



**Gobierno de Chile  
Ministerio de Hacienda  
Dirección de Presupuestos**

# **Evaluación Ex-Post: Conceptos y Metodologías.**

**División de Control de Gestión**

**Santiago, Junio 2015**

**CHILE**

## **Presentación**

El presente trabajo representa un esfuerzo por actualizar e integrar en un solo documento los conceptos y metodologías aplicadas en la evaluación ex post de programas e instituciones públicas desarrollados por la División de Control de Gestión de la Dirección de Presupuestos.

Así, este Documento Metodológico tiene por objetivo presentar los instrumentos de evaluación de que dispone la Dirección de Presupuestos en el contexto del Sistema de Evaluación y Control de Gestión y Presupuesto por Resultados y, a la vez, ser una referencia para los distintos actores que participan en la implementación de éstos en lo concerniente a las metodologías que se utilizan para ello.

Su elaboración se basa en los diferentes manuales y guías metodológicas preparadas con anterioridad en la División de Control de Gestión, en elementos conceptuales extraídos de la literatura especializada, y en la experiencia recogida por los profesionales de la División de Control de Gestión en los últimos años.

La Sección I del documento define brevemente el Rol de Estado y la generación de Políticas Públicas, contextualiza el rol de las evaluaciones gubernamentales en el ámbito del Presupuesto y se detallan las líneas de evaluación que ha desarrollado la Dirección de Presupuestos. La sección II explica los conceptos básicos, de carácter transversal, a tener en cuenta al momento de evaluar programas e instituciones. La Sección III describe los principales ámbitos que deben ser analizados para evaluar un programa o institución, y la sección IV describe las metodologías aplicadas por DIPRES en las distintas líneas de Evaluación.

**PAULA DARVILLE ÁLVAREZ**  
**Jefa División de Control de Gestión**  
**Dirección de Presupuestos**

## Contenidos

I.	INTRODUCCIÓN .....	4
I.1	Rol del Estado .....	4
I.2	Ciclo Presupuestario y Rol de las Evaluaciones.....	5
I.3	Sistema de Evaluación Ex Post en Dipres .....	7
II.	DEFINICIÓN DE PROGRAMA Y RELACIONES DE CAUSALIDAD .....	8
II.1	Definición de Programa.....	8
II.2	Relaciones de Causalidad .....	9
II.3	Definición de Población Potencial, Objetivo y Beneficiarios Efectivos .....	9
II.3.1	Focalización .....	10
III.	ÁMBITOS DE LA EVALUACIÓN EX - POST DE PROGRAMAS E INSTITUCIONES 11	
III.1	Diseño.....	11
III.2	Organización y Gestión.....	12
III.3	Dimensiones o Foco del Desempeño .....	12
III.4	Sostenibilidad .....	15
III.5	Enfoque de Género .....	15
IV.	METODOLOGÍAS E INSTRUMENTOS .....	15
IV.1	Metodología de Marco Lógico .....	16
IV.1.1	Objetivos .....	18
IV.1.2	Indicadores.....	19
IV.1.3	Medios de Verificación.....	21
IV.1.4	Supuestos .....	21
IV.1.5	Lógica Vertical y Horizontal.....	21
IV.2	Metodologías de Evaluación de Impacto .....	22
IV.2.1	Estimación del Impacto de una Intervención.....	23
IV.2.2	Efectos Heterogéneos .....	37
IV.2.3	Población Objetivo y Marco Muestral.....	39
IV.3	Metodología Evaluación del Gasto Institucional (EGI).....	41
IV.4	Metodología Evaluación de Programas Nuevos .....	42

IV.5 Instrumentos de Recolección de Información .....	42
Anexo I: Rol del Estado y Fallas de Mercado.....	44
Anexo II: Definición de Conceptos para el Análisis de Uso de Recursos.....	46
Anexo III: Diseño Muestral .....	51
Anexo IV: Diseños de Regresión Discontinua.....	54
Anexo V: Análisis de Género en la Formulación de Programas .....	56
BIBLIOGRAFÍA.....	58

## **I. INTRODUCCIÓN**

La Dirección de Presupuestos, ha desarrollado e introducido en el proceso presupuestario diferentes instrumentos de monitoreo y evaluación en las últimas décadas, con el objeto de mejorar el análisis, evaluación y formulación del presupuesto público, lo que ha contribuido a orientar los recursos hacia el logro de las políticas y objetivos de gobierno.

