Clase 22: Resultados Potenciales en Regresiones y Ejemplos Haciendo Economía I Econ 2205

Ignacio Sarmiento-Barbieri

Universidad de los Andes

October 31, 2023

Anuncios

- Las referencias para esta clase siguen siendo:
 - 1 Mastering Metrics de Angrist y Pischke (cap 1)
 - 2 Mostly Harmless Econometrics de Angrist y Pischke (cap 2)
 - 3 Causal Inference: The Mixtape de Cunningham (cap 4) (disponible online en su pagina web)
- ▶ Quiz sobre Resultados Potenciales → Martes 7 de Noviembre

Plan para hoy

Anuncios

- Modelo de Resultados Potenciales Recap
 - Ejemplo: El Experimento STAR
- Análisis de Regresión en Experimentos
 - Ejemplo: El Experimento STAR (cont.)
 - Ejemplo: Aprendiendo sobre resultados de prueba de HIV

▶ Una variable binaria de tratamiento (e.g.,va al hospital, no va):

$$D_i = \{0, 1\} \tag{1}$$

▶ Resultado (Outcome) para el individuo *i* (e.g., salud):

$$Y_i$$
 (2)

- ▶ El modelo tiene varios nombres...
 - Splawa-Neyman marco de resultados potenciales
 - ▶ Rubin modelo causal
 - ► Neyman-Rubin (Splawa-Neyman-Rubin)

- Pregunta de investigación: D_i afecta a Y_i ?
- ▶ Para cada individuo *i*, hay dos **resultados potenciales** (con D_i binario)
 - ▶ Y_{1i} si D_i = 1 Resultado de i si va al hospital
 - ► Y_{0i} si D_i = 0 Resultado de i si no va al hospital
- La diferencia entre los dos resultados nos da el efecto causal del tratamiento de ir al hospital, i.e.,

$$\tau_i = Y_{1i} - Y_{0i} \tag{3}$$

Sarmiento-Barbieri (Uniandes) Clase 22 October 31, 2023 4/38

Problemas

► Esta ecuación simple:

$$\tau_i = \mathbf{Y}_{1i} - \mathbf{Y}_{0i} \tag{4}$$

- ▶ nos lleva al **problema fundamental de inferencia causal**.
- ▶ Nunca podemos observar simultáneamente Y_{1i} and Y_{0i} .



Source: https://tinyurl.com/yv3v5tum

Solución Propuesta

- Comparamos
 - resultado para los que van al hospital $(Y_{1i} | D_i = 1)$
 - resultados para los que **no** van al hospital $(Y_{0i} | D_i = 0)$

$$E\left[Y_i \mid D_i = 1\right] - E\left[Y_i \mid D_i = 0\right] \tag{5}$$

• que nos da la diferencia observada en los resultados de salud

Solución Propuesta

Entonces

$$E[Y_i \mid D_i = 1] - E[Y_i \mid D_i = 0] = E[Y_{1i} \mid D_i = 1] - E[Y_{0i} \mid D_i = 0]$$
(6)

Haciendo un poco de matemática creativa

$$= E[Y_{1i} \mid D_i = 1] - E[Y_{0i} \mid D_i = 1] + E[Y_{0i} \mid D_i = 1] - E[Y_{0i} \mid D_i = 0]$$
(7)

Solución Propuesta

Que nos dice esto entonces?

$$E[Y_i \mid D_i = 1] - E[Y_i \mid D_i = 0] =$$
(8)

$$E[Y_{1i} | D_i = 1] - E[Y_{0i} | D_i = 1] + E[Y_{0i} | D_i = 1] - E[Y_{0i} | D_i = 0]$$
(9)

Efecto Promedio sobre los tratados (ATT) ©

Solución Propuesta

Que nos dice esto entonces?

$$E[Y_i \mid \mathbf{D}_i = 1] - E[Y_i \mid \mathbf{D}_i = 0] = \tag{10}$$

$$=\underbrace{E\left[\mathbf{Y}_{1i}\mid\mathbf{D}_{i}=1\right]-E\left[\mathbf{Y}_{0i}\mid\mathbf{D}_{i}=1\right]}_{\text{Efecto Promedio sobre los tratados (ATT) }\odot}+\underbrace{E\left[\mathbf{Y}_{0i}\mid\mathbf{D}_{i}=1\right]-E\left[\mathbf{Y}_{0i}\mid\mathbf{D}_{i}=0\right]}_{\text{Sesgo de Selección }\odot}$$
(11)

