

ABDUL LATIF JAMEEL
Poverty Action Lab



TRANSLATING RESEARCH INTO ACTION

Evaluación de Programas Sociales

Universidad de Los Andes
Julio 12–16, 2010

Educación Ejecutiva del
Abdul Latif Jameel Poverty Action Lab



Evaluación de Programas Sociales Programa de Presentaciones 2010

12 de julio del 2010 – 16 de julio del 2010

Lunes, 12 de julio

9:00 a.m. – 9:30 a.m.	Registro
9:30 a.m.	Palabras de Apertura
9:30 a.m. – 11:00 a.m.	Presentación 1: ¿Que es una Evaluación? <i>Presentador: Dan Levy</i>
11:00 a.m. – 11:15 a.m.	Pausa para café
11:15 a.m. – 1:00 p.m.	Discusión en grupos y caso 1
1:00 p.m. – 2:00 p.m.	Almuerzo
2:00 p.m. – 3:30 p.m.	Presentación 2: Marco lógico, indicadores y medición del impacto <i>Presentador: Juan Saavedra</i>
3:30 p.m. – 4:00 p.m.	Pausa para café
4:00 p.m. – 6:00 p.m.	Trabajo en grupo

Martes, 13 de julio

9:30 a.m. – 11:15 a.m.	Discusión en grupos del caso 2
11:15 a.m. – 11:30 a.m.	Pausa para café
11:30 a.m. – 1:00 p.m.	Presentación 3: ¿Por qué Aleatorizar? <i>Presentador: Dan Levy</i>
1:00 p.m. – 2:30 p.m.	Almuerzo
2:30 p.m. – 4:00 p.m.	Presentación 4: ¿Como Aleatorizar? <i>Presentador: Francisco Gallego</i>
4:00 p.m. – 4:15 p.m.	Pausa para café
4:15 p.m. – 5:00 p.m.	Trabajo en grupo
5:00 p.m. – 6:00 p.m.	Discusión en grupos del caso 3



Miércoles, 14 de julio

9:00 a.m. – 11:00 a.m.	Trabajo en grupo
11:00 a.m.-11:15 a.m.	Pausa para café
11:15 a.m. – 1:30 p.m.	Presentación 5: Tamaño de la muestra y Cálculos de Poder <i>Presentador: Francisco Gallego</i>
1:30 p.m. – 2:30 p.m.	Almuerzo
2:30 p.m. – 4:00 p.m.	Presentación 6: Implementar una evaluación <i>Presentador: Ernesto Schargrodsky</i>
4:00 p.m. – 4:15 p.m.	Pausa para café
4:15 p.m. – 6:00 p.m.	Trabajo en grupo

Jueves, 15 de julio

9:00 a.m. – 9:30 a.m.	Trabajo en grupo
9:30 a.m. – 11:00 a.m.	Discusión en grupos del caso 4 (Desparasitación)
11:00 a.m. – 11:15 a.m.	Pausa para café
11:15 a.m. – 12:45 p.m.	Presentación 7: Analizando Datos <i>Presentador: Juan Saavedra</i>
12:45 p.m. – 2:00 p.m.	Almuerzo
2:00 p.m. – 3:30 p.m.	Presentación 8: Evaluaciones desde el comienzo hasta el final <i>Presentador: Claudia Martínez</i>
3.30 p.m. – 3:45 p.m.	Pausa para café
3:45 p.m. – 6:00 p.m.	Trabajo en grupo
8:00 p.m. – 10:30 p.m.	Cena <i>Ubicación: hotel de la Opera Calle 10 No.5-72 La Candelaria</i>



Viernes, 16 de julio

9:00 a.m. – 10:00 a.m.	Finalización de Proyectos Grupales
10:00 a.m. – 11:00 p.m.	Presentaciones Grupales
11:00 a.m. – 11:15 a.m.	Pausa para café
11:15 a.m. – 1:00 p.m.	Presentaciones Grupales (continuación) Comentarios de los Instructores
1:00 p.m. – 2:00 p.m.	Almuerzo
2:00 p.m. – 3:30 p.m.	Presentaciones Grupales (continuación) Comentarios de los Instructores
3:30 p.m.	Comentarios de Clausura

Curso de Educación Ejecutiva

Universidad de Los Andes

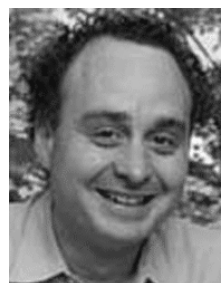
Bogotá, 12 al 16 de Julio de 2010

Francisco Gallego (Pontificia Universidad Católica de Chile)



Francisco Gallego es Doctor en Economía por el MIT y Profesor Asistente del Instituto de Economía en la Pontificia Universidad Católica de Chile (PUC). Se especializa en temas de Desarrollo Económico, Economía de la Educación y Economía Política. Ha trabajado en evaluaciones educativas enfocadas en los efectos de la entrega de información sobre las decisiones de los padres de los estudiantes. Es Director Científico de J-PAL LatAm.

Dan Levy (Harvard University)



Dan Levy es profesor de Políticas Públicas en la Escuela de Gobierno John F. Kennedy en la Universidad de Harvard, y director de capacitación del Abdul Latif Jameel Poverty Action Lab. Actualmente está trabajando en la evaluación de un programa de transferencias condicionales en Jamaica y a un conjunto de intervenciones educacionales en Burkina Faso. También provee asistencia técnica y entrenamiento en el Ministerio de Desarrollo Social de México (Sedesol) sobre evaluaciones de impacto de diversos programas sociales.

Claudia Martínez (Universidad de Chile)



Claudia Martínez A. es Profesora Asistente de Economía en la Universidad de Chile e Investigadora en el Centro de Microdatos de esa misma Universidad. Claudia también es miembro del *International Advisory Panel* para la evaluación de nuevos programas en Chile. Su investigación se enfoca en evaluaciones de la distribución de poder en el interior del hogar, migración, educación y Políticas Públicas en países en desarrollo. Proyectos previos y actuales incluyen: el estudio del impacto de controles en remesas y ahorros de familias transnacionales de El Salvador, y expectativas educativas en Chile.



Juan Saavedra (Universidad de Los Andes, Colombia)



Juan E. Saavedra es Doctor en Políticas Públicas por la Universidad de Harvard. Sus principales áreas de interés son la Economía Laboral, Economía de la Educación, Finanzas Públicas y Econometría Aplicada. Sus investigaciones exploran, por ejemplo, los efectos de las tutorías en el desempeño académico, el uso del tiempo y la delincuencia juvenil. Otros trabajos exploran las causas de heterogeneidad en los sistemas de vouchers, los mecanismos por los cuales éstos mejoran los resultados académicos en escuelas secundarias, y las implicancias que esto tiene en el diseño de Políticas Públicas educativas, todo esto apoyado por evidencia proveniente de experimentos aleatorios “naturales” en Colombia. Actualmente es profesor asistente de Políticas Públicas en la Escuela de Gobierno Alberto Lleras Camargo, en la Universidad de Los Andes, Colombia.

Ernesto Schargrodsky (Universidad Torcuato Di Tella, Argentina)



Ernesto Schargrodsky recibió su Doctorado en Economía en la Universidad de Harvard en 1998. Es el Decano de la Escuela de Negocios de la Universidad Torcuato Di Tella en Buenos Aires, Argentina. Ha sido Profesor Visitante Edward Laroque Tinker en la Universidad de Stanford e Investigador Visitante De Fortabat en la Universidad de Harvard. Su investigación incluye estudios sobre el efecto de la presencia policial en el crimen, el análisis de apoyo popular a privatizaciones, la relación entre salarios a funcionarios públicos y la corrupción, el efecto del uso de sistemas electrónicos para el pago de programas de seguridad social, y el efecto de la entrega de títulos de propiedad de tierra en áreas marginales. Sus trabajos se ha publicado en: American Economic Review, Journal of Political Economy, Quarterly Journal of Economics, Journal of Law and Economics, Journal of Public Economics, y Journal of Development Economics entre otros.

ABDUL LATIF JAMEEL

Poverty Action Lab



TRANSLATING RESEARCH INTO ACTION



Caso 1: Las Mujeres gobernando sus Villas

Midiendo los efectos de las cuotas políticas
Reflexionando sobre mediciones y resultados

Este caso se basa en "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India," por Raghavendra Chattopadhyay y Esther Duflo (2004a), *Econometrica* 72(5), 1409-1443.

J-PAL agradece a los autores su permiso para utilizar esta publicación

En 1992, India modificó su constitución federal con la intención de descentralizar el poder de los estados hacia los concejos rurales o Gram Panchayats (Concejos de Aldea). Actualmente, los Concejos de Aldeas, planifican e implementan los programas de desarrollo que desean llevar a cabo y asignan recursos del presupuesto a proyectos de inversión. También se requiere que los estados reserven un tercio de las plazas en los Concejos de Aldeas y un tercio de los puestos de Presidente de Concejo de Aldea para mujeres. En la mayoría de los estados, los ciclos de asignación de plazas reservadas entre los distintos Concejos de Aldea se determinan aleatoriamente. Esto crea la oportunidad de evaluar rigurosamente el impacto de las cuotas en la política y el gobierno: ¿Las políticas públicas difieren cuando hay un mayor número de mujeres en el gobierno? ¿Se reflejan las prioridades de las mujeres en las políticas públicas? Debido a que la aleatorización formó parte del diseño de este programa del Gobierno de India, la planificación de la evaluación se centró en recolectar los datos necesarios para medir el impacto. Entre las cuestiones a considerar se incluye: qué datos recolectar, qué instrumentos de recolección de datos usar, y para qué tamaño de muestra era necesario planificar.

Empoderando al Panchayati Raj

Los Panchayats, o Concejos de Aldea, tienen una larga tradición en las zonas rurales de India. Una asamblea (*yat*) de cinco (*panch*) ancianos, elegidos por la comunidad, se reúnan a mediar disputas entre individuos o aldeas. En los tiempos modernos, los Panchayats han sido formalizados dentro de las instituciones del auto-gobierno local.

El ímpetu por formalizarlos surgió de los líderes de la independencia, quienes abogaban por un gobierno descentralizado. Gandhi favorecía la idea que las aldeas se auto-gobernaran, un sistema en donde cada aldea sería “auto sustentable y capaz de gestionar sus propios asuntos”. El Primer Ministro Nehru concordó, otorgando a los Panchayats un “gran poder” para que las aldeas pudiesen “tener un verdadero swarji (auto-gobierno) en sus propias aldeas.”

Por tanto, el Artículo 40 de la Constitución India –adoptado durante la Independencia– estipula que los estados deben asegurar que los Panchayats “funcionen como unidades autónomas del gobierno”. Las guías de implementación recomiendan un sistema de tres niveles, siendo el concejo de aldea (gram panchayat) la base de la estructura. ¹La mayoría de los estados siguieron esta guía y dirección, lo cual culminó en la formalización de los Panchayats a principios de los 1950s. Sin embargo, en la década de 1960, al no tener un poder real y apoyo político y financiero del gobierno federal, los Panchayats desaparecieron de la mayoría de los estados. No revivieron hasta la década de 1990.

El renacimiento de los Panchayat se hizo mediante una nueva modificación a la constitución. En 1992, se promulgó la enmienda 73^o que exigía a los estados establecer el sistema de tres niveles de Panchayati Raj y tener elecciones de Panchayat cada cinco años. Los miembros del concejo son electos por voto popular y representan a cada distrito electoral. Luego, los miembros del concejo eligen entre ellos un presidente llamado

¹ Concejos de Aldea, llamados Gram Panchayats o GPs, constituyen la unidad básica de los Panchayat Raj. Los Presidentes de los Concejos elegidos por los miembros del concejo, son miembros del Concejo de subdistrito (*panchayat samiti*). En la cima del sistema se encuentra el concejo de distrito (*zilla parishad*) conformado por los presidentes de los concejos de subdistritos.

Caso 1: Reflexionando sobre mediciones y resultados

“pradhan”. Las decisiones se toman por mayoría de votos y el pradhan no tiene poder de veto. Pero al ser el único concejero a tiempo completo, el pradhan ejerce un poder efectivo en el concejo.

La enmienda 73 tenía el propósito de descentralizar la provisión de bienes y servicios públicos esenciales para el desarrollo de las áreas rurales. A los estados se les instruyó delegar su poder, de manera que los Panchayats planificaran e implementaran programas de desarrollo locales. Los fondos aún provienen del gobierno central pero ya no están sujetos a programas específicos. En cambio, ahora son los Concejos de Aldea quienes deciden qué programas implementar y cómo distribuir su presupuesto. Los Concejos de Aldeas pueden escoger programas en 29 áreas específicas, incluyendo servicios de bienestar social (por ejemplo, asistencia para viudas, personas de tercera edad, maternidad, cuidado prenatal y salud de los niños) y obras públicas (por ejemplo, acceso a agua potable, caminos, vivienda, edificios comunitarios, electricidad, irrigación y educación).

Empoderando a las mujeres en el Panchayati Raj

Los Concejos de Aldea son grandes y diversos. En Bengala occidental, por ejemplo, cada uno representa a 12 aldeas y aproximadamente a 10,000 personas, quienes difieren en términos de religión, casta y por supuesto, género. La voz y poder político también varían de acuerdo a las identidades de los grupos. Si las preferencias de políticas públicas varían de acuerdo a la identidad del grupo y si los autores de políticas públicas influyen en la selección de políticas, entonces los grupos sub representados en la política y el gobierno pueden ser marginados, ya que los Concejos de Aldeas pueden ignorar las prioridades de estos grupos. Existían temores de que los Concejos de Aldeas al poseer nuevas responsabilidades, reprimirían las prioridades de los grupos tradicionalmente marginados, como las mujeres. Para evitar esto, la enmienda 73 incluyó dos mandatos para asegurar que las inversiones reflejasen las necesidades de todas las personas en el Concejo de Aldea.

El primer mandato busca asegurar el aporte de la comunidad. Para que las inversiones del GP reflejen las prioridades de la comunidad, primero los miembros del Concejo deben tener conocimiento de estas prioridades. Por lo tanto, se requiere que los Concejos de Aldea celebren una asamblea general cada seis meses o una vez al año, para informar a la comunidad sobre las actividades del periodo anterior y presentar la propuesta de presupuesto a la comunidad para su ratificación. Adicionalmente, se requiere que los Presidentes del Concejo, establezcan horarios de oficinas regulares que permitan a los constituyentes presentar formalmente sus peticiones y quejas. Ambos requerimientos permiten a los constituyentes articular sus preferencias en materia de políticas públicas.

El segundo mandato busca asegurar la representación de las mujeres en el concejo. Se requiere que los estados reserven al menos un tercio de todos los puestos en el Concejo y cargos de Presidente de Concejo para mujeres. Además de esto, los estados deben asegurar que los puestos reservados para las mujeres sean “asignados por rotación a los distintos constituyentes en un Panchayat”. En otras palabras, se debe asegurar que los puestos en el Concejo y el cargo de presidente de Concejo se roten de manera igualitaria entre y dentro de los Concejos de Aldeas.

La aleatorización de cuotas en la India: ¿Qué podemos aprender?

A su equipo de evaluación se le otorga la oportunidad de medir el impacto de las cuotas de puestos políticos para mujeres en los Concejos de Aldeas. Su evaluación debe considerar todas las dimensiones en que las cuotas para mujeres generan un cambio en las comunidades locales en India. ¿Qué datos va a recolectar? ¿Qué instrumentos va a usar? ¿De qué tamaño va a ser la muestra?

