

*¿Cómo aumentar la relevancia
de las evaluaciones de impacto
en la definición de políticas
públicas? Algunas reflexiones*

Emanuela Galasso

Grupo de Investigación sobre el Desarrollo

Banco Mundial

¿Qué hemos aprendido a la fecha?

- Informe de Investigación de Políticas sobre CCTs (a publicarse)
- Las evaluaciones de impacto de los CCTs ha provisto importantes resultados de forma consistente :
 - Altas ganancias en la utilización de servicios (matriculación/asistencia, atención preventiva de salud) en todos los países con evaluaciones de impacto
 - El aumento en utilización es mayor entre los más pobres y grupos de mayor riesgo (donde la brecha inicial era mayor)
- Las evaluaciones de impacto han tenido un valor intrínstico
 - Innovación: Políticas públicas basadas en evidencias empíricas. Rendición de cuentas del sector público. Quiebra con prácticas del pasado.

Sin embargo...aún quedan aspectos claves de política pública sin respuesta

- Conocimiento limitado sobre la importancia relativa entre el ingreso vs. efectos de sustitución (rol del tamaño de la transferencia vs. la condicionalidad).
 - Problemas de implementación han provisto algunas evidencias (ex. Condicionalidades no monitoreadas aún cuando los hogares perciben lo contrario) o de simulaciones-modelos estructurales sugieren que las condicionalidades son más importantes para educación que para salud.

Sin embargo...aún quedan aspectos claves de política pública sin respuesta

- Evidencia mixta/escaza sobre el impacto sobre resultados finales:
 - Aprendizajes: Hogares households with two extra years of Progresar/Op. receive 1/5 more years of education. No effect on language or math tests (Mexico, Ecuador)
 - Nutrición: initial differences in height (1 cm) dissipates after 5 years (Mexico). No effect in Honduras, Ecuador. Efectos positivos en Colombia y Nicaragua
 - Anemia: mixto
 - Salud de adultos: algunos efectos positivos sobre ADLs, obesidad, hipertensión, diabetes.
- Información limitada sobre el *costo-efectividad*

Sin embargo...aún quedan aspectos claves de política pública sin respuesta

- Evidencias limitada sobre los impactos de *largo plazo*
 - México (Gertler, Martinez, Rubio): cerca del 12% de las transferencias son invertidas en activos productivos (microempresas, agrícola)
 - Chile (Galasso, Carneiro, 2007): impacto sobre el ingreso y la incidencia de la pobreza/indigencia reaparece después de 2-4 años, principalmente a través de mejores resultados laborales
 - Veredicto sobre la graduación aún abierto

Rol de los CCT dentro de una agenda social más amplia

- *El objetivo principal de la evaluación es verificar si los programas están logrando sus objetivos*
- *Cuales son los objetivos de los CCT?*
 - Alivio de la pobreza (corto/largo plazo, crónico vs transitorio)
 - Redistribución
 - Desarrollo del Capital Humano
 - “protección social”: protección del riesgo no asegurado
- **Balance depende sobre el contrato social**
- **identificar programas complementarios /alternativos con objetivos similares (rango de instrumentos)**
 - La inversión en capital humano requiere una perspectiva de largo plazo
 - Protección Social: instrumentos que permiten a las familias responder/protegerse de shocks/riesgos en el corto plazo

¿Cuáles son las preguntas claves de política pública que las evaluaciones deben responder?

- ***¿Cuáles son los resultados que interesan desde una perspectiva de política pública?*** ¿Cuáles conductas esperamos que las intervenciones afecten? ¿Y para que grupos?
 - Aliviar la pobreza (permanente/temporal), ingreso actual/consumo vs salida sostenida de superación de la pobreza
 - Resultados finales de educación y salud (antropometría, anemia, desarrollo infantil, asistencia/retención escolar/repitencia, transición primaria/secundaria, resultados de aprendizaje)
 - Resultados Laborales (jóvenes? adultos?)

¿Cuáles son las preguntas claves de política pública que las evaluaciones deben responder?