Ejemplo: Tamaño de la clase y aprendizaje de estudiantes

- Uno de los insumos más costoso en la función de producción educativa es el tamaño de la clase:
- ► Clases mas pequeñas → mas maestros
- Fundamental entender el beneficio de tamaño de la clase más pequeño
- Estudios **observacionales** (no experimentales) sugieren que existe poca o ninguna relación entre el tamaño de la clase y el aprendizaje de los estudiantes.
- ▶ De ser cierto, esto implicaría que se puede ahorrar dinero contratando menos maestros sin la consecuente reducción en el rendimiento.

Ejemplo: Tamaño de la clase y aprendizaje de estudiantes

- Como formalizamos esta preocupación en nuestro modelo?
- Evaluación observacional de programas

$$E[Calificacion_i | Clase Pequeña_i = 1] - E[Calificacion_i | Clase Pequeña_i = 0] = (12)$$

Ejemplo: Tamaño de la clase y aprendizaje de estudiantes

- ► Como formalizamos esta preocupación en nuestro modelo?
- Evaluación observacional de programas

$$E[Calificacion_i | Clase Pequeña_i = 1] - E[Calificacion_i | Clase Pequeña_i = 0] = (12)$$

$$\underbrace{E\left[\text{Calificacion}_{1i} \mid \text{Clase Pequeña}_{i} = 1\right] - E\left[\text{Calificacion}_{0i} \mid \text{Clase Pequeña}_{i} = 1\right]}_{\text{Efecto causal promedio } i.e., \, \overline{\tau}} + \underbrace{E\left[\text{Calificacion}_{0i} \mid \text{Clase Pequeña}_{i} = 1\right]}_{\text{Efecto causal promedio } i.e., \, \overline{\tau}}$$

(13)

11/38

$$E\left[\text{Calificacion}_{0i} \mid \text{Clase Pequeña}_{i} = 1\right] - E\left[\text{Calificacion}_{0i} \mid \text{Clase Pequeña}_{i} = 0\right]$$
(14)

Sesgo de Seleccion

► Si las clases pequeña atrae/selecciona individuos que en promedio, tienen calificaciones menores en ausencia de clase pequeñas, entonces tenemos selección negativa () () ()

Sarmiento-Barbieri (Uniandes) Clase 22 October 31, 2023

Experimentos

▶ Como hacen los experimentos para resolver el sesgo de selección?

Experimentos

- Como hacen los experimentos para resolver el sesgo de selección?
- Los experimentos rompen el link entre los resultados potenciales y el tratamiento

Experimentos

- Como hacen los experimentos para resolver el sesgo de selección?
- ▶ Los experimentos rompen el link entre los resultados potenciales y el tratamiento
- ► En otras palabras, asignando aleatoriamente D_i hace que D_i sea independiente de que resultado observamos (es decir Y_{1i} o Y_{0i})

$$\{\mathbf{Y}_{1i}, \mathbf{Y}_{0i}\} \perp D_i \tag{15}$$

- La asignación aleatoriaimplica que los grupos de tratamiento y control salen de la misma población subyacente.
- ▶ Son iguales en todos los sentidos, incluyendo sus $E[Y_{0i}]$

Experimentos

Diferencia de medias con asignación aleatoria D_i

$$E[Y_i \mid \mathbf{D}_i = 1] - E[Y_i \mid \mathbf{D}_i = 0] = \tag{16}$$

Experimentos

Diferencia de medias con asignación aleatoria D_i

$$E[Y_i \mid D_i = 1] - E[Y_i \mid D_i = 0] =$$
 (16)

$$= E[Y_{1i} \mid D_i = 1] - E[Y_{0i} \mid D_i = 0]$$
(17)

Experimentos

▶ Diferencia de medias con asignación aleatoria D_i

$$E[Y_i \mid D_i = 1] - E[Y_i \mid D_i = 0] =$$
 (16)

$$= E[Y_{1i} \mid D_i = 1] - E[Y_{0i} \mid D_i = 0]$$
(17)

$$= E[Y_{1i} \mid D_i = 1] - E[Y_{0i} \mid D_i = 1]$$
(18)