El primer paso es comprender la política de reserva de puestos. ¿Cuáles eran las necesidades? ¿Cuáles son las ventajas y desventajas de la política de cuotas? ¿Qué podemos aprender de la misma?

Tema de Discusión 1: Cuotas de Género en los Concejos de Aldea

1. ¿Cuáles son los principales objetivos del Concejo de Aldea?
2. Las mujeres están sub-representadas en la política y el gobierno. Sólo el 10 por ciento de los miembros de la asamblea nacional de la India son mujeres, en comparación a un 17 por ciento a nivel mundial.

¿Tiene alguna importancia que las mujeres estén sub-representadas? ¿Por qué si y por qué no?

3. ¿Qué buscaban lograr los creadores de la enmienda 73 al introducir cuotas para mujeres?

Usualmente, las cuotas de puestos en base al género, han estado acompañadas por aumentos dramáticos en la representación de las mujeres en la política. Por ejemplo, Ruanda pasó del puesto 24 en la clasificación de “mujeres en parlamento” al primer lugar (con un 49 por ciento) luego de la introducción de cuotas en 1996. Se han observado cambios similares en Argentina, Burundi, Costa Rica, Iraq, Mozambique y Sudáfrica. De hecho, 17 de los primeros 20 países con más representación femenina en el parlamento tienen un sistema de cuotas.

Imagine que su equipo es el parlamento nacional de un país que está decidiendo si debe o no adoptar un sistema de cuotas para mujeres en el parlamento nacional. Divida a su equipo aleatoriamente en dos partidos; uno a favor y uno en contra de las cuotas.

4. Debata las ventajas y desventajas de las cuotas. Al final del debate, debería tener una lista de las ventajas y desventajas de las cuotas.
5. ¿Que evidencia recolectaría para estudiar el punto de vista de cada partido?

¿Qué datos recolectar?

Primero, se necesita claridad respecto al probable impacto del programa. Los datos se recolectarán sobre la base de las dimensiones que usted piense que serán afectadas por el programa. ¿Cuáles son las principales áreas en las que la política de cuotas debe ser evaluada? ¿En qué áreas espera ver una diferencia como resultado de las cuotas?

¿Cuáles son todos los efectos posibles de las cuotas?

Tema de Discusión 2: Utilizando un marco lógico para delinear los resultados de interés intermedios y finales

1. Hagan una lluvia de ideas de los posibles efectos de las cuotas, tanto positivos como negativos.
Sugerencia : Use sus respuestas del Tópico de Discusión 1 como punto de partida.
2. Para cada efecto potencial en su lista, haga un listado de los indicador(es) que usaría para tal efecto. Por ejemplo, si dice que las cuotas van a afectar la participación política de las mujeres, el indicador podría ser "cantidad de mujeres que asisten a la Asamblea General."
3. Supongan que tienen todos los recursos y dinero disponibles en el mundo, y que podrían recolectar datos sobre cada uno de esos indicadores en comunidades con cuotas y sin cuotas, y compararlos. ¿Cuántos indicadores recolectaría?

Los resultados múltiples son difíciles de interpretar, mejor defina una hipótesis.

Las cuotas políticas para mujeres podrían afectar de distintas formas y direcciones a un gran número de resultados. Por ejemplo, podría mejorar el suministro de agua potable y empeorar el suministro de irrigación. Sin una hipótesis *ex-ante* sobre la dirección en que las diferentes variables pueden ser afectadas por la política de cuotas, será bastante difícil darle sentido a cualquier resultado que encontremos. Piense en lo siguiente: si tomara 500 aldeas, y las asignara de forma aleatoria en su computadora a un grupo "tratamiento" y un grupo "control", y luego hiciera una regresión para determinar si las aldeas se ven diferentes a lo largo de los 100 resultados distintos, ¿esperaría ver algunas diferencias entre éstas? ¿Tendría algún sentido interpretar esos resultados *ex-post*?

Lo mismo sucede en este caso: si usted presenta un informe frente a la comisión, explicando que las cuotas para mujeres tuvieron un efecto sobre algunas variables y no otras, ¿Qué se puede hacer con eso? ¿Cómo van a saber que las diferencias no se deben al azar en vez de la política? Es fundamental presentar una hipótesis clara respecto a cómo las cuotas cambian la creación de las políticas públicas, lo que le permitirá realizar predicciones sobre que resultados se verán afectados.

Tema de Discusión 2 (continúa):

4. ¿Cuál podría ser un ejemplo de una hipótesis clave que se pondría a prueba? Elija una.
5. ¿Qué indicadores o combinaciones de indicadores usaría para poner a prueba su hipótesis clave?

Use un marco lógico para delinear resultados finales e intermedios

Una buena manera de identificar resultados importantes es presentar su teoría del cambio, esto es, hacer un marco lógico que vincule la intervención- paso a paso- con los resultados claves finales.

Tema de Discusión 2 (continúa):

6. ¿Cuál es una posible cadena de resultados en el caso de las cuotas políticas?
7. ¿Cuáles son los principales pasos críticos que se necesitan para obtener los resultados finales? ¿Cuáles son las condiciones necesarias que deben alcanzarse en cada paso?
8. Para cada paso en el marco lógico, ¿qué variables debería tratar de obtener?
9. Utilizando los resultados y condiciones, realice un marco lógico, vinculando la intervención y los resultados finales.

ABDUL LATIF JAMEEL

Poverty Action Lab



TRANSLATING RESEARCH INTO ACTION



Caso 2: Aprendiendo a Leer Evaluaciones

Evaluación de la campaña Read India. Cómo leer y evaluar evaluaciones

Este estudio de caso se basa en “Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in India”, de Abhijit Banerjee (MIT), Rukmini Banerjee (Pratham), Esther Duflo (MIT), Rachel Glennerster (J-PAL) y Stuti Khemani (Banco Mundial)

J-PAL agradece a los autores la autorización para utilizar su trabajo

¿Por qué Learn to Read (L2R) (*Aprende a Leer*)?

En una encuesta de gran escala realizada en 2004, Pratham descubrió que solo el 39% de los niños entre 7 y 14 años en el área rural de Uttar Pradesh podían leer y comprender un cuento simple, y que alrededor de 15% no podían reconocer ni siquiera una letra.

Durante este período, Pratham desarrollaba el módulo “Learn-to-Read” (L2R) de su campaña Read India (*Lee India*). L2R incluía una pedagogía única que enseñaba habilidades básicas de alfabetización, combinado con un esfuerzo de organización local para reclutar voluntarios dispuestos a enseñar.

Este programa permitió a la comunidad participar de manera más directa en la educación de los niños, a través de reuniones de aldeas, donde el personal de Pratham compartió información sobre el estado de alfabetización en el pueblo y los derechos de los niños a la educación. En estas reuniones, Pratham identificó a los miembros de la comunidad que estaban dispuestos a enseñar. Los voluntarios asistieron a una sesión de capacitación acerca de la pedagogía, después de la cual pudieron realizar clases de lectura (después de la escuela) para los niños, utilizando materiales diseñados y proporcionados por Pratham. El personal de Pratham realizó visitas ocasionales a estos campamentos para asegurar que las clases se estaban realizando y para proporcionar capacitación adicional según fuera necesario.

¿Funcionó el proyecto Learn to Read?

¿Funcionó el programa “Learn to Read” de Pratham? ¿Qué se requiere para que midamos si un programa funcionó, si tuvo un impacto?

En general, preguntarse si un programa funciona es preguntarse si el programa logra su objetivo de cambiar ciertos resultados para sus participantes, y que esos cambios no estén provocados por otros factores o eventos que ocurren al mismo tiempo. Para mostrar que el programa *provoca* los cambios observados, necesitamos mostrar de manera simultánea que si el programa no hubiera sido implementado, los cambios observados no habrían ocurrido. ¿Pero cómo sabemos *lo que habría ocurrido*? Si el programa ocurrió, ocurrió. Medir *lo que habría ocurrido* requiere ingresar a un mundo imaginario en el que el programa *nunca se aplicó a estos participantes*. Los resultados de estos mismos participantes en este mundo imaginario se denominan el *contrafactual*. Dado que no podemos observar el verdadero contrafactual, lo mejor que podemos hacer es estimarlo o buscar replicarlo.

El desafío clave de una evaluación de impacto es construir o replicar el contrafactual. Normalmente hacemos esto mediante la selección de un grupo de personas que se parecen a los participantes lo más posible, pero que no participaron en el programa. Este grupo se denomina el grupo de comparación. Ya que deseamos poder decir que fue el programa y no algún otro factor lo que provocó los cambios en los resultados, es importante que la única diferencia entre el grupo de comparación y los participantes sea que el grupo de comparación no participó en el programa. Entonces calculamos el “impacto” como la diferencia observada al final del programa entre los resultados del grupo de comparación y los de los participantes del programa.

La precisión de la estimación de impacto depende directamente de qué tan exitoso es el grupo de comparación en replicar al contrafactual. Si el grupo de comparación representa de forma deficiente al contrafactual, el impacto (en la mayoría de los casos) se calcula de forma deficiente. Por lo tanto, el método utilizado para seleccionar el grupo de comparación es una decisión clave en el diseño de cualquier evaluación de impacto.

Eso nos devuelve a nuestras preguntas: ¿Funcionó el proyecto “Learn to Read”? ¿Cuál fue el impacto en los niveles de lectura de los niños?

En este caso, la intención del programa es “mejorar los niveles de lectura de los niños” y el nivel de lectura es la medida de resultado. Entonces, cuando nos preguntamos si el proyecto “Learn to Read” funcionó, estamos preguntándonos si mejoró el nivel de lectura de los niños. El impacto es la diferencia entre los niveles de lectura después de que los niños hayan tomado las clases de lectura y el que habría sido su nivel de lectura si las clases nunca se hubieran realizado.

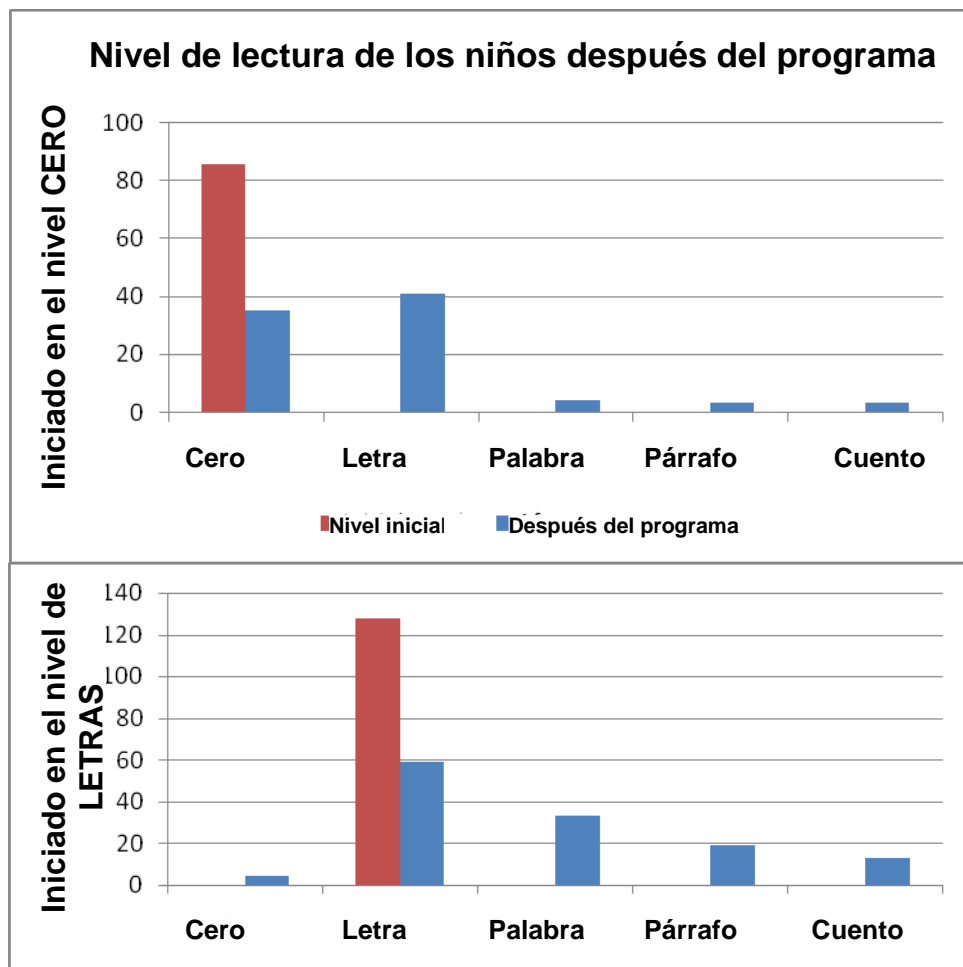
¿Qué grupos de comparación podemos utilizar? Los siguientes expertos ilustran diferentes métodos de evaluación de impacto.

Estimando el impacto del proyecto “Learn to Read”

Método 1:

Comunicado de prensa: Read India ayuda a los niños a aprender a leer.

Pratham celebra el éxito de su programa “Learn to Read”, parte de la iniciativa “Read India”. El programa ha hecho avances importantes en su objetivo de mejorar los índices de alfabetización de los niños a través de mejores materiales de aprendizaje, métodos pedagógicos y lo más importante, voluntarios comprometidos. El logro del programa “Learn to Read” (L2R) demuestra que un plan de estudios bien planificado y galvanizado por la movilización de la comunidad puede producir ganancias importantes. Los gastos masivos del gobierno en almuerzos escolares y construcción de escuela no han logrado resultados similares. En menos de un año, los niveles de lectura de los niños que se inscribieron en los campamentos de L2R mejoraron de manera considerable.



Justo antes del inicio del programa, la mitad de estos niños no podía reconocer palabras en Hindu, muchos no reconocían nada. Sin embargo, después de sólo algunos meses en las clases de lectura de Pratham, más de la mitad mejoró en al menos un nivel de lectura, con un número importante capaz de reconocer palabras y varios capaces de leer párrafos completos y cuentos. *En promedio, el indicador del alfabetismo de estos estudiantes mejoró en aproximadamente un nivel de lectura completo durante este período.*

Tema de discusión 1:

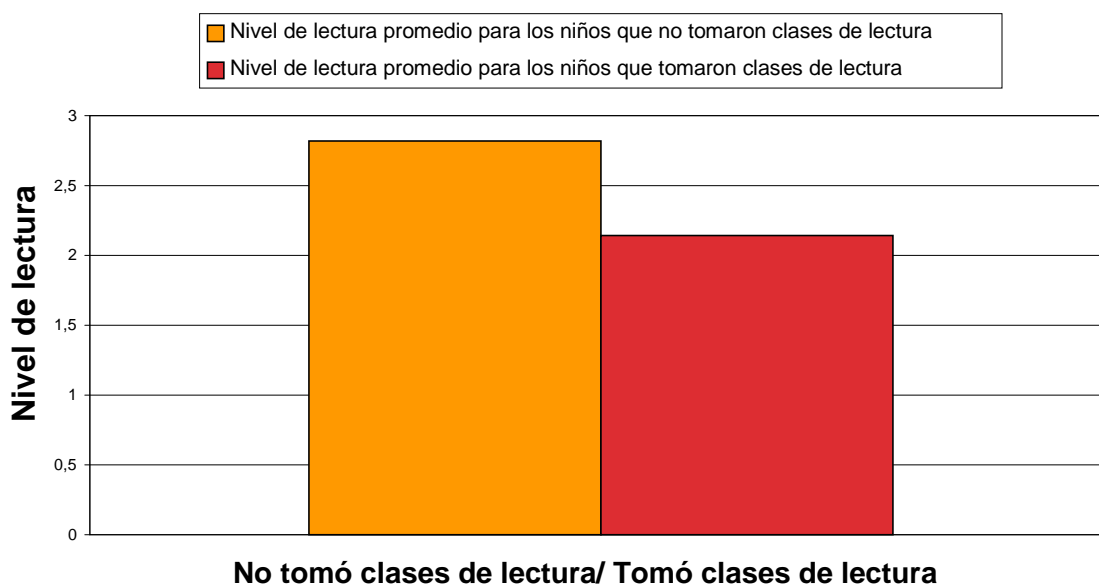
1. ¿Qué tipo de evaluación implica este comunicado de prensa?
2. ¿Cómo intenta a estimar el contrafactual este método? ¿Qué usa para el grupo de comparación?
3. ¿Cuáles podrían ser los problemas con este tipo de evaluación (utilice ejemplos concretos)?