- Por cada uno de estos resultados: ¿Cuáles son los programas sectoriales complementarios? ¿y cual es la ventaja comparativa de los CCTs en mejorar estos resultados?
 - ¿Cuál es el counterfactual analizado? (a la fecha, se a evaluado CCT vs sin programa)
 - Evaluación del sistema vs. políticas específicas?
Necesidad de innovaciones metodológicas para evaluar los “sistemas”

Entendiendo la 'caja negra'

- No solo si el programa es efectivo o no, pero también ¿porqué? \Rightarrow *identificar los canales de intervención*
 - Incluir indicadores relevantes que nos permitan desagregar los efectos relativos de cada canal
 - Combinar analisis cualitativos/cuantitativos para identificar las hipótesis a ser provadas
 - Investigación operativa y de proceso para identificar caminos (exitosos y cuellos de botella)
- Regresar a lo básico: Temas de diseño y cambios en conducta (rol de la condicionalidad, tamaño de las transferencias, incorporación) vs. temas de implementación, rol de la oferta

Asumir seriamente la heterogeneidad de impactos

- *¿Cuál es el parámetro de interés relevante?*
 - Interés focalizado en métodos que permitan medir los efectos causales de los programas que intervienen sobre los participantes. (validez interna)
- Qué es relevante para la política pública?
Impacto sobre la población objetivo, impacto marginal sobre los beneficiarios cubierto por el programa, superar el análisis de los promedios e identificar que grupos y tipo de hogares pueden beneficiarse más de las intervenciones (heterogeneidad de impacto)
 - Entender la heterogeneidad is clave para la validez externa

Validez Externa, forecasting

- *Temas de escala:*
 - Cuán diferente es un piloto vs. una versión ampliada del programa? Evaluaciones de Impacto ayudan a predecir resultados en entornos distintos (rural vs urbano, diferentes características de los hogares, características de la oferta, particularidades institucionales)
- *Flexibilidad en el diseño en la evaluación:* los programas evolucionan en el tiempo.
 - Evaluación debe acomodar nuevas preguntas de política, cambio en la naturaleza del tratamiento, permitir medidas cumulativas de efecto del programa en el tiempo.

Validez Externa, forecasting

- Las evaluaciones de impacto nos dan estimados de ‘efectos de política’ dado un diseño de política específico para una población objetivo (incluyendo las respuestas de los hogares)
- Los modelos estructurales que usan datos de Progresas/Oportunidades para proponer cambios en el diseño del programa (variaciones sobre límites de edad/montos de transferencias)
 - Attanasio, Meghir, Santiago
 - Todd, Wolpin

Algunos temas de discusión para mejorar la relevancia de la intervención

- Prioridad de responder las preguntas de política pública sobre metodologías específicas
- Una evaluación no es suficiente para responder a todas las preguntas relevantes para la política pública
- Los impactos pueden tomar tiempo en manifestarse, y el tiempo de las evaluaciones puede no corresponder con las necesidades operativas
- Cómo evaluar lo que funciona?

Evaluación de impacto básica para rendición de cuentas políticas, apoyo sostenido

- Argumento de economía política: objetivo de redistribución “efectiva” hacia los pobres (o vulnerables): llegar a los beneficiarios meta y proveer evidencia de cumplimiento de la condicionalidad
 - Se necesitan diagnósticos del desempeño de la focalización al programa (enfoque en minimizar los errores de inclusión, a expensas de los errores de exclusión)...
 - Combinar con evidencia acerca ganancias en acceso y el uso de los servicios

Evaluación de impacto básica para rendición de cuentas políticas, apoyo sostenido

- Rol de los datos administrativos, generalmente subenfatisado, como sustituto o complemento a los datos de encuestas:
 - La efectividad de la focalización puede ser evaluada usando la información de las fichas de ingreso (o indicadores de vulnerabilidad de Chile)
 - Acceso/utilización de servicios pueden ser recabadas de los datos administrativos propios de los oferentes (escuelas, records de los puestos de salud) – generalmente más disponibles en educación que en salud

Estudios que usan datos administrativos:

- **Ex1. Bolsa Escola Brazil: de Janvry, Finan, Sadoulet (2007):**
 - Records escolares seleccionados de manera aleatoria de dos escuela en cada municipio (asistencia, promoción de grado/repitencia, tasas de deserción, pruebas de aprendizaje).
 - Estatus de beneficiarios de los datos administrativos de Bolsa Escola. Consolidación por nombres
 - Encuesta municipal sobre la capacidad administrativa, influencia politica e implementación del programa
 - Censo de Población
 - Métodos no-experimentales (diferencias en diferencias, child fixed effects)
 - El program redujo la deserción en 7.8%, aumento la repitencia en 0.8%.
 - Enorme variación en el desempeño entre municipalidades (mayor en municipalidades con procesos democráticos competitivos, monitoreo más estricto de las reglas del programa)