Experimentos

Diferencia de medias con asignación aleatoria D_i

$$E[Y_i \mid D_i = 1] - E[Y_i \mid D_i = 0] =$$

$$(16)$$

$$= E[Y_{1i} \mid D_i = 1] - E[Y_{0i} \mid D_i = 0]$$
(17)

$$= E[Y_{1i} | D_i = 1] - E[Y_{0i} | D_i = 1]$$
(18)

$$= E[Y_{1i} - Y_{0i} \mid D_i = 1]$$
 (19)

Experimentos

▶ Diferencia de medias con asignación aleatoria D_i

$$E[Y_i \mid \mathbf{D}_i = 1] - E[Y_i \mid \mathbf{D}_i = 0] = \tag{16}$$

$$= E[Y_{1i} \mid D_i = 1] - E[Y_{0i} \mid D_i = 0]$$
(17)

$$= E[Y_{1i} \mid D_i = 1] - E[Y_{0i} \mid D_i = 1]$$
(18)

$$= E[Y_{1i} - Y_{0i} \mid D_i = 1]$$
 (19)

$$= E\left[\tau_i \mid D_i = 1\right] \tag{20}$$

Experimentos

Diferencia de medias con asignación aleatoria D_i

$$E[Y_i \mid \mathbf{D}_i = 1] - E[Y_i \mid \mathbf{D}_i = 0] = \tag{16}$$

$$= E[Y_{1i} \mid D_i = 1] - E[Y_{0i} \mid D_i = 0]$$
(17)

$$= E[Y_{1i} | D_i = 1] - E[Y_{0i} | D_i = 1]$$
(18)

$$= E[Y_{1i} - Y_{0i} \mid D_i = 1]$$
 (19)

$$= E\left[\tau_i \mid D_i = 1\right] \tag{20}$$

$$= E\left[\tau_i\right] \tag{21}$$

La asignación aleatoria de D_i rompe el sesgo de selección

Experimentos: Asignación aleatoria de tratamiento

- La clave para evitar el sesgo de selección es la asignación aleatoria
- Esta nos da:

$$E[Y_{0i} | D_i = 0] = E[Y_{0i} | D_i = 1]$$
 (22)

- ▶ El grupo de control nos da un buen contrafactual para la media del grupo de tratamiento
- ► En otras palabras no hay sesgo de selección: = $E[Y_{0i} | D_i = 1] E[Y_{0i} | D_i = 0] = 0$

Sarmiento-Barbieri (Uniandes)

Experimentos: Asignación aleatoria de tratamiento

- Un beneficio adicional de la aleatorización
- ▶ El efecto de tratamiento promedio (ATE) es ahora representativo de la población, y no solo del promedio del grupo de tratamiento (ATT)

$$E\left[\tau_{i} \mid D_{i} = 1\right] = E\left[\tau_{i} \mid D_{i} = 0\right] = E\left[\tau_{i}\right]$$
(23)

- Estudios **observacionales** (no experimentales) sugieren que existe poca o ninguna relación entre el tamaño de la clase y el aprendizaje de los estudiantes.
- ▶ De ser cierto, esto implicaría que se puede ahorrar dinero contratando menos maestros sin la consecuente reducción en el rendimiento.
- ▶ Problema: estudiantes más débiles suelen ser colocados en clases más pequeñas

- Estudios **observacionales** (no experimentales) sugieren que existe poca o ninguna relación entre el tamaño de la clase y el aprendizaje de los estudiantes.
- ▶ De ser cierto, esto implicaría que se puede ahorrar dinero contratando menos maestros sin la consecuente reducción en el rendimiento.
- Problema: estudiantes más débiles suelen ser colocados en clases más pequeñas
- Un experimento aleatorio puede solucionar esto: aleatorizar estudiantes a clases de diferentes tamaños
- ▶ Esta es la idea del proyecto STAR de Tennessee

- ► El experimento STAR fue ambicioso e influyente
- ▶ El experimento asignó a los estudiantes a uno de tres tratamientos:
 - 1 Clases **pequeñas** con 13 17 niños,
 - 2 Clases **regulares** con 22 25 estudiantes y un profesor asistente tiempo parcial **(grupo control)**,
 - 3 Clases **regulares** con un profesor asistente de tiempo completo
- Costó alrededor de \$12 millones y se implementó para una cohorte de niños de jardín de infantes en 1985/86.
- ▶ El estudio duró cuatro años e involucró a unos 11.600 niños.