Método 2:

Opinión: El proyecto “Read India” no está consiguiendo los resultados esperados

Pratham ha recaudado millones de dólares, expandiéndose rápidamente para cubrir todo India con su programa llamado “Learn-to-Read”, pero ¿aprenden realmente a leer los estudiantes? Evidencia reciente sugiere lo contrario. Un equipo de evaluadores de *Education for All* encontró que los niños que tomaron clases de lectura terminaron con niveles de alfabetización mucho menores que los de sus pares en el pueblo. Después de un año de asistir a los campamentos de lectura de Pratham, los estudiantes sólo podían reconocer palabras, mientras que quienes no participaron de los programas de Pratham pudieron leer párrafos completos.

Comparación de los niveles de lectura de los niños que tomaron clases de lectura y los niños que no las tomaron



Notas: El nivel de lectura es un indicador variable que adquiere valor 0 si el niño no puede leer nada, 1 si conoce el alfabeto, 2 si puede reconocer palabras, 3 si puede leer un párrafo y 4 si puede leer un cuento completo.

Si tiene cinco centavos que le sobren y desea realizar un aporte a la educación de los niños analfabetos en India, piénselo dos veces antes de arrojarlos al pozo de las promesas de Pratham.

Tema de discusión 2:

1. ¿Qué tipo de evaluación utiliza este documento de opinión?
2. ¿Cómo intenta a estimar el contrafactual este método? ¿Qué usa para el grupo de comparación?
3. ¿Cuál podría ser el problema con este tipo de evaluación (utilice ejemplos concretos)?

Método 3:

Carta al editor: *Education for All* debe considerar evaluar de manera justa y precisa

Han habido varios informes injustos en la prensa relacionados con programas implementados por la ONG Pratham. Un artículo reciente de un ex burócrata de *Education for All* señala que Pratham en realidad daña a los niños que recluta para sus campamentos de “Learn-to-Read”. Sin embargo, el análisis de la EFA utiliza una métrica errónea para medir el impacto. Compara los *niveles* de lectura de los estudiantes de Pratham con los de los otros niños en el pueblo, sin tomar en cuenta el hecho de que Pratham apunta a aquellos cuyos niveles de alfabetización son particularmente deficientes desde el comienzo. Si Pratham simplemente reclutara a los niños más alfabetizados en sus programas y los comparara con sus contrapartes que presentan más deficiencias, podrían declarar su éxito sin realizar una sola clase. Pero Pratham no hace esto. Y de manera realista, Pratham no espera que sus niños inicialmente analfabetos superen a los estudiantes más avanzados en el pueblo. Simplemente intenta iniciar mejoras sobre el estado actual. Por lo tanto, la métrica debería ser la *mejora* en los niveles de lectura, no el nivel final. Cuando repetimos el análisis de la EFA utilizando este indicador de resultados más adecuado, los niños de Pratham mejoraron a razón del doble que los niños que no pertenecían a Pratham (aumento en el nivel de lectura de 0,6 en comparación con 0,3). Esta diferencia es estadísticamente significativa.

Si los evaluadores de EFA hubieran usado estos resultados más adecuados, reconocerían el increíble éxito de Read India. Probablemente ellos mismos deberían inscribirse en algunas clases de Pratham.

Tema de discusión 3:

1. ¿Qué tipo de evaluación utiliza esta carta?
2. ¿Cómo intenta estimar el contrafactual este método? ¿Qué usa para el grupo de comparación?
3. ¿Cuál podría ser el problema con este tipo de evaluación (utilice ejemplos concretos)?

Método 4:

Los datos nunca mienten siempre cuando el investigador no está durmiendo

Pratham celebra la victoria, sus oponentes denuncian fraude. Una mirada más cercana muestra que, como es habitual, la verdad se encuentra en algún lugar intermedio.

Ha habido una guerra en la prensa entre los partidarios y los detractores de Pratham. Pratham y sus defensores sostienen que la campaña “Read India” ha tenido como resultado grandes aumentos en la alfabetización de los niños. Varios detractores reclaman que los programas de Pratham, al quitar la atención de las escuelas, están de hecho provocando un daño importante a los estudiantes. Lamentablemente, esta batalla se está disputando con el uso de instrumentos de análisis que tienen graves fallas. La víctima final es el público que busca una respuesta a la pregunta: ¿ayuda Pratham a sus presuntos beneficiarios?

Este informe utiliza métodos estadísticos sofisticados para medir el verdadero impacto de los programas de Pratham. Nuestra mayor preocupación eran otras variables que pudieran estar

confundiendo resultados previos. Por lo tanto, realizamos una encuesta en estos pueblos para recolectar información acerca de la edad del niño, el grado que cursa y el nivel educacional de los padres, y la utilizamos para predecir los puntajes en la prueba del niño.

Tabla 1: Resultados de Lectura							
	Nivel		Mejora				
	(1)	(2)	(3)	(4)			
Clases de lectura	-0.68 (0.0829)	** 0.04 (0.1031)	0.24 (0.0628)	** 0.11 (0.1081)			
Nivel de lectura previo		0.71 (0.0215)	**				
Edad		0.00 (0.0182)		-0.01 (0.0194)			
Sexo		-0.01 (0.0469)		0.05 (0.0514)			
Estándar		0.02 (0.0174)		-0.08 (0.0171)	**		
Alfabetización de los padres		0.04 (0.0457)		0.13 (0.0506)	**		
Constante	2.82 (0.0239)	0.36 (0.2648)	0.37 (0.0157)	0.75 (0.3293)			
Controles de tipo escolar	No	Sí	No	0.37			

Notas: La categoría omitida para el tipo escolar es el de "No asiste a la escuela". El nivel de lectura es una variable indicador que adquiere valor 0 si el niño no puede leer nada, 1 si conoce el alfabeto, 2 si puede reconocer palabras, 3 si puede leer un párrafo y 4 si puede leer una historia completa

Al observar la Tabla 1, encontramos algunos resultados positivos, algunos resultados negativos y algunos "sin resultados", dependiendo de las variables por las cuales controlemos. Los resultados de la columna (1) sugieren que el programa de Pratham daña a los niños. Existe una correlación negativa entre recibir clases de Pratham y los resultados finales de lectura (-0,68). La columna (3), que evalúa las mejoras, sugiere resultados impresionantes (0,24). Sin embargo, al observar los resultados del niño (ya sea en términos de nivel o mejora) *que controlan* por los niveles iniciales de lectura, edad, género, nivel de educación normal y de los padres, todos determinantes de los niveles de lectura de los niños, no encontramos impacto de los programas de Pratham.

Por lo tanto, al controlar por las variables correctas, hemos descubierto que por un lado, Pratham no ha provocado el daño que delataban ciertos oponentes, pero por otro lado, no ha ayudado a los niños a aprender. Por lo tanto, Pratham ha fallado en su esfuerzo por convencernos de que puede gastar el dinero de los donantes de forma efectiva.

Tema de discusión 4:

1. ¿Qué tipo de evaluación utiliza este informe?
2. ¿Cómo intenta a estimar el contrafactual este método? ¿Qué usa para el grupo de comparación?
3. ¿Cuál podría ser el problema con este tipo de evaluación (utilice ejemplos concretos)?

		Metodología	Descripción	¿Quién es el grupo de comparación?	Supuestos requeridos	Requerimiento de Datos
		Pre-Post	Mide variación en resultados de participantes en el tiempo. La diferencia entre la situación anterior a un tratamiento y la situación posterior a este.	Los mismos participantes del programa -justo antes de aplicarse el programa- representan el grupo de comparación.	Se asume que el programa es el único factor que influyó cualquier cambio en el resultado obtenido.	Datos de los participantes antes y después del programa.
		Diferencia Simple	Mide las diferencias entre aquellos que participaron en el programa y otros que no lo hicieron, después de completado el programa.	Los individuos que no participaron en el programa (por alguna razón), pero para los cuales tenemos datos después de completado el programa.	Los no-participantes en promedio son idénticos a los participantes excepto por la participación en el programa, y tenían la misma probabilidad de participar en el programa antes de que comenzara.	Datos de los participantes y no participantes para después del programa.
		Diferencias en Diferencias	Compara la variación en resultados a lo largo del tiempo de los participantes con la de los que no participaron del programa.	Los individuos que no participaron en el programa (por alguna razón), pero de quienes se recolectó información tanto antes como después del programa. Hay que buscar idealmente no participantes que sean "similares" a los participantes.	Asume que si el programa no existiera, los dos grupos tendrían idénticas trayectorias a lo largo de este periodo.	Datos de ambos participantes, antes y después del programa.
		Regresión Multivariada	Los individuos que recibieron el programa son comparados con los que no recibieron, y otros factores que pueden explicar las diferencias de resultados son tomados en consideración (controlados).	Los individuos que no participan en el programa (por cualquier motivo), pero de quienes se recolectó la información antes y/o después del programa. En este caso la información no esta comprometida sólo con indicadores de resultados, sino también por otras variables explicativas.	Los factores que fueron excluidos (porque eran no observables y/o que no fueron medidos) no sesgan los resultados porque no se correlacionan con los resultados o no difieren entre los que participan y los que no (al menos una de estas condiciones debe cumplirse para que no sesguen la estimación).	Los resultados así como las "variables control" para los que participan y los que no.
		"Matching"	Los individuos del grupo tratado son comparados con individuos similares en el grupo de comparación.	Matching exacto: Para cada participante, se elige al menos un no-participante que es idéntico en las características seleccionadas. Propensity score matching: los que no participan tienen una serie de características muy parecidas a los participantes, lo que predice que hubiesen podido ser parte del programa. Se busca que el promedio ponderado de las características de los individuos de cada par sean similares.	Los no-participantes en promedio son idénticos a los participantes excepto por la participación en el programa, y tenían la misma probabilidad de participar en el programa antes de que comenzara.	Resultados, así como también "variables de coincidencia o matching" para ambos grupos, los que participan y los que no.
		Diseño de Regresión Discontinua	Los individuos son clasificados en base a criterios medibles específicos. Hay una línea de corte que determina si un individuo es elegible o no. Los participantes que se encuentran en la vecindad del límite después son comparados entre sí y el criterio de elegibilidad es controlado.	Los individuos que están cerca de la línea de corte, pero que caen en el lado "equivocado", y por tanto no pueden participar del programa.	Después de controlar el criterio (y otras medidas de decisión), las diferencias restantes entre los individuos directamente debajo de la línea de corte o por encima de ella no son estadísticamente significativas y no producirán sesgo en los resultados. Un requerimiento necesario y suficiente para mantener este supuesto es que la línea de corte es un criterio estricto de adherencia.	Resultados, así como también todas las medidas para el criterio de selección (y cualquier otra medida de control).
		Variables instrumentales	La participación puede ser predicha por un factor accidental (aleatorio o casi aleatorio), que no está "instrumental", que no está correlacionada con el resultado, de manera distinta al hecho de que predice la participación (y la participación afecta el resultado).	Individuos que, por su cercanía al factor aleatorio, se predice que no participaran y (posiblemente como resultado) no participan.	Si la variable instrumental no tuviese la habilidad de predecir la participación, este "instrumento" no tendría efecto sobre o estaría no correlacionado con el resultado.	Requiere información del resultado, el "instrumento", y otras variables de control.
Método Experimental		Evaluación Aleatoria	Método experimental para medir relaciones causales entre dos variables comparando participantes y no participantes cuando la participación fue determinada por sorteo.	La participación en el programa es asignada de forma aleatoria de un grupo de potenciales participantes elegibles mayor al número de cupos para el programa	La Aleatorización "funciona", es decir, los dos grupos son estadísticamente idénticos (en factores observable y no observables)	Requiere información posterior para los grupos tratamiento y control. El tamaño de muestra debe ser lo suficientemente grande para detectar un resultado significativo.

Ejercicio 1: La mecánica de la asignación aleatoria usando MS Excel ®

Parte 1: aleatorización simple

Como la mayoría de los programas de planillas de cálculo, MS Excel cuenta con la función generación de número aleatorio.

Digamos que tenemos una lista de escuelas y queremos asignar a una mitad al tratamiento y a la otra mitad al control.

(1) Tenemos toda la lista de escuelas.

	A	B	C	D
1	Código de escuela	Nombre de escuela	Número aleatorio	asignación
2	101	Babajpura G.M.M.Kumar shala No. 1		
3	103	Babajpura Kanya Shala No. 3		
4	107	Babajpura Mishra Shala No. 7		
5	108	Babajpura Mishra Shala No. 8		
6	112	Babajpura Marathi Mishra Shala No. 12		
7	113	Babajpura Kanya Shala No. 13		
8	114	Babajpura Mishra Shala No. 14		
9	117	Babajpura Kumar Shala No. 17		
10	118	Babajpura Mishra Shala No. 18		
11	119	Babajpura Mishra Shala No. 19		
12	120	Babajpura Mishra Shala No. 20		
13	121	Babajpura Mishra Shala No. 21		
14	125	Babajpura Kumar Shala No. 25		
15	126	Babajpura Kanya Shala No. 26		
16	127	Babajpura Mishra Shala No. 27		
17	128	Babajpura Mishra Shala No. 28		
18	130	Babajpura Hindi Mishra Shala No. 30		
19	131	Babajpura Mishra Shala No. 31		
20	132	Babajpura Mishra Shala No. 32		
21	201	Fatehpura Kumar Shala No. 1		
22	202	Fatehpura Mishra Shala No. 2		
23	209	Fatehpura Mishra Shala No. 9		
24	210	Fatehpura Kanya Shala No. 10		
25	211	Fatehpura Mishra Shala No. 11		
26	213	Fatehpura Kumar Shala No. 13		
27	215	Fatehpura Hindi Mishra Shala No. 15		
28	216	Fatehpura Mishra Shala No. 16		
29	218	Fatehpura Mishra Shala No. 18		
30	219	Fatehpura Mishra Shala No. 19		
31	301	N. Sayajganj Mishra Shala No. 1 (center)		
32	303	N. Sayajganj Marathi Mishra Shala No. 3		
33	305	Sayajganj Mishra Shala No. 5		
34	306	Sayajganj Kumar Shala No. 6		
35	307	Sayajganj Mishra Shala No. 7		

(2) Asigne un número aleatorio a cada escuela:

La función **ALEATORIO()** (o **RAND()** si tienes Excel en inglés) es el generador de número aleatorio de Excel. Para usarlo, en la Columna C, escriba lo siguiente =ALEATORIO() en cada celda adyacente a cada nombre. O usted puede escribir esta función en la fila superior (fila 2) y simplemente copiar y pegar en toda la columna o hacer click y arrastrar.