Para esto, se ha implementado, gradual y progresivamente, el Sistema de Evaluación y Control de Gestión, que tiene por objeto promover la eficiencia y la eficacia en la asignación y uso de los recursos públicos, a través de la aplicación de técnicas e instrumentos de evaluación y control de gestión ligados al presupuesto. Este Sistema, estrechamente vinculado al ciclo presupuestario, ha permitido disponer de un modelo de evaluación y seguimiento del desempeño de la gestión presupuestaria, retroalimentando la toma de decisiones.

Así, el Sistema de Evaluación y Control de Gestión comprende distintos instrumentos integrados en forma sinérgica al ciclo presupuestario, asociados a las cuatro etapas de la gestión presupuestaria: a) Formulación, b) Discusión/Aprobación, c) Ejecución y d) Evaluación. El avance en el desarrollo y consolidación de cada uno de los instrumentos, ha logrado fomentar e impulsar el mejoramiento de la gestión en los programas e instituciones públicas<sup>1</sup>.

### **I.1 Rol del Estado**

En general la literatura define tres funciones en las cuales el Estado puede intervenir en los mercados<sup>2</sup>. El primero de ellos tiene relación con la estabilización, es decir, mantener a la economía en su nivel de pleno empleo y precios estables. Una segunda función se relaciona con la asignación, esto es la manera en que se ubican los recursos dentro de la economía. Una tercera función mediante la cual el Estado puede intervenir es la distribución de los recursos de la economía entre los diferentes miembros de la sociedad<sup>3</sup>.

Los argumentos presentados para las dos primeras funciones se basan en fallas de mercado que impiden a la economía alcanzar su punto de equilibrio eficiente. Sin embargo, aún si la economía fuese eficiente, existen otros argumentos que justifican la intervención del Estado. El principal de ellos tiene relación con que un equilibrio eficiente en la economía no necesariamente coincide con la distribución de ingresos socialmente deseada. En este sentido, el Estado tiene como una de sus actividades más importantes la redistribución de la renta, la cual se concreta a través de los programas sociales implementados por los diferentes Ministerios.

---

<sup>1</sup> Dirección de Presupuestos (2013), pág. 85.

<sup>2</sup> Para descripción más detallada del rol de Asignación del Estado y las posibles fallas de mercado, ver Anexo 1.

<sup>3</sup> Stiglitz (2000).

El Ministerio de Hacienda<sup>4</sup> recoge las funciones del Estado explicitadas anteriormente y las operacionaliza en tareas concretas, entre ellas:

- Creación de mercados cuando estos no existen.
- Corrección de las fallas de mercado, cuando éstas sean verificables y existan mecanismos técnicos apropiados para hacerlo.
- Regulación de mercados que no son plenamente competitivos o donde las externalidades son importantes.
- Generación de políticas que permitan a todos los ciudadanos alcanzar condiciones sociales mínimas, las que se reflejan en los programas impulsados en las áreas de salud, educación y vivienda.

El Estado, como generador de planes, programas y políticas, también debe velar por el diseño, implementación y resultados de cada uno de estos cursos de acción. Esto se fundamenta en la necesidad de “asegurar la calidad de un gasto que es realizado con recursos públicos y que, por lo tanto, concierne a todos los habitantes del país”.