Variable	Pequeña	Regular	Regular con Asistente	P-value conjunto
41 0 0	0.45	0.40	0.50	0.00
Almuerzo Gratis	0.47	0.48	0.50	0.09
Blanco/Asiático	0.68	0.67	0.66	0.26
Edad en 1985	5.44	5.43	5.42	0.32
Tamaño de clase	15.10	22.40	22.80	0.00
Percentil Resultado de Prueba	54.70	48.90	50.00	0.00

Fuente: Adaptación tabla 2.2.1 MHE

1 Anuncios

- 2 Modelo de Resultados Potenciales Recap
 - Ejemplo: El Experimento STAR

- 3 Análisis de Regresión en Experimentos
 - Ejemplo: El Experimento STAR (cont.)
 - Ejemplo: Aprendiendo sobre resultados de prueba de HIV

- La tabla previa estima/compara el efecto del tratamiento (ATE) haciendo diferencias de medias .
- Podemos hacer lo mismo con una regresión.
- Específicamente, si hacemos la regresión del "outcome" (percentil de la prueba) en una variable dummy para cada grupo de tratamiento.

► Asumimos que el efecto del tratamiento es constante ("homogeneo")

$$Y_{1i} - Y_{0i} = \tau \quad \forall i \tag{24}$$

Entonces podemos escribir

$$Y_i = Y_{0i} + D_i (Y_{1i} - Y_{0i})$$
 (25)

como

$$Y_{i} = \underbrace{\alpha}_{E[Y_{0i}]} + D_{i} \underbrace{\tau}_{Y_{1i} - Y_{0i}} + \underbrace{\eta_{i}}_{Y_{0i} - E[Y_{0i}]}$$
(26)

$$Y_i = \alpha + D_i \tau + \eta_i \tag{27}$$

El Experimento STAR

Variable

Clase Pequeña

Regular + Asistente

Blanco/Asiatico

Mujer

Almuerzo Gratis

Efecto Fijo Escuela

Fuente: Adaptación tabla 2.2.2 MHE

El Experimento STAR

Variable	(1)			
Cl. P. ~	4.00			
Clase Pequeña	4.82 (2.19)			
Regular + Asistente	0.12			
Blanco/Asiatico	(2.23)			
Mujer				
Almuerzo Gratis				
Efecto Fijo Escuela	No			
Euente: Adaptación tabla 2.2.2 MHE				

Fuente: Adaptación tabla 2.2.2 MHE

El Experimento STAR

Variable	(1)	(2)
Clase Pequeña	4.82 (2.19)	5.37 (1.26)
Regular + Asistente	0.12	0.29
Blanco/Asiatico	(2.23)	(1.13)
Mujer		
Almuerzo Gratis		
Efecto Fijo Escuela	No	Si

Fuente: Adaptación tabla 2.2.2 MHE

El Experimento STAR

▶ Si se asigno aleatoriamente y esto elimina el sesgo de selección por qué la regresión incluye controles?

El Experimento STAR

Si se asigno aleatoriamente y esto elimina el sesgo de selección por qué la regresión incluye controles?

Los controles juegan dos roles en los análisis de regresión de datos experimentales.

Il El diseño experimental STAR utilizó asignación aleatoria condicional: la asignación a clases de diferentes tamaños fue aleatoria dentro de las escuelas, pero no entre escuelas.

Start with raw data. Correlation between X and Y: -0.039

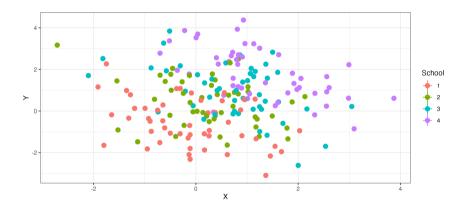
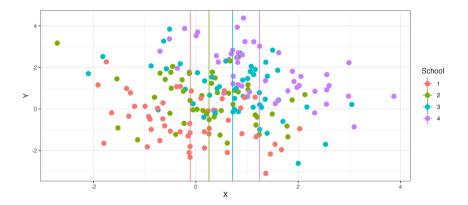