	A	B	C	D
1	Código de escuela	Nombre de escuela	Número aleatorio asignación	
2	101	Babajpura G.M.M.Kumar shala No. 1	=aleatorio()	
3	103	Babajpura Kanya Shala No. 3		
4	107	Babajpura Mishra Shala No. 7		
5	108	Babajpura Mishra Shala No. 8		
6	112	Babajpura Marathi Mishra Shala No. 12		
7	113	Babajpura Kanya Shala No. 13		
8	114	Babajpura Mishra Shala No. 14		
9	117	Babajpura Kumar Shala No. 17		
10	118	Babajpura Mishra Shala No. 18		
11	119	Babajpura Mishra Shala No. 19		
12	120	Babajpura Mishra Shala No. 20		
13	121	Babajpura Mishra Shala No. 21		
14	125	Babajpura Kumar Shala No. 25		
15	126	Babajpura Kanya Shala No. 26		
16	127	Babajpura Mishra Shala No. 27		
17	128	Babajpura Mishra Shala No. 28		
18	130	Babajpura Hindi Mishra Shala No. 30		
19	131	Babajpura Mishra Shala No. 31		
20	132	Babajpura Mishra Shala No. 32		
21	201	Fatehpura Kumar Shala No. 1		
22	202	Fatehpura Mishra Shala No. 2		
23	209	Fatehpura Mishra Shala No. 9		
24	210	Fatehpura Kanya Shala No. 10		
25	211	Fatehpura Mishra Shala No. 11		
26	213	Fatehpura Kumar Shala No. 13		
27	215	Fatehpura Hindi Mishra Shala No. 15		
28	216	Fatehpura Mishra Shala No. 16		
29	218	Fatehpura Mishra Shala No. 18		
30	219	Fatehpura Mishra Shala No. 19		
31	301	N. Sayajiganj Mishra Shala No. 1 (center)		
32	303	N. Sayajiganj Marathi Mishra Shala No. 3		
33	305	Sayajiganj Mishra Shala No. 5		
34	306	Sayajiganj Kumar Shala No. 6		
35	307	Sayajiganj Mishra Shala No. 7		
36	308	Savainani Kumar Shala No. 8		

Escribir = ALEATORIO() coloca un número aleatorio de 15 dígitos entre 0 y 1 en la celda.

	A	B	C	D
1	Código de escuela	Nombre de escuela	Número aleatorio asignación	
2	101	Babajpura G.M.M.Kumar shala No. 1	0.947982734	
3	103	Babajpura Kanya Shala No. 3	0.250582855	
4	107	Babajpura Mishra Shala No. 7	0.044329091	
5	108	Babajpura Mishra Shala No. 8	0.278495399	
6	112	Babajpura Marathi Mishra Shala No. 12	0.627344366	
7	113	Babajpura Kanya Shala No. 13	0.218143899	
8	114	Babajpura Mishra Shala No. 14	0.828606736	
9	117	Babajpura Kumar Shala No. 17	0.652040312	
10	118	Babajpura Mishra Shala No. 18	0.341206087	
11	119	Babajpura Mishra Shala No. 19	0.167396304	
12	120	Babajpura Mishra Shala No. 20	0.130426588	
13	121	Babajpura Mishra Shala No. 21	0.747381659	
14	125	Babajpura Kumar Shala No. 25	0.474458661	
15	126	Babajpura Kanya Shala No. 26	0.110874057	
16	127	Babajpura Mishra Shala No. 27	0.946000482	
17	128	Babajpura Mishra Shala No. 28	0.029266351	
18	130	Babajpura Hindi Mishra Shala No. 30	0.05150847	
19	131	Babajpura Mishra Shala No. 31	0.479153942	
20	132	Babajpura Mishra Shala No. 32	0.492306686	
21	201	Fatehpura Kumar Shala No. 1	0.295018795	
22	202	Fatehpura Mishra Shala No. 2	0.055901542	
23	209	Fatehpura Mishra Shala No. 9	0.130469506	
24	210	Fatehpura Kanya Shala No. 10	0.92361372	
25	211	Fatehpura Mishra Shala No. 11	0.294344947	
26	213	Fatehpura Kumar Shala No. 13	0.675364861	
27	215	Fatehpura Hindi Mishra Shala No. 15	0.498212689	
28	216	Fatehpura Mishra Shala No. 16	0.920947921	
29	218	Fatehpura Mishra Shala No. 18	0.645629999	
30	219	Fatehpura Mishra Shala No. 19	0.746717479	
31	301	N. Sayajiganj Mishra Shala No. 1 (center)	0.126801429	
32	303	N. Sayajiganj Marathi Mishra Shala No. 3	0.427312788	
33	305	Sayajiganj Mishra Shala No. 5	0.845369665	
34	306	Sayajiganj Kumar Shala No. 6	0.416067614	
35	307	Sayajiganj Mishra Shala No. 7	0.825719923	
36	308	Savainani Kumar Shala No. 8	0.879144666	

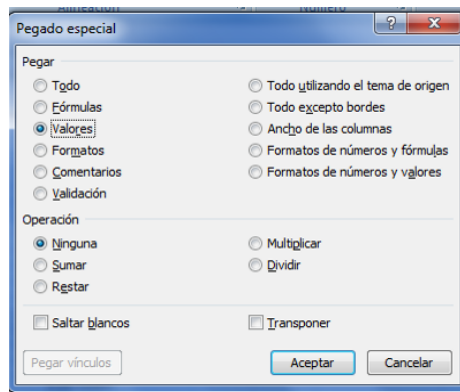
(3) Copie las celdas en la Columna C, luego pegue los valores en las mismas celdas

La función =ALEATORIO() volverá a aleatorizar cada vez que usted realice cambios en cualquier otra parte de la planilla de cálculo. Excel hace esto debido a que recalcula todos los valores que tengan cualquier cambio en cualquier celda. (Usted también puede inducir el re-cálculo y por lo tanto la re-aleatorización presionando la tecla F9.)

Sin embargo, esto puede confundir. Una vez que hemos generado nuestra columna de números aleatorios no necesitamos volver a aleatorizar. Ya tenemos una columna limpia de valores aleatorios. Para hacer que Excel deje de recalcular, usted puede reemplazar las “funciones” en esta columna con “valores”.

Para hacer esto, seleccione todos los valores en la Columna C. Luego, haga click con el botón derecho en cualquier parte de la columna destacada y escoja la opción “copiar”.

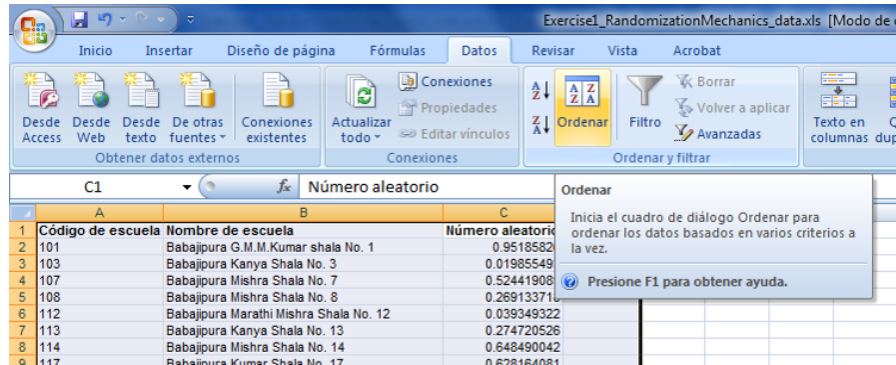
Luego, haga click con el botón derecho en cualquier parte de esa columna y escoja la opción “Pegado Especial”. Aparecerá la ventana de “Pegado Especial”. Haga click en “Valores”.



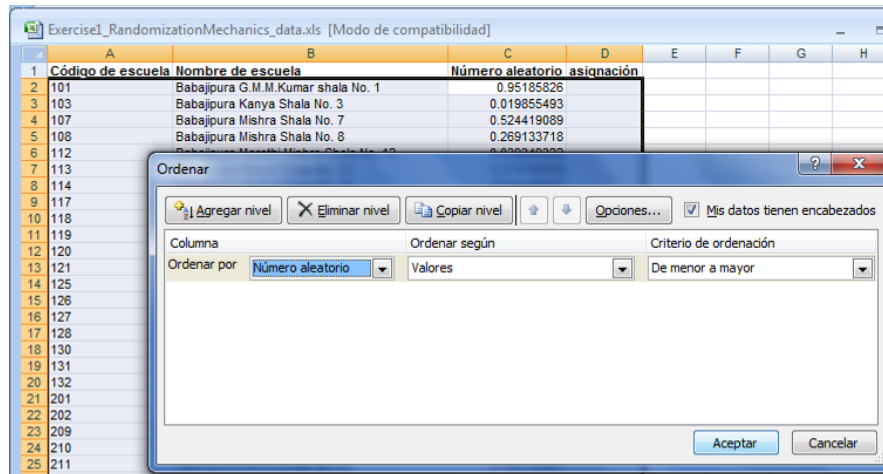
Incorporando la asignación aleatoria al diseño de investigación

(4) Ordene las columnas de acuerdo a la Columna C ya sea en orden ascendiente o descendiente:

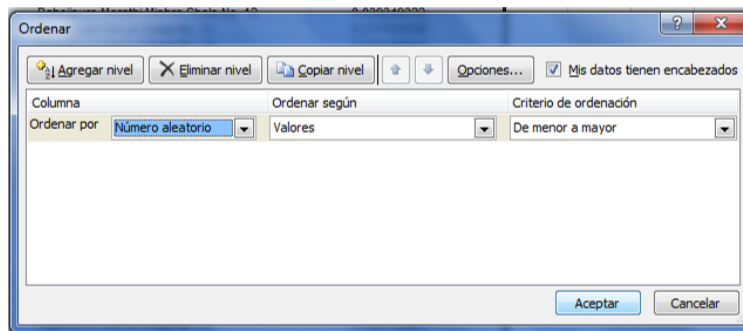
Seleccione las columnas A, B, y C. En la pestaña datos presione el botón ordenar:



Aparecerá una ventana “ordenar”.



En Ordenar por columna, seleccione “# aleatorio”. Haga click en Aceptar. Hacienda esto ordena la lista por el número aleatorio en orden ascendiente o descendiente. El que usted escoja.



¡Listo! Usted tiene una lista ordenada aleatoriamente.

	A	B	C	D
1	Código de escuela	Nombre de escuela	Número aleatorio	asignación
2	103	Babajpura Kanya Shala No. 3	0.019855493	
3	526	Shaher Vibhag Mishra Shala No. 26	0.030204755	
4	209	Fatehpura Mishra Shala No. 9	0.035708391	
5	112	Babajpura Marathi Mishra Shala No. 12	0.039349322	
6	326	Sayajiganj Hindi Mishra Shala No. 26	0.045140569	
7	410	Raopura Kumar Shala No. 10	0.062378333	
8	118	Babajpura Mishra Shala No. 18	0.076918506	
9	506	Shaher Vibhag Mishra Shala No. 6	0.076921029	
10	216	Fatehpura Mishra Shala No. 16	0.082951096	
11	523	Shaher Vibhag Marathi Mishra Shala No. 23	0.083529625	
12	411	Raopura Kumar Shala No. 11	0.0850926	
13	337	Sayajiganj Kanya Shala No. 37	0.085750965	
14	529	Shaher Vibhag Mishra Shala No. 29	0.110202952	
15	522	Shaher Vibhag Mishra Shala No. 22	0.116180033	
16	331	Sayajiganj Mishra Shala No. 31	0.121187102	
17	502	Shaher Vibhag Mishra Shala No. 2	0.121863691	
18	120	Babajpura Mishra Shala No. 20	0.123780268	
19	202	Fatehpura Mishra Shala No. 2	0.124786504	
20	127	Babajpura Mishra Shala No. 27	0.126571213	
21	308	Sayajiganj Kumar Shala No. 8	0.140000294	
22	602	Wadi Mishra Shala No. 2	0.143501187	
23	634	Wadi Mishra Shala No. 34	0.153474111	
24	215	Fatehpura Hindi Mishra Shala No. 15	0.160532217	

(5) Ordene las columnas ya sea en orden ascendente o descendente de acuerdo a la Columna C:

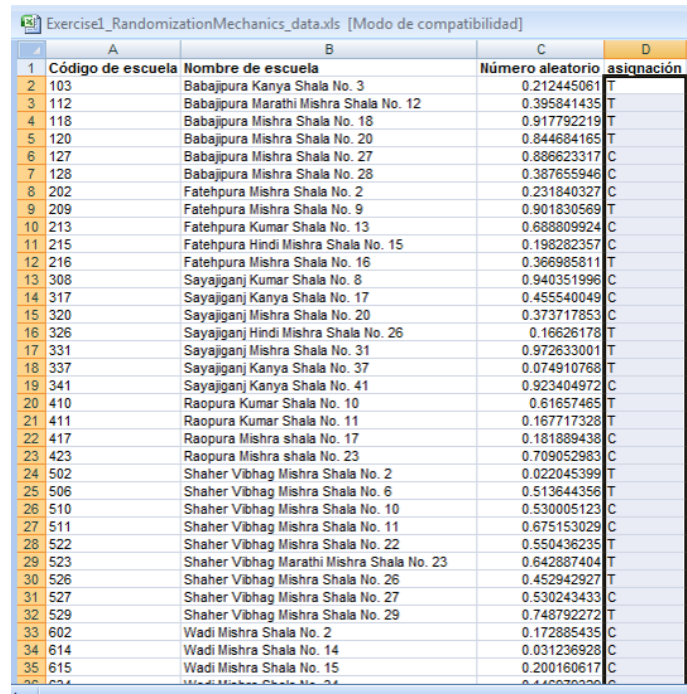
Debido a que su lista está ordenada aleatoriamente, es completamente aleatorio si las escuelas se encuentran en la mitad superior o inferior de la lista. Por lo tanto, si usted asigna la mitad superior al grupo de tratamiento y la mitad inferior al grupo de control, sus escuelas habrán sido “asignadas aleatoriamente”.

En la columna D escriba “T” para la primera mitad de las filas (filas 2-61). Para la segunda mitad de las filas (filas 62-123), escriba “C”.