Estudios que usan datos administrativos :

- **ex2: Incidencia de la Mortalidad Infantil y Materna en Progres/Oportunidades (Barham 2005)**
 - Usa estadísticas vitales municipales sobre número de nacimientos/muertes en las áreas rurales
 - Método de diferencias-sobre-diferencias
 - Muertes infantiles 2/1000 menores en comunidades del programa (reducción del 11%)
- **Ex3: Visitas de las clínicas de salud Progres (Gertler 2000) –ejercicio de validación**
 - Usa la muestra experimental
 - Tendencias en el número de visitas a los puestos de salud es similar en comunidades control y tratamiento antes del programa
 - Las comunidades en el programa muestran un mayor número de visitas (resultados similares que si usara la encuesta de la evaluación de impacto)

Estudios que usan datos administrativos :

- ***ex 4: Efectos de largo plazo del programa de becas para la educación secundaria en Colombia (Angrist et al 2004)***
 - Programa se focalizó en familias de bajos ingresos. Voucher para cubrir el costo promedio de la educación privada
 - Voucher para secundaria asignados por lotería : compara los ganadores-perdedores de la lotería.
 - La evaluación uso encuestas hechas con entrevistas telefónico sobre los que solicitaron el programa (Angrist et al 2003)
 - Efectos de corto/mediano plazo : ganadores 16% mayor probabilidad de asistir a las escuelas privadas. Efecto sobre terminación del 8th grado (10%), a través de menor repitencia, mayores resultados en pruebas de desempeño (+0.2SD)
 - Analysis de largo plazo: Unifica una muestra de interesados en participar en el programa con los records administrativos si estos se registraron en un examen de ingreso a la universidad (ICFES test). La unificación se da utilizando los números de identidad nacional
 - Efectos de largo plazo: PACES aumentó la tasa de terminación de la escuela secundaria por 15-20% y en las pruebas por 0.2SD

Uso datos administrativos :

- *Ventajas*
 - Costo-efectivo
 - Muestras amplias (por definición)
 - Ganancias en el acceso puede ser evaluado en un punto en el tiempo a en periodos más amplios (impactos de largo plazo, aspectos dinámicos). Vinculación natural con la evaluación de largo plazo
 - Lecciones importantes en como armar una base de datos administrativos que pueden estar ligados a otras bases de datos (identificación nacional, identificación geográfica, claves de escuelas/puestos de salud), estos son actualizados de forma regular y tienen controles de calidad internos (experiencia del SIIS en Chile)
 - Es necesario incorporar indicadores de calidad dentro del sistema

Uso de datos administrativos :

- *Limitaciones:*

- Los datos administrativos suelen tener información limitada sobre impactos finales o resultados intermedios. Es más difícil establecer una fotografía de la historia completa.
- Generalmente los datos están limitados a características observables a nivel de los beneficiarios o los hogares (es más difícil tratar la heterogeneidad de los impactos).
- No siempre es posible observar niños/hogares que se mueven de la escuela o de la municipalidad? Otros hogares fuera del sistema?

Evaluación de Impacto usando encuestas propias ('rolls royce')

- Escala más amplia (el foco no es sólo en acceso/utilización, sino en mayores resultados, heterogeneidad, spillovers, impactos de largo plazo)
- Analizar en profundidad si un diseño específico funciona, por qué y para quién?
- Not a one-shot deal
- Las metodologías se han tenido que adaptar a las necesidades cambiantes: Aleatorización en las etapas tempranas de Progresión (randomized rolled-out) tuvo que ser complementado con métodos no experimentales para analizar impactos de medio y largo plazo (Berhman, Todd, Parker , Parker, Todd Wolpin)

Cómo se puede maximizar el impacto con diseños complementarios-alternativos?

- Depende del resultado en el que estemos interesados
- Uso de análisis cualitativo, investigación operativa iterativa para ayudar a identificar las preguntas
- Experimentación a nivel de pilotos menores con diferentes modelos de entrega
- Establecer asociaciones formales con diferentes actores:
 - Instituciones de investigación local (eg. INSP)
 - Agencias de implementación local
 - Ministerios Sectoriales (salud, educación, trabajo)
 - Institutos de Estadísticas Nacionales

Qué funciona para efectos específicos?