Figure out any between-School differences in X



Remove all between-School differences in X

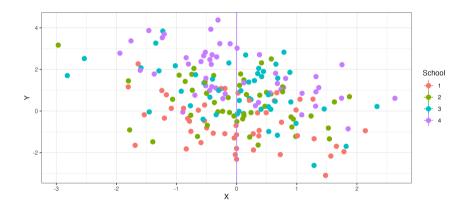
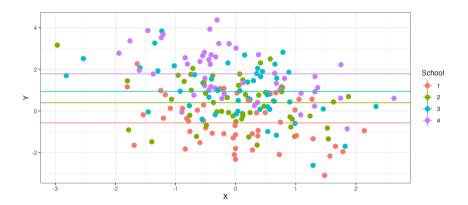
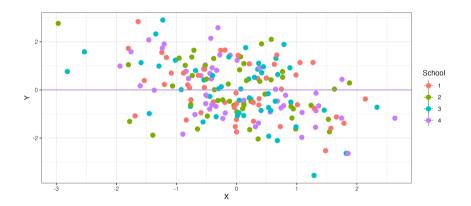


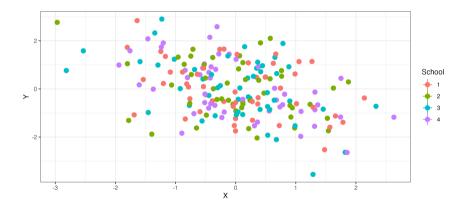
Figure out any between-School differences in Y



Remove all between-School differences in Y



Analyze what's left! Within-School Correlation Between X and Y: -0.429



Análisis de Regresión en Experimentos

El Experimento STAR

Variable	(1)	(2)	(3)
Clase Pequeña	4.82	5.37	5.36
	(2.19)	(1.26)	(1.21)
Regular + Asistente	0.12	0.29	0.53
	(2.23)	(1.13)	(1.09)
Blanco/Asiatico			8.35
			(1.35)
Mujer			4.48
			(0.63)
Almuerzo Gratis			-13.15
			(0.77)
Efecto Fijo Escuela	No	Si	Si

Fuente: Adaptación tabla 2.2.2 MHE

El Experimento STAR

Si se asigno aleatoriamente y esto elimina el sesgo de selección por qué la regresión incluye controles?

Los controles juegan dos roles en los análisis de regresión de datos experimentales.

- Il El diseño experimental STAR utilizó asignación aleatoria condicional: la asignación a clases de diferentes tamaños fue aleatoria dentro de las escuelas, pero no entre escuelas.
- 2 Como vimos en la tabla $Cov(D_i, X_i) = 0$, sin embargo la inclusión de control aumenta la precisión.

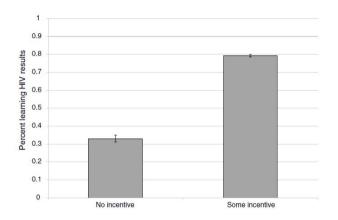
1 Anuncios

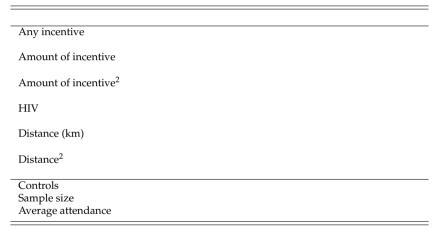
- 2 Modelo de Resultados Potenciales Recap
 - Ejemplo: El Experimento STAR

- 3 Análisis de Regresión en Experimentos
 - Ejemplo: El Experimento STAR (cont.)
 - Ejemplo: Aprendiendo sobre resultados de prueba de HIV

- La idea es que si las personas conocen su estado sobre VIH tomarían precauciones en caso de ser positivo, reduciendo la tasa de infección
- ▶ Si usamos datos observacionales, el problema que surge es que las personas se auto seleccionan a aprender sobre su estado de salud.
- Individuos que se testean también son mas probables que tengan conductas menos riesgosas.
- ▶ Para romper esta dependencia es necesario un experimento

- ▶ Thornton fue hasta Malawi rural e hizo un experimento
- ▶ Fueron puerta a puerta ofreciendo pruebas de HIV gratuitos
- Les dieron aleatoriamente vouchers (o no) entre \$ 1 y \$ 3
- La gente los podía cambiar un vez que visitaban el centro de pruebas mas cercanos





Nota: *** Significativamente diferente de cero a un nivel de confianza del 99 por ciento. ** Significativamente diferente de cero a un nivel de confianza del 95 por ciento. * Significativamente diferente de cero a un nivel de confianza del 90 por ciento.