	A	B	C	D
1	Código de escuela	Nombre de escuela	Número aleatorio	asignación
2	103	Babajpura Kanya Shala No. 3	0.212445061	T
3	526	Shaher Vibhag Mishra Shala No. 26	0.452942927	T
4	209	Fatehpura Mishra Shala No. 9	0.901830569	T
5	112	Babajpura Marathi Mishra Shala No. 12	0.395841435	T
6	326	Sayajiganj Hindi Mishra Shala No. 26	0.16626178	T
7	410	Raopura Kumar Shala No. 10	0.61657465	T
8	118	Babajpura Mishra Shala No. 18	0.917792219	T
9	506	Shaher Vibhag Mishra Shala No. 6	0.513644356	T
10	216	Fatehpura Mishra Shala No. 16	0.366985811	T
11	523	Shaher Vibhag Marathi Mishra Shala No. 23	0.642887404	T
12	411	Raopura Kumar Shala No. 11	0.167717328	T
13	337	Sayajiganj Kanya Shala No. 37	0.074910768	T
14	529	Shaher Vibhag Mishra Shala No. 29	0.748792272	T
15	522	Shaher Vibhag Mishra Shala No. 22	0.550436235	T
16	331	Sayajiganj Mishra Shala No. 31	0.972633001	T
17	502	Shaher Vibhag Mishra Shala No. 2	0.022045399	T
18	120	Babajpura Mishra Shala No. 20	0.844684165	T
19	202	Fatehpura Mishra Shala No. 2	0.231840327	C
20	127	Babajpura Mishra Shala No. 27	0.88623317	C
21	308	Sayajiganj Kumar Shala No. 8	0.940351996	C
22	602	Wadi Mishra Shala No. 2	0.172885435	C
23	634	Wadi Mishra Shala No. 34	0.146979329	C
24	215	Fatehpura Hindi Mishra Shala No. 15	0.198282357	C
25	511	Shaher Vibhag Mishra Shala No. 11	0.675153029	C
26	614	Wadi Mishra Shala No. 14	0.031236928	C
27	341	Sayajiganj Kanya Shala No. 41	0.923404972	C
28	320	Sayajiganj Mishra Shala No. 20	0.373717853	C
29	527	Shaher Vibhag Mishra Shala No. 27	0.530243433	C
30	615	Wadi Mishra Shala No. 15	0.200160617	C
31	417	Raopura Mishra shala No. 17	0.181889438	C
32	423	Raopura Mishra shala No. 23	0.709052983	C
33	510	Shaher Vibhag Mishra Shala No. 10	0.530005123	C
34	128	Babajpura Mishra Shala No. 28	0.387655946	C
35	317	Sayajiganj Kanya Shala No. 17	0.455540049	C

Incorporando la asignación aleatoria al diseño de investigación

Vuelva a ordenar su lista por identificación de escuela. Verá que sus escuelas han sido asignadas aleatoriamente a grupos de tratamiento y de control



	A	B	C	D
	Código de escuela	Nombre de escuela	Número aleatorio	asignación
1	103	Babajpura Kanya Shala No. 3	0.212445061	T
2	112	Babajpura Marathi Mishra Shala No. 12	0.395841435	T
3	118	Babajpura Mishra Shala No. 18	0.917792219	T
4	120	Babajpura Mishra Shala No. 20	0.844684165	T
5	127	Babajpura Mishra Shala No. 27	0.886623317	C
6	128	Babajpura Mishra Shala No. 28	0.387655946	C
7	202	Fatehpura Mishra Shala No. 2	0.231840327	C
8	209	Fatehpura Mishra Shala No. 9	0.901830569	T
9	213	Fatehpura Kumar Shala No. 13	0.688809924	C
10	215	Fatehpura Hindi Mishra Shala No. 15	0.198282357	C
11	216	Fatehpura Mishra Shala No. 16	0.368985811	T
12	308	Sayajganj Kumar Shala No. 8	0.940351996	C
13	317	Sayajganj Kanya Shala No. 17	0.455540049	C
14	320	Sayajganj Mishra Shala No. 20	0.373717853	C
15	326	Sayajganj Hindi Mishra Shala No. 26	0.16626178	T
16	331	Sayajganj Mishra Shala No. 31	0.972633001	T
17	337	Sayajganj Kanya Shala No. 37	0.074910768	T
18	341	Sayajganj Kanya Shala No. 41	0.923404972	C
19	410	Raopura Kumar Shala No. 10	0.61657465	T
20	411	Raopura Kumar Shala No. 11	0.167717328	T
21	417	Raopura Mishra shala No. 17	0.181889438	C
22	423	Raopura Mishra shala No. 23	0.709052983	C
23	502	Shaher Vibhag Mishra Shala No. 2	0.022045399	T
24	506	Shaher Vibhag Mishra Shala No. 6	0.513644356	T
25	510	Shaher Vibhag Mishra Shala No. 10	0.530005123	C
26	511	Shaher Vibhag Mishra Shala No. 11	0.675153029	C
27	522	Shaher Vibhag Mishra Shala No. 22	0.550436235	T
28	523	Shaher Vibhag Marathi Mishra Shala No. 23	0.642887404	T
29	526	Shaher Vibhag Mishra Shala No. 26	0.452942927	T
30	527	Shaher Vibhag Mishra Shala No. 27	0.530243433	C
31	529	Shaher Vibhag Mishra Shala No. 29	0.748792272	T
32	602	Wadi Mishra Shala No. 2	0.172885435	C
33	614	Wadi Mishra Shala No. 14	0.031236928	C
34	615	Wadi Mishra Shala No. 15	0.200160617	C
35	624	Wadi Mishra Shala No. 24	0.440696932	C

Parte 2: Aleatorización estratificada

La estratificación es el proceso de dividir una muestra en grupos y luego asignar los individuos aleatoriamente dentro de cada grupo al tratamiento o control. Las razones para hacer esto son más bien técnicas. Una razón para estratificar es que aumenta la probabilidad de que los subgrupos sean equilibrados, haciendo que sea más fácil realizar ciertos análisis de subgrupos sea más fácil. Por ejemplo, si usted desea probar la efectividad en un nuevo programa separadamente para las escuelas donde los niños se les enseña Hindú versus las escuelas donde se les enseña Gujarati, usted puede estratificar por “idioma de instrucción” y asegurarse de que hay igual número de escuelas en cada idioma en los grupos de tratamiento y control.

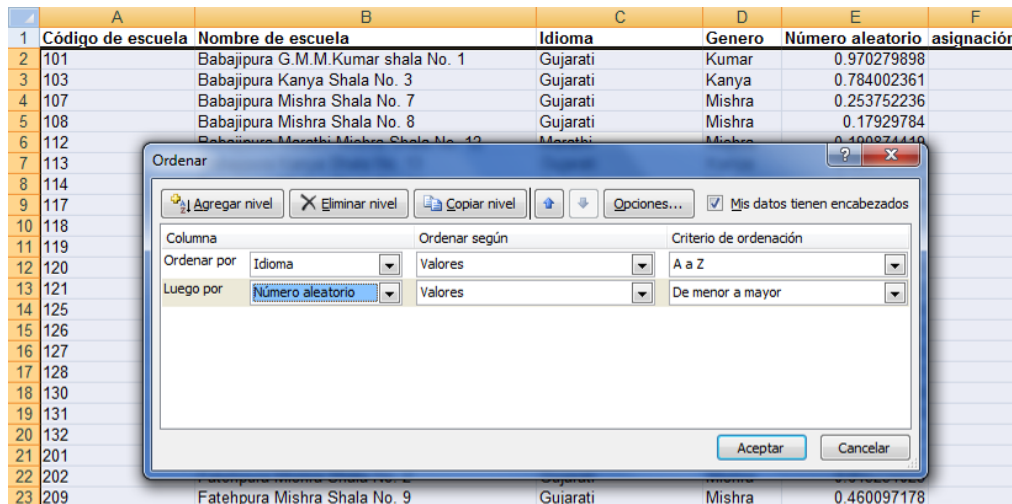
(1) Tenemos toda nuestra lista de escuelas y potenciales “estratos”.

Mecánicamente, la única diferencia en ordenar aleatoriamente es que en vez de ordenar simplemente por número aleatorio, usted primero ordenaría por idioma y luego por número aleatorio. Obviamente, el primer paso es para asegurar que usted tiene las variables por las cuales desea estratificar. Para hacer este ejercicio, haga click en la segunda hoja del archivo, “Estratificado”

(2) Ordene por estratos y luego por número aleatorio

Asumiendo que cuenta con todas las variables que necesita: en la pestaña de datos, seleccione “Ordenar”. Aparecerá la ventana de Ordenar. Ordene por “Idioma”. Presione

el botón “Agregar Nivel” y seleccione la columna: “Número aleatorio”.



Asigne Tratamiento - Control para cada grupo.

Dentro de cada grupo de idiomas, escriba “T” para la primera mitad de filas y “C” para la segunda mitad.

	A	B	C	D	E	F
100	132	Babajipura Mishra Shala No. 32	Gujarati	Mishra	0.8931975	C
101	615	Wadi Mishra Shala No. 15	Gujarati	Mishra	0.9142383	C
102	618	Wadi Kumar Shala No. 18	Gujarati	Kumar	0.9229356	C
103	408	Raopura Kanya Shala No. 8	Gujarati	Kanya	0.9285077	C
104	502	Shaher Vibhag Mishra Shala No. 2	Gujarati	Mishra	0.9549163	C
105	311	Sayajiganj Mishra Shala No. 11	Gujarati	Mishra	0.9595266	C
106	344	Sayajiganj Mishra Shala No. 44	Gujarati	Mishra	0.9688854	C
107	347	Sayajiganj Hindi Mishra Shala No. 47	Hindi	Mishra	0.0163449	T
108	332	Sayajiganj Hindi Mishra Shala No. 32	Hindi	Mishra	0.1528766	T
109	342	Sayajiganj Hindi Mishra Shala No. 42	Hindi	Mishra	0.2646791	T
110	215	Fatehpura Hindi Mishra Shala No. 15	Hindi	Mishra	0.3142377	T
111	326	Sayajiganj Hindi Mishra Shala No. 26	Hindi	Mishra	0.4291559	T
112	638	Wadi Hindi Mishra Shala No. 38	Hindi	Mishra	0.6772441	C
113	130	Babajipura Hindi Mishra Shala No. 30	Hindi	Mishra	0.7053783	C
114	315	Sayajiganj Hindi Mishra Shala No. 15	Hindi	Mishra	0.7955641	C
115	626	Wadi Hindi Mishra Shala No. 26	Hindi	Mishra	0.8918818	C
116	346	Sayajiganj Hindi Mishra Shala No. 46	Hindi	Mishra	0.9051467	C
117	303	N. Sayajiganj Marathi Mishra Shala No. 3	Marathi	Mishra	0.0354843	T
118	523	Shaher Vibhag Marathi Mishra Shala No. 23	Marathi	Mishra	0.1834626	T
119	409	Raopura Marathi Mishra Shala No. 9	Marathi	Mishra	0.7676874	T
120	611	Wadi Marathi Mishra Shala No. 11	Marathi	Mishra	0.8847497	T
121	329	Sayajiganj Marathi Mishra Shala No. 29	Marathi	Mishra	0.8992905	C
122	112	Babajipura Marathi Mishra Shala No. 12	Marathi	Mishra	0.9430188	C
123	327	Sayajiganj Marathi Mishra Shala No. 27	Marathi	Mishra	0.9515261	C
124	617	Wadi Marathi Mishra Shala No. 17	Marathi	Mishra	0.9648498	C

ABDUL LATIF JAMEEL

Poverty Action Lab



TRANSLATING RESEARCH INTO ACTION



Caso 3: Programa de Profesores adicionales Diseñando una evaluación para responder tres preguntas claves sobre políticas educativas

Este estudio se basa en el documento: "Peer Effects and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya," de Esther Duflo (MIT), Pascaline Dupas (UCLA), y Michael Kremer (Harvard)

J-PAL agradece a los autores la autorización para utilizar su publicación.

Al confrontarse con escuelas con demasiados estudiantes y una escasez de profesores, en 2005 la ONG *International Child Support Africa* (ICS) se ofreció a apoyar el sistema escolar de Kenia Occidental mediante la contratación de profesores en 140 escuelas primarias. En su programa de dos años, ICS proporcionó fondos a estas escuelas para que cada una contratara un profesor adicional. En contraste con los funcionarios civiles contratados por el Ministerio de Educación, los profesores por contrato son contratados localmente por comités escolares. ICS esperaba que este programa mejorara el aprendizaje de los estudiantes mediante, entre otras cosas, la disminución del tamaño de las clases y la utilización de profesores que rindieran cuentas directamente a las comunidades a las que prestan servicios. Sin embargo, los profesores por contrato tienden a tener menos capacitación y recibir un salario mensual menor que los que son funcionarios civiles. Por esta razón, existía preocupación acerca de si estos profesores estaban lo suficientemente motivados, dada su remuneración, o calificados, dadas sus credenciales.

¿Qué diseños experimentales podrían probar el impacto de esta intervención en el logro educativo? ¿Cuál de estos cambios en el entorno escolar es el principal responsable del mejor desempeño de los estudiantes?

Escuelas con exceso de alumnos

Al igual que muchos otros países en vías de desarrollo, Kenia ha realizado rápidos progresos hacia el cumplimiento del objetivo de educación primaria universal de los Objetivos de Desarrollo del Milenio. En gran parte gracias a la eliminación de las tarifas escolares en 2003, la matrícula en las escuelas primarias aumentó casi un 30 por ciento, de 5,9 millones a 7,6 millones entre 2002 y 2005.

Sin embargo, sin una inversión correspondiente por parte del gobierno, este progreso ha dado origen a un conjunto de nuevos desafíos en Kenia:

- 1) **Tamaño de clases demasiado grandes:** Debido a las restricciones presupuestarias, el aumento en la matrícula escolar no ha coincidido con un aumento proporcional en el número de profesores. (Los salarios de los Profesores ya son el componente más grande del gasto en educación.) El resultado ha sido clases muy grandes, particularmente en los grados menores. Por ejemplo, en una muestra de escuelas en Kenia Occidental, la clase de primer grado en 2005 tenía un promedio de 83 estudiantes. Esto es preocupante porque típicamente es considerado que las clases pequeñas son más importantes para los estudiantes más jóvenes, que aún se están acostumbrando al entorno escolar. La Kenyan National Union of Teachers (Sindicato Nacional de Profesores de Kenia) calcula que el país necesita 60.000 profesores de escuela primaria adicionales, además de los 175.000 ya existentes, para llegar a todos los estudiantes de escuela primaria y reducir los tamaños de las clases.

Programa de Profesores Adicionales en Kenia

- 2) **Ausentismo de los profesores:** Exacerbando aun más el problema de la relación profesor-alumno, el absentismo de los profesores sigue siendo alto, llegando a cerca del 20% en algunas áreas de Kenia.

Normalmente no existen reemplazos para los profesores ausentes, de modo que los estudiantes simplemente dan vueltas, se van a casa o se integran en otra clase, por lo general de otro grado. Las escuelas pequeñas, que prevalecen en áreas rurales de los países en vías de desarrollo, pueden cerrarse completamente por la ausencia de un profesor. Las familias deben considerar si la escuela estará abierta en el momento de decidir si envían o no a sus hijos a la escuela. Un resultado obvio es la baja asistencia de los estudiantes, incluso en los días en que la escuela está abierta.

- 3) **Clases heterogéneas:** Las clases en Kenia también son muy heterogéneas, con estudiantes que varían ampliamente en términos de preparación escolar y el apoyo que reciben en el hogar.

Agrupar a los estudiantes en clases clasificadas por aptitud (*subdivisión o segregación*) es una idea controversial entre los académicos y los elaboradores de políticas públicas. Por un lado, si los profesores encuentran más fácil enseñar a un grupo homogéneo de estudiantes, la subdivisión podría mejorar la efectividad de las escuelas y los puntajes en los exámenes. Por otra parte, si los estudiantes aprenden en parte de sus compañeros, la subdivisión podría convertirse en una desventaja para los estudiantes con malos resultados y un beneficio para los que tiene buenos resultados, lo que exacerbaría la desigualdad.