Para ello, la Dirección de Presupuestos (en adelante DIPRES) tiene como misión velar por una asignación y uso eficiente de los recursos públicos en el marco de la política fiscal<sup>5</sup>. En este contexto, la evaluación de políticas, programas e instituciones públicas es esencial para cumplir con esta misión, ya que entrega información respecto de su funcionamiento, así como de las fortalezas y debilidades de éstas.

## **I.2 Ciclo Presupuestario y Rol de las Evaluaciones**

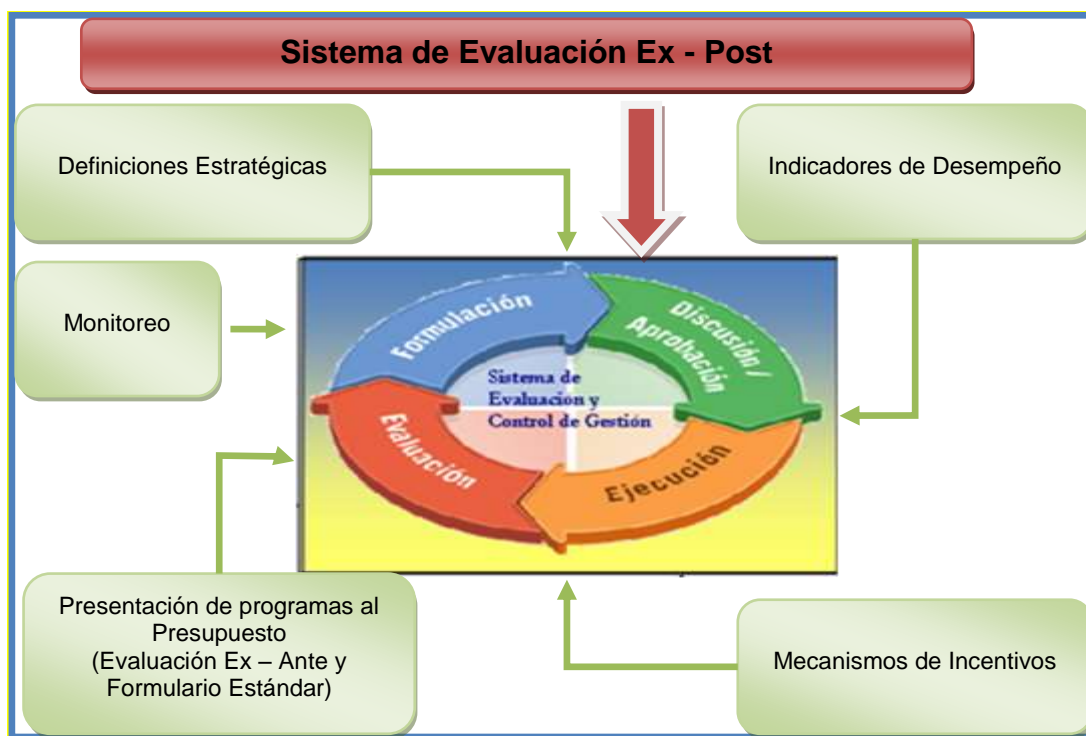
Los instrumentos comprendidos en el sistema de evaluación y control de gestión se encuentran integrados al ciclo presupuestario, en sus distintas etapas, tal como se expone en la Figura 1.

---

<sup>4</sup><http://www.hacienda.cl/preguntas-frecuentes/deberes-y-funciones/cual-es-la-politica-economica-del.html>

<sup>5</sup> <http://www.dipres.gob.cl/594/w3-propertyvalue-2128.html>

**Figura 1: Ciclo Presupuestario y Sistema de Evaluación y Control de Gestión**



Fuente: DIPRES

Como es posible observar, existen diversos instrumentos para dar cuenta del objetivo propuesto de velar por la calidad del gasto. El desarrollo y perfeccionamiento de éstos se ha realizado en forma gradual, dando forma al actual “Sistema de Evaluación y Control de Gestión” de DIPRES. En efecto, el año 1993 y a partir de una experiencia piloto se desarrollaron indicadores para un conjunto acotado de Servicios. A partir de 2001, se solicita información de indicadores de desempeño, a través de un formato estándar único (Formulario H), que se elaboran a partir de las Definiciones Estratégicas<sup>6</sup> presentadas por las instituciones; permitiendo, de esta forma, alinear los procesos de planificación con los recursos disponibles.