	1
Any incentive	0.431***
•	(0.023)
Amount of incentive	
Amount of incentive ²	
HIV	-0.055*
	(0.031)
Distance (km)	
2	
Distance ²	
Controls	Yes
Sample size	2,812
Average attendance	0.69

Nota: *** Significativamente diferente de cero a un nivel de confianza del 99 por ciento. ** Significativamente diferente de cero a un nivel de confianza del 95 por ciento. * Significativamente diferente de cero a un nivel de confianza del 90 por ciento.

	1	2
Any incentive	0.431***	0.309***
•	(0.023)	(0.026)
Amount of incentive		0.091***
		(0.012)
Amount of incentive ²		
HIV	-0.055*	-0.052
	(0.031)	(0.032)
Distance (km)	` /	, ,
Distance ²		
Controls	Yes	Yes
Sample size	2,812	2,812
Average attendance	0.69	0.69

Nota: *** Significativamente diferente de cero a un nivel de confianza del 99 por ciento. ** Significativamente diferente de cero a un nivel de confianza del 95 por ciento. * Significativamente diferente de cero a un nivel de confianza del 90 por ciento.

	1	2	3
Any incentive	0.431***	0.309***	0.219***
•	(0.023)	(0.026)	(0.029)
Amount of incentive		0.091***	0.274***
		(0.012)	(0.036)
Amount of incentive ²			-0.063***
			(0.011)
HIV	-0.055*	-0.052	-0.05
	(0.031)	(0.032)	(0.032)
Distance (km)			
Distance ²			
Distance			
Controls	Yes	Yes	Yes
Sample size	2,812	2,812	2,812
Average attendance	0.69	0.69	0.69

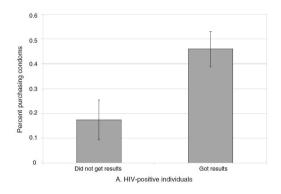
Nota: *** Significativamente diferente de cero a un nivel de confianza del 99 por ciento. ** Significativamente diferente de cero a un nivel de confianza del 95 por ciento. * Significativamente diferente de cero a un nivel de confianza del 90 por ciento.

	1	2	3	4
Any incentive	0.431***	0.309***	0.219***	0.220***
	(0.023)	(0.026)	(0.029)	(0.029)
Amount of incentive		0.091***	0.274***	0.274***
		(0.012)	(0.036)	(0.035)
Amount of incentive ²			-0.063***	-0.063***
			(0.011)	(0.011)
HIV	-0.055*	-0.052	-0.05	-0.058*
	(0.031)	(0.032)	(0.032)	(0.031)
istance (km)				-0.076***
				(0.027)
Distance ²				0.010**
				(0.005)
Controls	Yes	Yes	Yes	Yes
Sample size	2,812	2,812	2,812	2,812
Average attendance	0.69	0.69	0.69	0.69

Nota: *** Significativamente diferente de cero a un nivel de confianza del 99 por ciento. ** Significativamente diferente de cero a un nivel de confianza del 95 por ciento. * Significativamente diferente de cero a un nivel de confianza del 90 por ciento.

	1	2	3	4	5
Any incentive	0.431***	0.309***	0.219***	0.220***	0.219 ***
	(0.023)	(0.026)	(0.029)	(0.029)	(0.029)
Amount of incentive		0.091***	0.274***	0.274***	0.273***
		(0.012)	(0.036)	(0.035)	(0.036)
Amount of incentive ²			-0.063***	-0.063***	-0.063***
			(0.011)	(0.011)	(0.011)
HIV	-0.055*	-0.052	-0.05	-0.058*	-0.055*
	(0.031)	(0.032)	(0.032)	(0.031)	(0.031)
Distance (km)				-0.076***	
				(0.027)	
Distance ²				0.010**	
				(0.005)	
Controls	Yes	Yes	Yes	Yes	Yes
Sample size	2,812	2,812	2,812	2,812	2,812
Average attendance	0.69	0.69	0.69	0.69	0.69

Nota: *** Significativamente diferente de cero a un nivel de confianza del 99 por ciento. ** Significativamente diferente de cero a un nivel de confianza del 95 por ciento. * Significativamente diferente de cero a un nivel de confianza del 90 por ciento.



Dependent variables:	Bought condoms	Number of condoms bought
	1	2
Got results	-0.022	-0.193
	(0.025)	(0.148)
Got results x HIV	0.418***	1.778**
	(0.143)	(0.564)
HIV	-0.175**	-0.873
	(0.085)	(0.275)
Controls	Yes	Yes
Sample size	1,008	1,008
Mean	0.26	0.95