- 4) **Escasez de materiales para la escuela:** Debido a los altos costos de materiales educativos y el creciente número de estudiantes, los recursos educacionales se han quedado cortos, y en algunos casos hasta cuatro estudiantes deben compartir un solo libro. Además, la ya sobrecargada infraestructura se deteriora más rápido cuando se ve obligada a servir a más niños.
- 5) **Bajos índices de finalización:** Como resultado de estos factores, los índices de finalización en Kenia son muy bajos, con sólo 45,1% de los niños y 43,3% de las niñas finalizando la primaria.

Considerando todos estos problemas, se presentan nuevos desafíos para las comunidades: ¿Cómo garantizar la calidad mínima de la educación, dadas las restricciones de presupuesto de Kenia?

¿Qué son los Profesores por Contrato?

Los gobiernos de varios países en vías de desarrollo han respondido a desafíos similares llenando las vacantes de profesores con profesores por contrato locales que no son funcionarios civiles. Las cuatro características principales de los profesores por contrato son: (1) se contratan con contratos de renovación anual, sin garantía de renovación del empleo (a diferencia de lo que ocurre con los profesores funcionarios); (2) con frecuencia

son menos calificados que los profesores regulares y es menos probable que tengan un certificado o grado de capacitación de profesor; (3) reciben salarios menores que los de profesores regulares (normalmente reciben menos de un quinto de los salarios que se pagan a los profesores) y (4) es más probable que sean del área local donde se encuentra la escuela.

¿Son efectivos los profesores por contrato?

El incremento de los profesores por contrato ha sido una de las innovaciones más importantes para las políticas públicas que buscan proporcionar educación primaria en los países en vías de desarrollo. Sin embargo, también ha sido una idea altamente controversial. Los partidarios señalan que utilizar profesores por contrato es una forma eficiente de ampliar el acceso y la calidad de la educación para una gran cantidad de alumnos que conforman una primera generación de estudiantes. Al saber que la recontractación al año siguiente depende de una decisión del comité de la escuela y que depende de su desempeño, los profesores por contrato tienen motivación para esforzarse más que sus colegas con cargos gubernamentales. Los profesores por contrato además son más similares a sus estudiantes en términos de raíces geográficas y culturales, y en situación socioeconómica. Los oponentes argumentan que utilizando profesores no capacitados y sin las certificaciones adecuadas se puede suplir personal en las salas de clases, pero no producirá resultados positivos de aprendizaje. Además, argumentan que el uso de profesores por contrato resta profesionalismo a la enseñanza, reduce el prestigio de toda la profesión y reduce la motivación de todos los profesores. Incluso si ayuda en el corto plazo, puede perjudicar los esfuerzos para reclutar a profesores calificados en el futuro.

Aunque el uso de los profesores por contrato ha sido controversial, existe muy poca evidencia respecto a la efectividad de los profesores por contrato para mejorar los resultados de aprendizaje de los estudiantes.

La evaluación aleatorizada del Programa de Profesores Adicionales

En enero de 2005, *International Child Support Africa* comenzó un programa de dos años para examinar el efecto de los profesores por contrato en la educación en Kenia. Bajo el programa, ICS entregó fondos a 140 comités de escuela locales para contratar a un profesor adicional para que enseñara una clase de primer grado adicional. El propósito de esta intervención era abordar los tres desafíos más importantes: tamaño de la clase, rendición de cuentas del profesor y heterogeneidad de las aptitudes. La evaluación se diseñó para medir el impacto de las reducciones de las clases, la efectividad relativa de los profesores por contrato y el impacto de una subdivisión de las clases por aptitudes (es decir, unas clases con los estudiantes de bajos rendimientos y otras con aquéllos que tienen buenos rendimientos).

Abordar varias preguntas de investigación a través de un diseño experimental

Programa de Profesores Adicionales en Kenia

Se pueden utilizar diferentes estrategias de aleatorización para responder preguntas difíciles. ¿Qué estrategia de aleatorización se podría utilizar para evaluar las siguientes preguntas? Concéntrese en la unidad (nivel) de aleatorización adecuado para cada pregunta.

Tema de discusión 1: Probar la efectividad de los profesores por contrato

1. ¿Cuál es la efectividad relativa de los profesores por contrato en comparación con la de los profesores del gobierno?

Tema de discusión 2: Observar enfoques más generales para mejorar la educación

2. ¿Cuál es el efecto de las clases más pequeñas en el desempeño de los estudiantes?
3. ¿Cuál es el impacto de agrupar a los estudiantes según sus aptitudes en el desempeño de los estudiantes?

Tema de discusión 3: Abordar todas las preguntas con una sola evaluación

4. ¿Podría una evaluación explorar todos estos problemas al mismo tiempo?
5. ¿Qué estrategia de aleatorización podría lograrlo?

Ejercicio 2: Entendiendo el funcionamiento del muestreo aleatorio y la ley de los grandes números

En este ejercicio exploraremos visualmente las muestras de diferentes tamaños de una población dada. En particular, trataremos de demostrar que los tamaños mayores de muestra tienden a reflejar mejor la población subyacente.

- 1) Abra el archivo “Ejercicio 2 - distribución de la muestra.xlsm”.
- 2) Si se pregunta o hay una barra “advertencia de seguridad”, haga click en “Opciones” y “Habilitar Macros”.
- 3) Vaya a la hoja “Aleatorizar” que le permite escoger una muestra aleatoria de tamaño “Tamaño de Muestra” de los datos contenidos en la hoja de “control”.
- 4) Ingrese “10” para “Tamaño de la Muestra” y haga click en el botón “Aleatorizar”. Observe la distribución de las distintas características entre Tratamiento, Control y Esperado. Con un tamaño de muestra tan pequeño, la diferencia porcentual del promedio esperado es bastante alta para los puntajes de lectura. Haga click en “Aleatorizar” varias veces y observe cómo cambia la distribución.
- 5) Ahora, intente con “50” para el tamaño de la muestra. ¿Qué le pasa a las distribuciones? Repita la aleatorización algunas veces y observe la diferencia porcentual para los puntajes de lectura.
- 6) Aumente el tamaño de la muestra a “500”, “2000” y “10000” y repita las observaciones del paso 5. ¿Qué podemos decir respecto de lo que pasa con tamaños de muestra más grandes? ¿Cómo afectan nuestras muestras de Tratamiento y Control? ¿Siempre debiera bajar la diferencia porcentual entre el Tratamiento, Control y Esperado a medida que aumentamos el tamaño de la muestra?

Cálculos del tamaño muestral

El estudio de caso del Programa “Extra Teacher” (ETP) discutía el concepto de ensayos aleatorios por conglomerado. El ejemplo Balsakhi usado en la clase anterior introdujo el concepto de cálculos de poder. En este último estábamos interesados en medir el efecto de un tratamiento (balsakhis en las salas de clase) sobre los resultados medidos a nivel individual – puntajes de prueba por niño. Sin embargo, la aleatorización de Balsakhis se realizó a nivel de sala de clase. Podría ser que nuestro resultado de interés está correlacionado con los estudiantes en la misma sala de clase por razones que no tienen conexión con el Balsakhi. Por ejemplo, todos los estudiantes en la sala de clase serán afectados por su profesor, por si su sala de clase está inusualmente oscura o si tienen pizarrón para tiza; estos factores significan que cuando a un estudiante en la sala le va particularmente bien por esta razón, a todos los estudiantes en esa sala probablemente les iría mejor también – lo cual podría no tener conexión con un Balsakhi.

Por lo tanto, si hacemos una muestra de 100 niños de 10 escuelas escogidas aleatoriamente, esa muestra es menos representativa de la población de escuelas en la ciudad que si seleccionásemos 100 niños al azar de toda la población de escuelas y, por lo tanto, absorbe menos varianza. De hecho, tenemos un tamaño muestral más pequeño de lo que pensamos. Esto llevará a tener más ruido en nuestra muestra y, por lo tanto, un error estándar mayor que en la caso común de muestreo independiente. Necesitamos considerar esto al momento de planificar tanto el tamaño de la muestra como la mejor manera de hacer una muestra de las salas de clases.

Este ejercicio le ayudará a comprender cómo hacer eso. ¿Debiera uno tener una muestra de cada estudiante en solo unas pocas escuelas? ¿Debiera uno tener una muestra de pocos estudiantes en muchas escuelas? ¿Cómo decidir?

Trabajaremos estas preguntas determinando el tamaño de la muestra que nos permita detectar un efecto específico con al menos un poder de 80%. Recuerde que el poder es la probabilidad de que cuando el tratamiento tiene un efecto, usted podrá distinguirlo de cero en su muestra.

En este ejemplo los “conglomerados” se refieren a “conglomerados de niños” – en otras palabras, “salas de clase” o “escuelas”. Este ejercicio le muestra cómo cambia el poder de su muestra con el número de conglomerados, el tamaño de los conglomerados, el tamaño del efecto del tratamiento y el Coeficiente de Correlación Intra-conglomerado. Usaremos un software desarrollado por Steve Raudebush con fondos de la Fundación William T. Grant. Puede encontrar recursos adicionales sobre diseños agregados en su página Web.

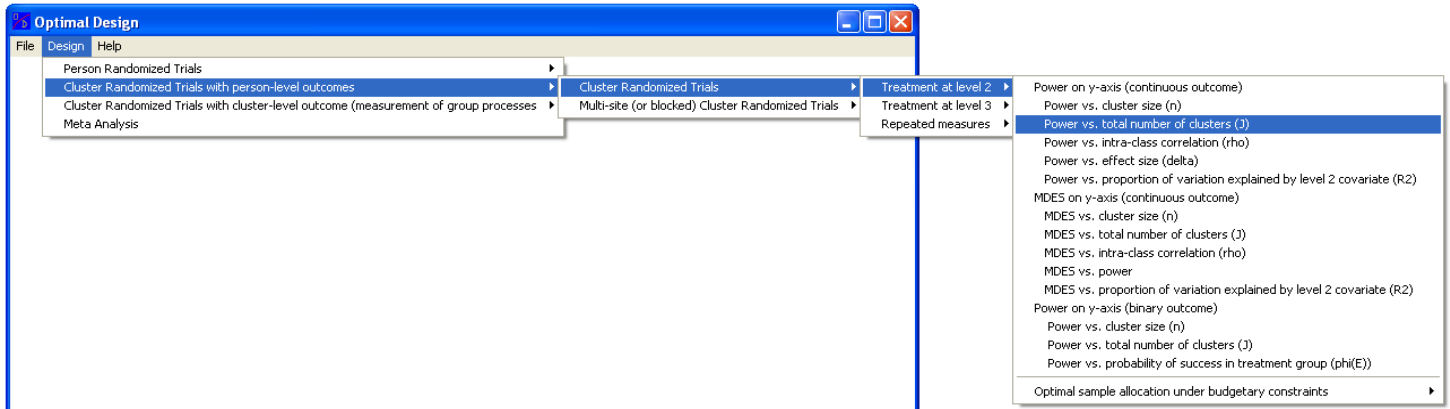
Sección 1: Usar el Software OD

Primero baje el Software OD del sitio web (se encuentra disponible también el manual del software):

http://sitemaker.umich.edu/group-based/optimal_design_software

Cuando lo abra, verá una pantalla que se parece a la de más abajo. Seleccione la opción de menú “Design” para ver el menú principal. Seleccione la opción “Cluster Randomized Trials with person-

level outcomes,” “Cluster Randomized Trials,” y luego “Treatment at level 2”. Usted verá varias opciones para generar gráficos, escoja “Power vs. Total number of clusters (J).”



Aparecerá una nueva ventana:



Seleccione α (alfa). Verá que ya está programada en 0.050 para un nivel de significancia de 95%.

Primero supongamos que deseamos probar 40 estudiantes por escuela. ¿Cuántas escuelas necesita visitar para contar con una respuesta estadísticamente significativa?

Presione **n** que representa el número de estudiantes por escuela. Ya que estamos testeando solo 40 estudiantes por escuela, solo complete n(1) con 40 y presione aceptar.

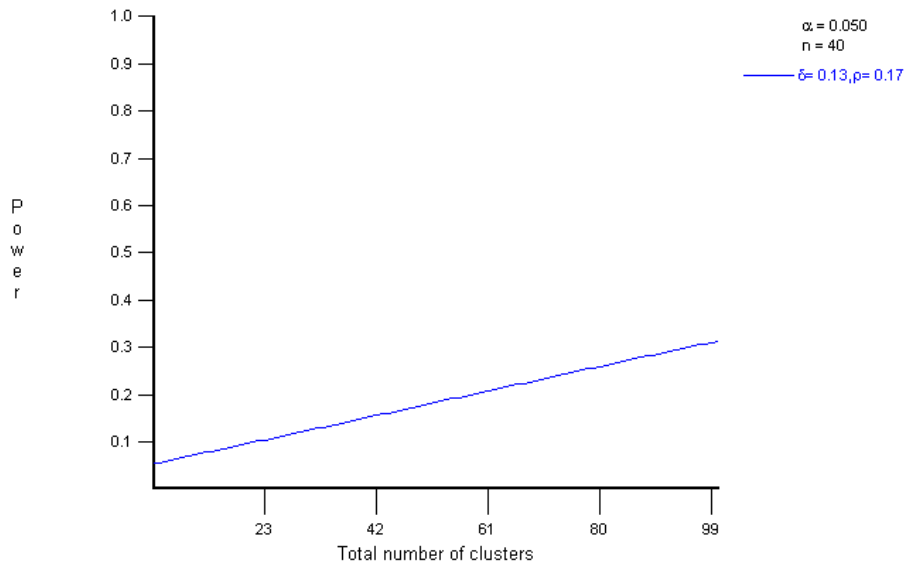
Ahora tenemos que determinar δ (delta), el tamaño estándar del efecto (el tamaño del efecto dividido por la desviación estándar de la variable de interés). Asumamos que estamos interesados en detectar si existe un aumento de 10% en los puntajes de pruebas (O con mayor precisión, no estamos interesados en detectar menos de 10%) Nuestra encuesta de línea base indicó que el puntaje promedio de prueba es 26 con una desviación estándar de 20. Deseamos detectar un tamaño de efecto de 10% de 26, el cual es 2,6. Dividimos 2,6 por la desviación estándar para obtener δ igual a 2,6/20 o 0,13.

Seleccione δ del menú. En el cuadro de diálogo que aparece hay un valor preestablecido de 0,200 para delta(1). Cambie el valor a 3,13 y cambie el valor de delta (2) a vacío. Presione aceptar.