Salud

- *Salud infantil = f(salud materna, educación materna, nutrición, exposición al medio ambiente en el que viven, cuidado de la salud)*
- Intervención sobre la demanda del CCT en salud
 - Efecto ingreso
 - Indirecto sobre conocimiento materno
- Cuales son los insumos complementarios para mejorar los resultados en nutrición y salud? Deben ser entregados como parte del CCT?

Qué funciona para efectos específicos?

Salud

- Alto retorno de inversión temprana. Retornos de largo plazo en cuanto a escolarización, desarrollo cognitivo.
- nutrición
 - Ventana de oportunidad: evidencia sólida que la altura se define en los primeros 2/3 años de vida
 - Suplementación en micronutrientes costo-efectiva
 - Consumo alimentario solo un insumo. Calidad del consumo en alimentos es igualmente importante
 - Protección contra la diarrea y enfermedades respiratorias
 - Prácticas de cuidado infantil (calidad de la dieta y prácticas de alimentación, estimulación temprana)

Qué funciona para efectos específicos?

Salud

- Porqué no encontramos efectos sistemáticos en nutrición? Necesidad de identificar hipótesis :
 - Falta de inversión durante el período de edad crítica?
Pérdida Irreversible
 - Falta de cambio en las conductas maternas? (calidad consejería/pláticas)
 - Suplementación alimentaria no están llegando a los beneficiarios metas (respuestas conductuales de las familias)

Qué funciona para efectos específicos?

Pruebas de aprendizaje-Desarrollo infantil

- *Desempeño/resultados cognitivos=f(insumos familiares, insumos escolares, atributos)*
- Proceso acumulativo: ‘destrezas engendran destrezas’, comienzo temprano, es importante contabilizar insumos pasados y corrientes
 - Evidencia US (Todd, Wolpin 2003, 2005, Carneiro Heckman Masterov 2002):
 - Insumos del hogares – crianza - (actual y pasada) son determinantes importantes de las brechas raciales en los resultados de las pruebas de rendimiento escolar
 - Insumos Escolares (tasa alumno/maestro, gastos por estudiante) significativo pero no explican mucho de las diferencias
 - incentivos (gestión alumno/maestro) sobre la oferta

Pruebas de aprendizaje-Desarrollo infantil

- Por qué no vemos efectos sistematicos sobre el aprendizaje? Necesidad de identificar hipótesis :
 - Los niños que accedan a la escuela como resultado del programa son más pobres que el resto
 - Baja calidad de las escuelas están concentrado en las escuelas que atienden a los niños en desventaja (eg. pertinencia curricular, métodos pedagógicos, bilinguismo)
 - Falta de insumos del hogar complementarios (crianza, nutrición, valor de la educación para la familia)

La ventaja de invertir in paneles de largo plazo

- Impactos de corto vs largo plazo (pobreza crónica vs transitoria). Cantidad de entradas y salidas de la pobreza
- Se puede analizar los insumos del pasado
- Se puede analizar la elegibilidad y el acceso a un mayor grupo de programas
- Ejemplo PSID (Panel Study on Income Dynamics) en los EEUU. Encuestas son el instrumento principal para analizar las reformas en el sistema de bienestar de los EEUU
- (en línea con el objetivo de analizar la eficiencia de los sistemas)
 - Mismas familias seguidas por 34 años (1968-2005)
 - Muestra inicial + seguimiento de familias que se desdoblaron
 - Actualización de la muestra (para mantener la representatividad de cada encuesta)
 - Rango de resultados (empleo, ingreso, gastos, participación en el programa, vivienda, crianza, historial matrimonio/fecundidad, salud, salud mental, educación, bienestar)
 - Módulos específicos fueron agregados
 - Equipo Interdisciplinario

La ventaja de invertir in paneles de largo plazo

- Ex. Pueden analizar los efectos de largo plazo de la participación en el programa (ex. Head Start)
 - Currie, Thomas (1995) analizaron a los adultos en el PSID y compararon a los herman@s que participaron vs los que no participaron en Head start (controlando por características del hogar). Aquellos que participaron en Head Start cuando jóvenes tienen mayores probabilidades de terminar la secundaria, universidad y posiblemente tener mayores ingresos.