Luego, se ha avanzado en modelos más completos y complejos de evaluación ex-post y modelos de evaluación ex-ante, cuyo objetivo es contar con programas bien diseñados, tanto en lo conceptual como en su diseño operativo.

Por último, se destaca que toda la información generada a través de este sistema es de carácter público y se envía al Congreso, propiciando la transparencia en el uso de los recursos públicos y la rendición de cuentas.

<sup>6</sup> La información corresponde a la misión, objetivos estratégicos, productos estratégicos (bienes y/o servicios) y usuarios, beneficiarios o clientes.

### I.3 Sistema de Evaluación Ex Post en Dipres<sup>7</sup>

En 1997 se inició el Programa de Evaluación de Programas Gubernamentales (línea EPG) que, en su fase piloto, contempló la evaluación de 20 programas públicos y que respondió a una serie de compromisos establecidos en el marco de la aprobación de la Ley de Presupuestos de ese año, tendientes a una mayor evaluación y transparencia de la gestión pública. A partir del año 2003 —en el marco de lo establecido en el artículo 52 del DL N°1.263 de 1975, Orgánico de Administración Financiera del Estado y su reglamento<sup>8</sup>— se le otorga la facultad a la Dipres para efectuar evaluaciones de los programas sociales, de fomento productivo y de desarrollo institucional incluidos en los presupuestos de los servicios públicos que se determinen mediante uno o más decretos del Ministerio de Hacienda.

Cada año, en el marco de la tramitación del Proyecto de Ley de Presupuestos, los programas públicos e instituciones a evaluar son acordados con el Congreso Nacional. Dicho compromiso se plasma en un Protocolo de Acuerdo entre el Ministerio de Hacienda y el Congreso Nacional.

Para asegurar la calidad, transparencia y uso de las evaluaciones, se han establecido, entre otros, requisitos de independencia de la institución evaluada, siendo éstas desarrolladas por entidades externas al sector público<sup>9</sup>, y transparencia al hacer públicos sus resultados. Dichas condiciones se han mantenido hasta hoy.

Gradualmente, el sistema ha desarrollado diferentes líneas de evaluación ex-post, que tienen distinto alcance y profundidad:

- **Evaluación de Programas Gubernamentales (EPG):** A través de esta línea se evalúa la consistencia de los objetivos y diseño del programa, aspectos de su organización y gestión y resultados a nivel de producto (cobertura, focalización, entre otros). Iniciada el año 1997.
- **Evaluación de Impacto de Programas (EI):** El foco de atención está centrado en evaluar resultados intermedios y finales. Se utilizan metodologías cuasi-experimentales. Iniciada el año 2001.
- **Evaluación del Gasto Institucional (EGI):** Tiene como objetivo evaluar el diseño y gestión institucional, así como los resultados y uso de recursos en la provisión de los productos estratégicos de la institución. Iniciada el año 2002.
- **Evaluación de Programas Nuevos (EPN):** Tiene como objetivo diseñar la evaluación desde que se planifica un nuevo programa público, incorporando preferentemente la construcción de un grupo de control, obtenido de forma experimental. Iniciada el año 2009.

La evaluación de programas ha conseguido importantes logros tal como lo señala el Banco Mundial<sup>10</sup> en la evaluación de impacto llevada a cabo del Programa de

<sup>7</sup> Dirección de Presupuestos (2013), pág. 86-87.

<sup>8</sup> Reglamento N° 1177, 17 de diciembre 2003.

