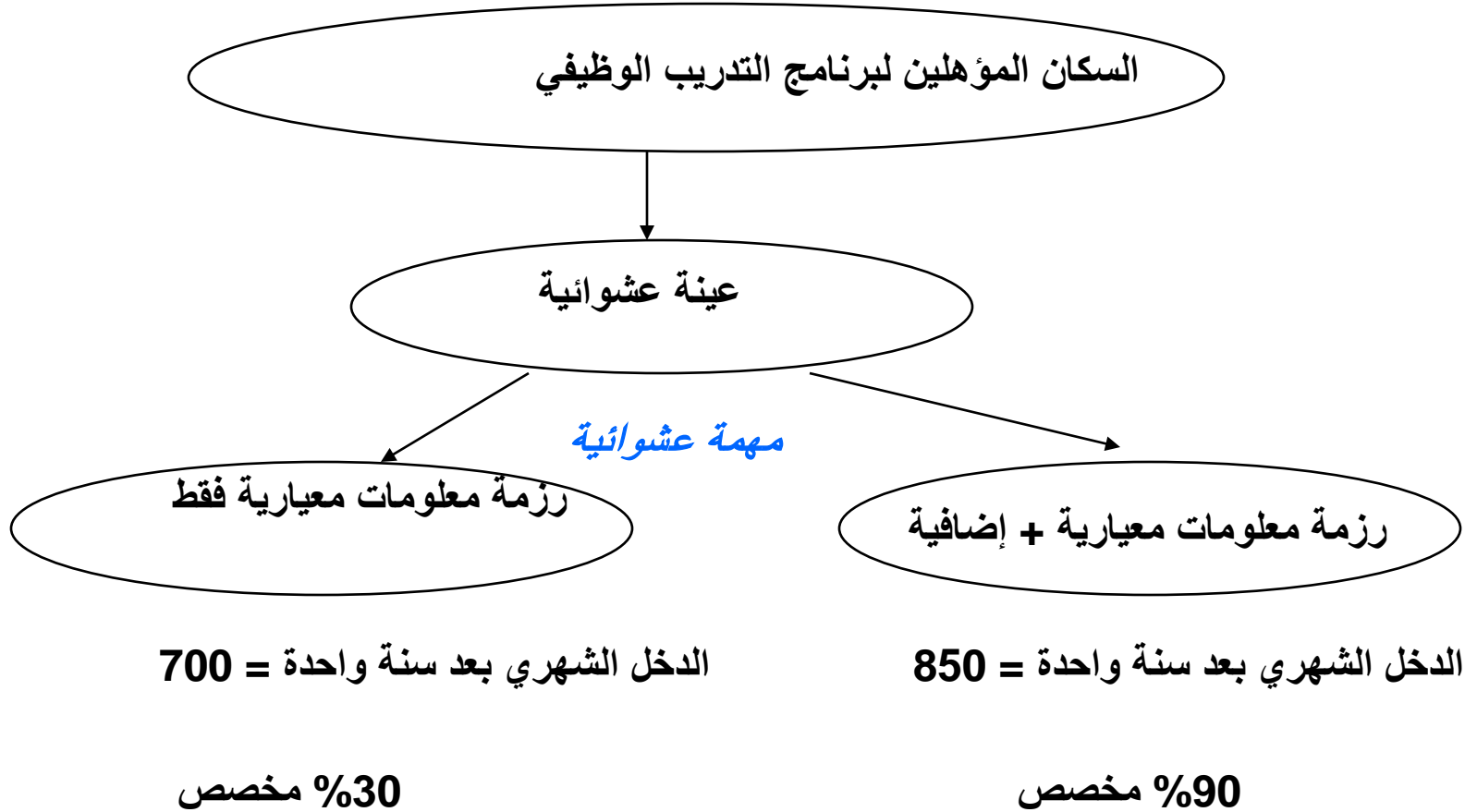
■ إذا كان كل واحد مؤهل للمشاركة في المعالجة

■ إلا أن للبعض معلومات أكثر من الآخرين

□ الإحتمالية الأكبر للمشاركة تكون للذي لديه معلومات أكثر

■ تقديم "معلومات إضافية" على أساس عشوائي

# برنامج التدريب الوظيفي الطوعي



سؤال : ما هو أثر برنامج التدريب الوظيفي؟

رزمة معلومات معيارية فقط

رزمة معلومات معيارية + إضافية



الدخل الشهري بعد سنة واحدة = 700

الدخل الشهري بعد سنة واحدة = 850

30% مخصص

90% مخصص

**سؤال: ما هو أثر برنامج التدريب الوظيفي؟**

• الفرق بين المجموعة "المطلعة جيدا" والمجموعة "غير المطلعة بصورة جيدة"

.....

• مصححة لمعدل التخصيص التفاضلي

.....

• عمليا : الأثر = .....



# الربط بصيغة التقدير

## □ المرحلة 1 :

- قم بإجراء ارتداد للمشاركة على التدريب، على نموذج، لمعرفة ما إذا تلقى الشخص رزمة معلومات إضافية (نموذج خطي)
- قم باحتساب قيمة المشاركة المتنبأ بها

## □ المرحلة 2:

- قم بإجراء ارتداد للأجور على قيمة المشاركة المتنبأ بها



# الإدارة الذاتية للمدارس في نيبال

## □ الهدف هو تقييم

- أ. الإدارة الذاتية للمدارس من قبل المجتمع
- ب. شهادات المدارس

## □ البيانات

- يمكننا دمج نحو 1000 مدرسة في التقييم
- كل مجتمع يختار بحرية المشاركة أو عدم المشاركة
- شهادات المدارس تقوم بها منظمات غير حكومية
- هناك مدرسة واحدة فقط في كل مجتمع

□ المهمة : تصميم تنفيذ البرنامج بحيث يتم تقييمه – اقتراح وسيلة تقييم

# الإدارة الذاتية للمدارس في نيبال

		التدخلات: ب. شهادات المدارس تدخل من المنظمات غير الحكومية		
		نعم	لا	المجموع
متغير مساعد للتدخل أ : تزور المنظمة غير الحكومية المجتمع لتبليغهم بالإجراءات لنقل المدرسة لإدارة المجتمع	نعم	300	300	600
	لا	200	200	400
	المجموع	500	500	1000

# تذكرة وكلمة تنبيه...

## □ المترابطتان $(Z, \varepsilon) = 0$

- مشكلة! إذا كانت المترابطتان  $(Z, \varepsilon) \neq 0$  “أداة سيئة”
- العثور على أداة جيدة صعب!
- استخدم كلاً من النظرية والمنطق السليم للعثور على واحدة!
- يمكننا التفكير في تصاميم تعطي أدوات جيدة.

## □ المترابطتان $(Z, T) \neq 0$

- “أدوات ضعيفة”: الترابط بين  $Z$  و  $D$  يجب أن يكون قوياً بما فيه الكفاية.
- إن لم يكن، يظل التحيز كبيراً حتى في العينات كبيرة الحجم.