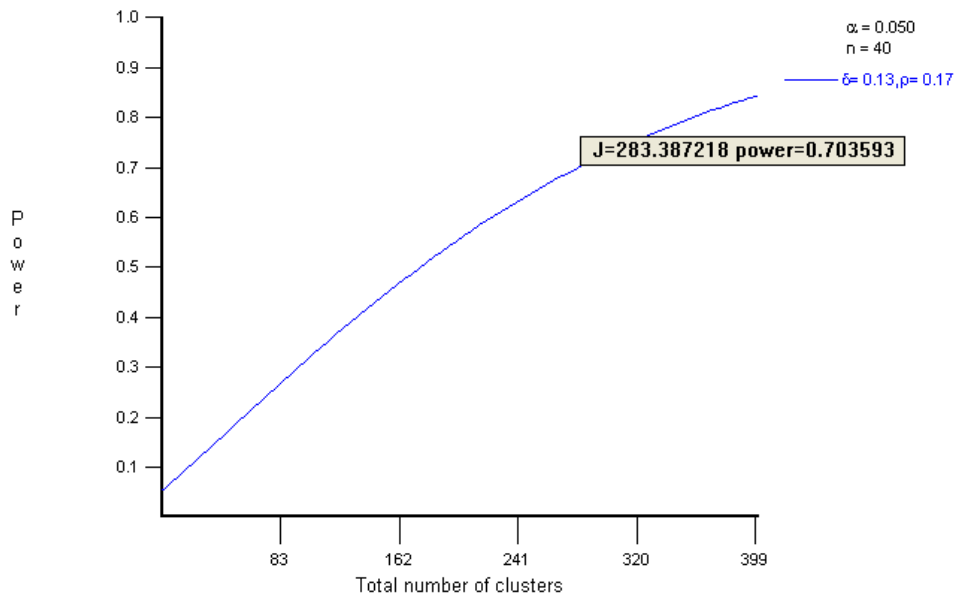
Finalmente necesitamos escoger ρ (rho), que es la correlación intra-conglomerado. ρ nos dice que tan fuertemente están correlacionados los resultados para las unidades dentro del mismo conglomerado. Si los estudiantes de la misma escuela fueran clones (sin variación) y todos obtuvieran puntajes iguales en la prueba, entonces ρ sería igual a 1. Si, por otro lado, los estudiantes de las mismas escuelas son de hecho independiente – y no existiese diferencias entre las escuelas, entonces ρ será igual a 0.

Usted ha determinado en su estudio piloto que ρ es 0,17. Complete rho(1) y deje rho (2) vacío.

Usted debiera ver un gráfico parecido al que aparece más abajo



Usted notará que su eje x no es lo suficiente largo para permitirle ver qué números de conglomerado le daría un poder de 80%. Haga click en para fijar su eje x en un máximo de 400. Luego, puede hacer click en el gráfico para ver el poder exacto y números de conglomerado para un punto en particular.



Ejercicio 3.1:

¿Cuántas escuelas se necesitan para alcanzar un poder de 80%? ¿de 90%?

Ahora usted ha visto cuántos conglomerados se necesita para un poder de 80%, haciendo una muestra de 40 estudiantes por escuela. Suponga, en cambio, que usted solo tiene la capacidad de llegar a 124 escuelas (este es el número real que se utilizó en el programa Balsakhi).

Ejercicio 3.2:

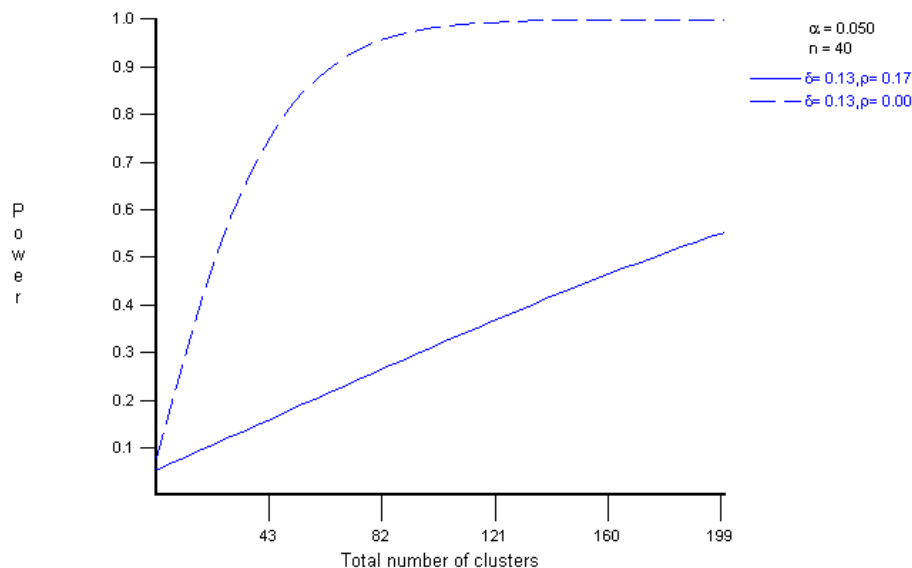
¿Cuántos niños por escuela se necesitan para alcanzar un poder de 80%? ¿de 90%? Escoja diferentes valores para n para ver cómo cambia su gráfico.


Finalmente, veamos cómo cambia el poder de una muestra dado del Coeficiente de Correlación Intra-conglomerado (ρ). Deje rho(1) en 0,17, pero para la comparación cambie rho(2) a 0,0.

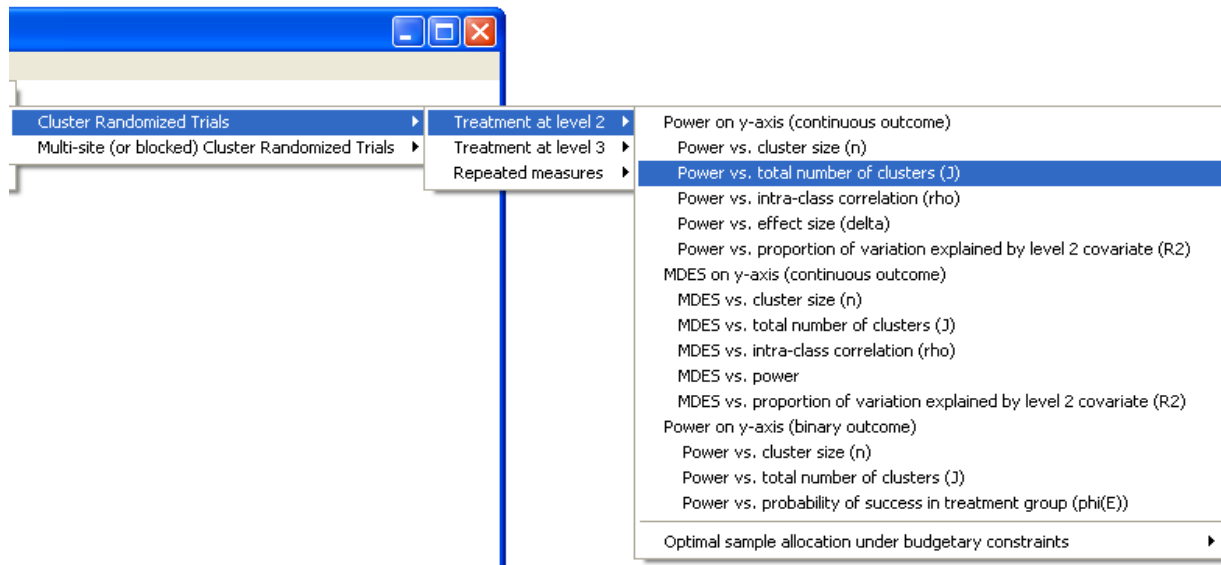
Usted debiera ver un gráfico como el siguiente. La curva sólida azul es la que tiene los parámetros que usted ha establecido basado en su estimaciones previas al testeó del efecto de las reservas de mujeres respecto del agua potable. La curva segmentada azul es para propósitos de comparación – para ver cuánto poder obtendría de su muestra si ρ fuese cero. Mire cuidadosamente el gráfico.

Ejercicio 3.3:

¿Cómo cambia el poder de la muestra con el Coeficiente de Correlación Intra-conglomerado (ρ)?



Mire algunas de las otras opciones del menú, cierre el gráfico haciendo click en  en la esquina superior derecha de la ventana interna. Seleccione nuevamente el menú Cluster Randomized Trial.

**Ejercicio 3.4:**

Intente generar gráficos para ver cómo cambia el poder con el tamaño de conglomerado (n), correlación Intra-conglomerado (ρ) y tamaño del efecto (δ).

Tendrá que volver a ingresar sus parámetros pre-testeo cada vez que usted abra un nuevo gráfico.

ABDUL LATIF JAMEEL

Poverty Action Lab



TRANSLATING RESEARCH INTO ACTION



Caso 4: Desparasitación en Kenia Enfrentando las amenazas a la integridad del experimento

Este estudio de caso está basado en el trabajo de Edward Miguel y Michael Kramer, "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica* 72(1): 159-217, 2004

J-PAL agradece a los autores el permiso para utilizar esta publicación

Caso 4: Enfrentando las amenazas a la integridad del experimento

Entre 1998 y 2001, la ONG *International Child Support Africa* implementó un programa de desparasitación masiva focalizado en los colegios, en 75 escuelas primarias en Kenia occidental. En el programa se trató a 30.000 pupilos inscritos en estas escuelas contra parásitos -anquilostomiasis, lombriz intestinal, triquina y esquistosomiasis. Las escuelas se fueron integrando al programa de modo gradual (escalonado) y aleatorio.

La aleatoriedad asegura que el grupo de tratamiento y el grupo de control sean comparables en el inicio, pero no puede asegurar que permanezcan comparables hacia el final del programa. Tampoco puede asegurar que la gente siga el tratamiento que les fue asignado. La vida continúa después de la selección aleatoria: otros eventos fuera del programa ocurren entre la selección aleatoria y el final. Estos eventos pueden reintroducir un sesgo de selección; pueden disminuir la validez de las estimaciones de impacto y son una amenaza a la integridad del experimento.

¿Cómo enfrentar amenazas que comúnmente afectan a la integridad de los experimentos?

Parásitos—un problema común con una solución barata

Las infecciones parasitarias representan más del 40 por ciento de la carga global de enfermedades tropicales. Las infecciones son frecuentes en áreas de escasa salubridad. Más de 2,000 millones de personas están infectadas. Los niños, los cuales todavía están aprendiendo hábitos sanitarios, son particularmente vulnerables: 400 millones de niños en edad escolar están infectados crónicamente con parásitos intestinales.

Los parásitos afectan más que la salud de los niños. Los síntomas incluyen apatía, diarrea, dolor abdominal y anemia. Más allá de su efecto en la salud y la nutrición, fuertes infecciones parasitarias pueden perjudicar el desarrollo físico y mental de los niños, además de reducir su asistencia y desempeño en la escuela.

Hábitos sanitarios e higiene deficientes facilitan la transmisión. La gente infectada excreta huevos de parásitos en sus heces y orina. En áreas de sanitación deficiente, los huevos contaminan el suelo o el agua. Otra gente es infectada cuando ingiere comida o tierra contaminada (anquilostomiasis, lombriz intestinal, triquina), o cuando larvas penetran su piel tras contacto con suelo contaminado (anquilostoma) o agua dulce (esquistosomiasis). Niños en edad escolar tienen mayor probabilidad de esparcir parásitos debido a que poseen hábitos de higiene más riesgosos (es más probable que naden en agua contaminada, más probable que no usen la letrina, más probable que no se laven las manos antes de comer). Así que tratar a un niño no sólo reduce su propia carga de parásitos, sino que puede reducir la transmisión de la enfermedad—y, así, beneficiar a la comunidad en general.

El tratamiento mata a los parásitos en el cuerpo, pero no previene la reinfección. Medicamentos orales que pueden matar el 99 por ciento de los parásitos en el cuerpo están disponibles: *albendazole* o *mebendazole* para tratar la anquilostomiasis, lombriz intestinal y triquina; y *praziquantel* para tratar la esquistosomiasis. Estos medicamentos son baratos y seguros. Una dosis de *albendazole* o *mebendazole* cuesta menos de 3 centavos de dólar mientras que una dosis de *praziquantel* cuesta menos de 20 centavos de dólar. Los medicamentos tienen muy pocos efectos secundarios.

Los parásitos colonizan los intestinos y el tracto urinario, pero no se reproducen en el cuerpo; su número aumenta solamente a través de contacto repetido con suelo y agua contaminados. La OMS recomienda que en áreas con alto grado de aparición, se realicen desparasitaciones masivas preventivas focalizadas en escuelas. Escuelas con

Parásitos: Evaluando el Programa de desparasitación en escuelas primarias

aparición de anquilostomiasis, lombriz intestinal y triquina de más del 50 por ciento debiesen tratarse masivamente con *albendazole* cada 6 meses, y escuelas con aparición de esquistosomiasis superior al 30 por ciento debiesen tratarse masivamente con *praziquantel* cada año.

Programa de desparasitación en escuelas primarias

International Child Support Africa (ICS) implementó el Programa de desparasitación en escuelas primarias (*Primary School Deworming Program* o PSDP) en el distrito de Busia en Kenia occidental, un área densamente poblada con un alto nivel de infección. El tratamiento siguió los lineamientos generales de la OMS. Los medicamentos fueron administrados por enfermeras provenientes de los servicios de salud pública y del Ministerio de salud y con la presencia de oficiales de salud del ICS.

El PSDP debía afectar la salud, nutrición y educación. Para medir su impacto ICS recolectó datos sobre una serie de resultados: presencia de infección parasitaria, carga parasitaria (severidad de la infección parasitaria); enfermedad (auto-reportada); tasa de participación escolar y puntajes de pruebas.

Diseño de la evaluación — El experimento según se planeó

Debido a restricciones administrativas y financieras, el PSDP no pudo ser implementado en todas las escuelas inmediatamente. En su lugar, las 75 escuelas fueron divididas aleatoriamente en 3 grupos de 25 escuelas, y gradualmente integradas a lo largo de 3 años. Las escuelas del Grupo 1 fueron tratadas comenzando tanto en 1998 como en 1999, las escuelas del Grupo 2 en 1999 y las del Grupo 3 comenzaron en 2001. Las escuelas del Grupo 1 era el grupo de tratamiento en 1998, mientras que las escuelas del Grupo 2 y Grupo 3 fueron el grupo de control. En 1999 las escuelas del Grupo 1 y Grupo 2 constituyeron el grupo de tratamiento y el Grupo 3 el grupo de control.

Figure 1: El experimento planeado: la línea de tiempo del PSDP mostrando los grupos experimentales en 1998 y 1999

	1998	1999	2001
Grupo 1	Tratamiento	Tratamiento	Tratamiento
Grupo 2	Control	Tratamiento	Tratamiento
Grupo 3	Control	Control	Tratamiento

Amenazas a la integridad del experimento planeado

Tema de discusión 1: Amenazas a la integridad del experimento

La aleatoriedad asegura que los grupos sean equivalentes, y por lo tanto comparables, al comienzo del programa. El impacto es estimado como la diferencia en el resultado promedio del grupo de tratamiento y el resultado promedio del grupo de control. Para poder afirmar que el programa causó el impacto es necesario poder asegurar que el programa fue la única diferencia entre el grupo de tratamiento y el de control a lo largo de la evaluación.

1. ¿Qué quiere decir que los grupos sean equivalentes al comenzar el programa?
2. ¿Se puede comprobar que los grupos son equivalentes al comienzo del programa?
¿Cómo?
3. ¿Qué puede ocurrir a lo largo de la evaluación que pudiese volver los grupos no equivalentes?
4. ¿Cómo amenaza la integridad del experimento la no equivalencia al término de la evaluación?
5. Usted aleatorizó, creando grupos equivalentes de tratamiento y control. Si los grupos permanecen equivalentes ¿Qué más puede suceder después de la asignación aleatoria, que pudiese amenazar la afirmación de que el programa fuese la única diferencia entre los dos grupos?
6. En el Caso 1, usted aprendió sobre otros métodos para estimar el impacto del programa, como diferencia simple, regresión múltiple, regresión múltiple con datos de panel y matching.
 - a. Para cada amenaza que usted acaba de identificar, explique si la amenaza existe para cada uno de estos métodos y de qué modo.
 - b. ¿Son estas amenazas a la integridad experimental exclusivas para las evaluaciones aleatorias?

Enfrentando el desgaste — cuando los grupos no se mantienen equivalentes.