<sup>9</sup> Para cumplir con el requisito de independencia se encargan las evaluaciones a Paneles de Expertos, integrados por 3 consultores de reconocido prestigio; o universidades o empresas consultoras a través de procesos de licitación pública.

<sup>10</sup>Banco Mundial (2005).



Evaluación de Dipres. Entre los resultados más importantes se destaca su contribución a la toma de decisiones presupuestarias, mejoramiento de la gestión y la rendición de cuentas; aportando, de esta forma, al mejoramiento de la eficacia y eficiencia del gasto público. Asimismo, el programa de evaluación ha permitido la formación de una cultura creciente de evaluación en el sector público y ha contribuido a la transparencia de la gestión gubernamental.

## **II. Definición de Programa y Relaciones de Causalidad**

### **II.1 Definición de Programa**

La puesta en práctica de las políticas públicas se materializa a través de programas, iniciativas, proyectos y del accionar en general de las instituciones públicas. Un programa es un conjunto de actividades necesarias, integradas y articuladas que proveen bienes y/o servicios (productos), tendientes a lograr un objetivo específico en una población determinada, de modo de resolver un problema o atender una necesidad que la afecta.

En términos operativos un programa:

- Debe ser fácilmente identificable dentro del ámbito de acción de una o más instituciones públicas. Ello significa que debe llevar asociado un nombre que lo caracterice, una estructura organizacional, un responsable asociado claramente identificable y un presupuesto consistente y pertinente de gastos asociados a los bienes y servicios que produce.
- Debe permitir identificar, total o parcialmente, su vinculación con una política pública vigente, ya sea de Estado o de Gobierno (sectorial o global, local o nacional). Adicionalmente, debe ser consistente con la misión y los objetivos estratégicos de la/las instituciones que lo implementan.
- Debe tener un diseño, que en su conjunto, dé cuenta de una estrategia para dar solución al problema y/o necesidad concreta que afecta a la población objetivo identificada.
- Debe identificar claramente la población objetivo que será beneficiaria de los bienes y/o servicios provistos.
- Debe identificar cuáles son los bienes o servicios públicos provistos por el programa, así como los resultados esperados de su ejecución y forma de medición. Respecto al alcance de los resultados, un programa debe estar explícitamente formulado para lograr un efecto o cambio en los beneficiarios; esto es: un resultado intermedio o final (ver sección IV.1.2)

Por otra parte, también es posible distinguir otras iniciativas programáticas, que si bien tienen un alcance más acotado que la definición de programa, es igualmente una forma en que el Estado concreta sus actividades, y por tanto, sujeto de ser evaluada. Nos referimos a otras iniciativas programáticas como un conjunto de actividades necesarias, integradas y articuladas para proveer de forma regular ciertos bienes y/o servicios a beneficiarios de la sociedad civil, o en su defecto a otras organizaciones del

Estado. Si bien pueden basar su diseño en una “hipótesis de cambio” y abordan problemas o necesidades de carácter público, tienen un alcance más acotado por lo cual no requieren identificar o medir sus resultados intermedios o finales (bonos o subsidios asistenciales, ciertas becas (alimentación), prestaciones curativas de salud, etc.).

## **II.2 Relaciones de Causalidad**

El impacto o efecto de un programa sobre su población beneficiaria es el cambio que se produce sobre variables específicas que el programa espera modificar en la población intervenida, aislando todos los demás factores que pueden afectarla. Dicho de otra forma, el impacto es la diferencia entre la situación real de la población beneficiaria después del programa y la situación en la que esa población habría estado si no hubiese participado en él.

Para conocer los efectos de un programa, en primer término se deben identificar las relaciones teóricas de causa-efecto entre los bienes o servicios que proporciona la intervención y los cambios en la situación de los beneficiarios. Cada bien o servicio entregado debería afectar una o más variables de interés, lo que determinará los resultados esperados.