Desgaste se le llama al hecho de que la gente se integre o abandone la muestra — tanto al grupo de tratamiento como al de control — a lo largo del experimento. Un ejemplo común en experimentos clínicos, es cuando la gente muere; es tan común que el desgaste se conoce en ocasiones como mortalidad experimental.

Tema de discusión 2: Enfrentando el desgaste

Usted está buscando los efectos sobre la salud de la desparasitación. En particular está observando la carga parasitaria (la severidad de la infección parasitaria). Las cargas parasitarias se miden según la siguiente escala: Cargas parasitarias grandes obtienen un puntaje de carga parasitaria de 3, cargas parasitarias medianas un puntaje de 2 y cargas parasitarias débiles un puntaje de 1.

El programa está focalizado en escuelas, por lo mismo es natural y efectivo en términos de costo recolectar datos en las escuelas — los niños están juntos en un lugar, por lo que el encuestador no necesita viajar al hogar de cada niño. El encuestador toma las mediciones para todos los niños en la escuela en un día elegido al azar (las autoridades escolares no son informadas previamente).

Existen 30.000 niños, 15.000 en escuelas de tratamiento y 15.000 en escuelas de control. Después de la asignación aleatoria se puede comprobar que los grupos son equivalentes, los niños de cada una de las tres categorías están igualmente representados.

El protocolo se sigue en un 100 por ciento: todos los niños que están en el grupo de tratamiento son tratados y ninguno de los niños en el grupo de control son tratados. La desparasitación al comienzo del año resulta en un nivel de carga parasitaria de 1 al finalizar el año debido a reinfección. Niños que tienen una carga parasitaria de 3 sólo asisten a la escuela la mitad de los días, y abandonan la escuela si no son tratados. El número de niños en cada categoría de carga parasitaria se muestran aquí tanto antes como después del experimento:

Carga parasitaria	Previo a experimento		Post experimento	
	Tratamiento	Control	Tratamiento	Control
3	5,000	5,000	0	Abandonan escuela
2	5,000	5,000	0	5,000
1	5,000	5,000	15,000	5,000
Total niños testeados en escuelas	15,000	15,000	15,000	10,000

- ¿Cuál es la carga parasitaria promedio para el grupo de tratamiento previo al experimento?
 - ¿Cuál es la carga parasitaria promedio para el grupo de control previo al experimento?
 - ¿Son los grupos equivalentes?
- ¿Cuál es la carga parasitaria promedio para el grupo de tratamiento posterior al experimento?
 - ¿Cuál es la carga parasitaria promedio para el grupo de control posterior al experimento?
 - ¿Cuál es la diferencia?
- Calcule la diferencia en el resultado al comenzar y al finalizar el año
 - ¿Es este diferencial de resultados un estimador preciso del impacto del programa?
 - Si no es preciso ¿Sobreestima o subestima el impacto real?
- Debido a que el grupo de tratamiento fue tratado, se espera que hubiese una diferencia entre los grupos a fin de año.

 - Si esta diferencia es un efecto ¿Cuál es la fuente de sesgo por desgaste, si es que lo hay?
 - ¿Cómo se puede resolver el problema para obtener una mejor estimación del impacto del programa?
- ¿Cuál es la carga parasitaria promedio del grupo de control posterior al experimento, si se incluyese a los 5.000 niños que abandonaron la escuela (asumiendo que hubiesen tenido carga parasitaria 3)?
 - Calcule el impacto del programa.
 - ¿Cuál es el tamaño del sesgo por desgaste?

Caso 4: Enfrentando las amenazas a la integridad del experimento

6. El PSPD también recolectó data sobre la tasa de asistencia escolar y resultados de pruebas.
 - a. ¿Distorsionarían algunos de estos resultados las diferencias de desgaste?
 - b. ¿El impacto sería sobreestimado o subestimado?
7. En el Caso 1, usted aprendió sobre otros métodos para estimar el impacto del programa, como diferencia simple, regresión múltiple, regresión múltiple con datos de panel y matching.
 - a. Para cada amenaza que usted acaba de identificar, explique si la amenaza existe para cada uno de estos métodos y de qué modo.
 - b. ¿Son estas amenazas a la integridad experimental exclusivas para las evaluaciones aleatorias?

Enfrentando cumplimiento parcial con el tratamiento—cuando el grupo de tratamiento no recibe el tratamiento o el grupo de control si lo recibe

Puede que algunas personas asignadas al grupo de tratamiento, al final no lo reciban. Por ejemplo, en un programa de tutorías después de clases, algunos niños a los que se las ha asignado tutorías pueden simplemente no asistir. Y otros asignados al grupo de control podrían obtener acceso al tratamiento ya sea desde el programa o una fuente externa. O niños del grupo de control podrían obtener ayuda extra de sus profesores o adquirir materiales del programa y métodos de sus compañeros de clase. De cualquier forma, esta gente no está cumpliendo íntegramente con su tarea en el experimento planeado. Esto se denomina “cumplimiento parcial con el tratamiento” o “difusión” o, menos benignamente, “contaminación”. En contraste con experimentos de laboratorio que son controlados cuidadosamente, la difusión es ubicua en programas sociales. Después de todo, la vida continúa, las personas serán personas, y no se tiene ningún control sobre que decidan hacer en el transcurso del experimento. Todo lo que se puede hacer es planear el experimento y ofrecerles tratamiento. ¿Cómo se pueden enfrentar las amenazas provenientes de cumplimiento parcial con el tratamiento?

Tema de discusión 3: Enfrentando cumplimiento parcial con el tratamiento

Ninguno de los niños de las familias más pobres poseen zapatos por lo que tienen cargas parasitarias de nivel 3. Aunque sus padres no han pagado los aranceles escolares, se les permitió a los niños quedarse en la escuela durante el año. Para tratar a los niños, se requirió un permiso parental y para consentir los padres debían ir a la escuela y firmar un documento de consentimiento en la oficina del director. Debido a que no habían pagado los aranceles, los padres más pobres estaban reacios a ir a la escuela. Por lo tanto, ninguno de los niños con carga parasitaria 3 fue tratado. Su carga parasitaria permaneció en 3 al final del año. Ninguno de los asignados al grupo de control fue tratado. A todos los niños que estaban en la muestra a principio de año, se les hizo un seguimiento ya fuera en la escuela o en sus hogares.

Carga parasitaria	Previo a experimento		Posterior a experimento	
	Tratamiento	Control	Tratamiento	Control
3	5,000	5,000	5,000	5,000
2	5,000	5,000	0	5,000
1	5,000	5,000	10,000	5,000
Total niños testeados	15,000	15,000	15,000	15,000

1.
 - a. Calcula el impacto estimado basado en las instrucciones originales.
 - b. ¿Qué mide el estimado de “intención a tratar” (ITT) de este grupo?
 - c. Esta es una estimación precisa del efecto del programa, pero ¿es una medida adecuada? ¿cuáles son las consideraciones? ¿Cuándo es útil? ¿Cuándo no lo es?

Usted está interesado en conocer el efecto del tratamiento sobre aquéllos que efectivamente fueron tratados.

Parásitos: Evaluando el Programa de desparasitación en escuelas primarias

2. Cinco de sus colegas están pasando por su escritorio; están de acuerdo en que debería calcular el efecto del tratamiento usando únicamente los 10.000 niños tratados.
 - a. ¿Cuál es el impacto de emplear solamente a los tratados?
 - b. ¿Es el consejo prudente? ¿por qué sí? ¿por qué no?
3. Otro colega afirma que no es una buena idea descartar completamente a los no tratados; debería incluirlos en el análisis pero considerarlos como parte del grupo de control.
 - a. ¿Cuál es el impacto estimado basado en esta estrategia?
 - b. ¿Es el consejo prudente? ¿por qué sí? ¿por qué no?
4. Otro colega sugiere que emplee la tasa de cumplimiento: la proporción de personas de cada grupo que se comportó de acuerdo con el tratamiento. Se debería dividir el estimador del grupo de "intención a tratar" con las diferencias en las tasas cumplimiento.
 - a. ¿Cuáles son las tasas de cumplimiento en el grupo de tratamiento y de control?
 - b. ¿Cuál es el estimado del impacto basado en esta estrategia?
 - c. ¿Es el consejo prudente? ¿por qué sí? ¿por qué no?
5. El programa elevó el nivel de conocimiento acerca de los parásitos, por lo que algunos padres en el grupo de control compraron los medicamentos y trataron a los niños en su propio hogar. En total 2.000 niños del grupo de control fueron tratados.

¿Cuál es el impacto estimado del "tratamiento sobre los tratados"?

Enfrentando externalidades —cuando el grupo de control, sin ser tratado, se beneficia del tratamiento del grupo tratado

Personas asignadas al grupo de control pueden beneficiarse indirectamente de aquéllos que reciben tratamiento. Por ejemplo, un programa que distribuye mosquiteros tratados con insecticida podría reducir la transmisión de malaria en la comunidad, indirectamente beneficiando a aquéllos que no duermen bajo las redes. Tales efectos son denominados *externalidades*.

Tema de discusión 4: Enfrentando externalidades

En el PSPD, la aleatorización ocurrió a nivel de escuela.

Las familias en las áreas de evaluación vivían en granjas cercanas una a la otra. Grupos de granjas pueden ser divididos en áreas de 3 Km de radio. Tres de tales áreas —Áreas A, B y C— se muestran en el diagrama a continuación. Las granjas están suficientemente cerca como para que niños de granjas vecinas jueguen entre ellos. Las familias también pueden elegir la escuela primaria a la cual enviar a sus hijos.

Existen tres escuelas en el área A, tres en el área B y cinco en el área C. Es común que los niños de granjas vecinas, o incluso hermanos, asistan a diferentes escuelas. Algunas de las escuelas en cada grupo de granjas formaban parte del grupo de tratamiento, otras eran de control. Las escuelas del Grupo 1 estaban en tratamiento el año 1 y los grupos 2 y 3 eran de control.

Cada escuela tiene 100 alumnos. El cumplimiento al protocolo (instrucciones) es del 100 por ciento: todos los niños en el grupo de tratamiento son tratados y todos los niños en el grupo de control no son tratados.

1. Estime el impacto comparando carga parasitaria en escuelas del grupo de tratamiento y escuelas del grupo de control.

¿Este estimado subestimaría o sobreestimaría el impacto?

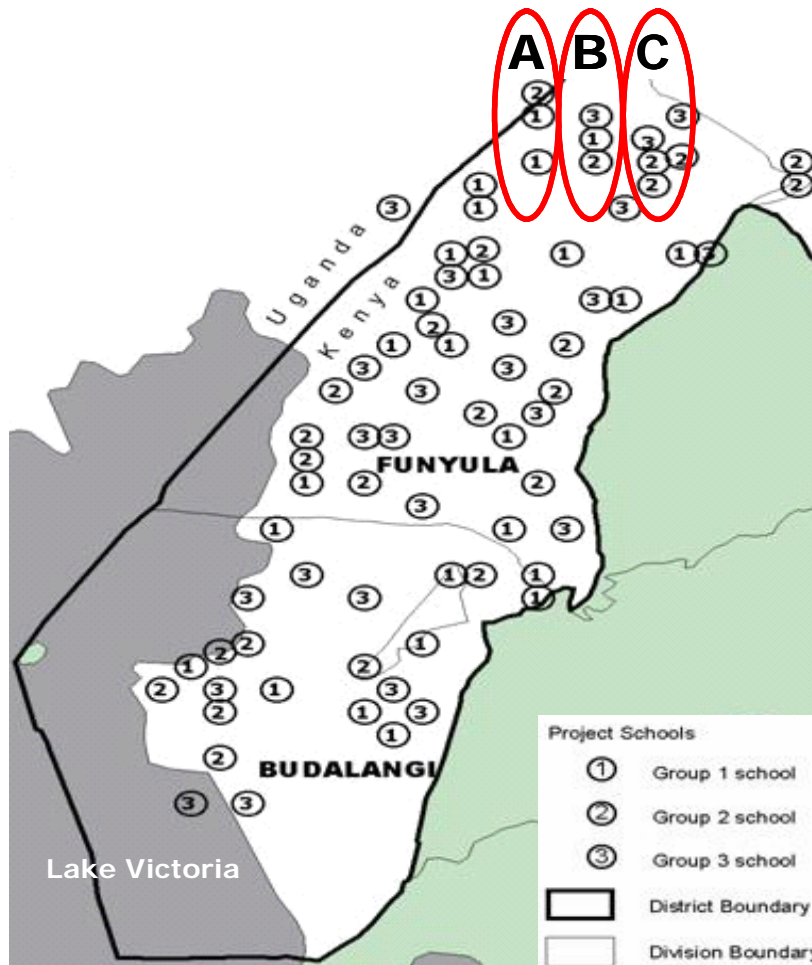
2. La densidad de tratamiento es la proporción de tratados a no tratados en un agrupamiento dado de personas.
 - a. ¿Cuál es la densidad de tratamiento en las escuelas del grupo de tratamiento en el año 1?
 - b. ¿Cuál es la densidad de tratamiento en las escuelas del grupo de control?
 - c. ¿Cuáles son las densidades de tratamiento en las áreas A, B y C en el año 1?
 - d. ¿Cuáles son las densidades de tratamiento en las áreas A, B y C en el año 2 y en

Caso 4: Enfrentando las amenazas a la integridad del experimento

Tema de discusión 4: Enfrentando externalidades

el año 3?

3.
 - a. Si hay externalidades, ¿de dónde se esperaría que provengan?
 - b. ¿Sería posible capturar el efecto de las externalidades dentro de las escuelas?
 - c. Si no espera ser capaz de capturar el efecto de las externalidades ¿Qué debiese hacer para ser capaz de capturarlo?
4. Establezca un ranking para las áreas A, B y C, en términos de la cantidad de externalidades esperadas para el año 1, 2 y 3.
5.
 - a. Si hubiese aleatorizado a nivel individual ¿Qué hubiera podido hacer para capturar externalidades interpersonales?
 - b. Si hubiese aleatorizado a nivel de escuelas ¿Qué hubiera podido hacer para capturar externalidades entre escuelas?
 - c. ¿Qué estrategia general sugieren estos resultados?



* Las ubicaciones GPS fueron recolectadas antes de Mayo del 2000, cuando los EE.UU todavía disminuía la precisión del GPS internacional. Las lecturas sólo son precisas dentro de varios cientos de metros. Así, una escuela del Grupo 3 parece encontrarse en Uganda, pero en realidad se encuentra en el lado keniano de la frontera. La escuela que parece estar en el Lago Victoria en realidad está en una isla pequeña.

**Parásitos: Evaluando el Programa de desparasitación en escuelas
primarias**

Referencias:

- Crompton, D.W.T. 1999. "How Much Helminthiasis Is There in the World?" *Journal of Parasitology* 85: 397 – 403.
- Kremer, Michael y Edward Miguel. 2007. "The Illusion of Sustainability," *Quarterly Journal of Economics* 122(3)
- Miguel, Edward, y Michael Kremer. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica* 72(1): 159-217.
- Shadish, William R, Thomas D. Cook, y Donald T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton Mifflin Company
- Banco Mundial. 2003. "School Deworming at a Glance," *Public Health at a Glance Series*. <http://www.worldbank.org/hnp>
- OMS. 1999. "The World Health Report 1999," Organización Mundial de la Salud, Geneva.
- OMS. 2004. "Action Against Worms" *Partners for Parasite Control Newsletter*, Issue #1, Enero 2004, www.who.int/wormcontrol/en/action_against_worms.pdf