En general, la gerencia de cada programa maneja explícita o implícitamente ciertas hipótesis respecto a la relación causa-efecto entre los componentes y los impactos buscados. El rol del evaluador es indagar sobre la real existencia de esas relaciones causa-efecto; validarlas si corresponde, o rechazarlas. Los instrumentos que se utilizan en DIPRES para ordenar estas relaciones causa-efecto son: la matriz de marco lógico y modelos conceptuales.

La evaluación debe demostrar empíricamente la existencia de las relaciones causa-efecto supuestas por los diseñadores y encargados de la intervención.

## **II.3 Definición de Población Potencial, Objetivo y Beneficiarios Efectivos**

- Población Potencial: Corresponde a la población que presenta una necesidad o problema público, identificado y definido como tal por una política diseñada para abordarlo.
- Criterio de Focalización: Condiciones o características que deben cumplir los beneficiarios potenciales (población potencial), para ser considerados como parte de la población objetivo del programa.
- Población Objetivo: Es aquella parte de la población potencial que cumple con los criterios de focalización y que, por lo tanto, el programa planifica atender en el mediano y largo plazo.
- Criterios de Priorización: Criterios objetivos que permitan ordenar el flujo de beneficiarios dentro de un plazo plurianual, determinando en forma no arbitraria a quiénes se atiende antes y a quiénes después.

- **Población Beneficiaria:** Es aquella parte de la población objetivo que cumple los criterios de priorización y que, por lo tanto, el programa planifica atender en un año determinado.

Para evaluar de mejor forma la capacidad de la intervención o programa para responder a sus compromisos, se debe distinguir entre el concepto clásico de cobertura del programa, calculado como la relación entre la población efectivamente atendida (beneficiarios efectivos) y la población potencial; de la cobertura planificada, es decir, la relación entre la población efectivamente atendida (beneficiarios efectivos) y la población que el programa tiene considerado o planificado atender (población objetivo).

Adicionalmente, se incorporan los criterios de priorización los cuales corresponden a criterios objetivos (verificables y no arbitrarios), que permitan ordenar el flujo de beneficiarios dentro de un plazo plurianual, determinando en forma no arbitraria a quiénes se atiende antes y a quienes después.

El análisis de la población potencial u objetivo es importante en la medida en que permite dimensionar el tamaño y duración del programa. No obstante, si el problema o necesidad que aborda el programa dice relación con un bien público o superior de alcance nacional o mayor, que no se manifiesta específicamente en un conjunto de población, sino que afecta al conjunto de la sociedad, el dimensionamiento del programa no estará basado en la población potencial u objetivo sino en aquellos elementos que serán utilizados como medio para producir el cambio deseado (por ejemplo, problemas relacionados con la contaminación del aire, del agua, el patrimonio cultural, biodiversidad).

La población potencial y objetivo se caracterizan y cuantifican en términos de: edad (tramos etarios si corresponde); educación (años de estudio, nivel educacional); sexo (por categorías: solamente mujeres, solamente hombres, preferentemente mujeres, preferentemente hombres, sin distinción de sexo); nivel socio-económico (quintil de ingreso, si corresponde); localización (rural/urbana); tamaño, región; principales actividades económicas, características culturales u otros atributos que sean pertinentes.

### **II.3.1 Focalización**

El concepto de focalización se refiere a concentrar la provisión de bienes y/o servicios en una población de beneficiarios objetivos claramente identificada, a la que se quiere llegar con determinado programa o proyecto (Franco, 1995).

Se estipulan criterios de focalización cuando existe un grupo dentro de la población potencial que tiene más restricciones para acceder de alguna forma a los bienes o servicios que ofrece el programa. Por su parte, la determinación de la focalización se encuentra asociada al proceso de diseño de la intervención o programa. Evaluar la focalización en un programa público implica centrar la atención en los elementos de diseño, aplicación y resultados.

En relación al diseño, se debe identificar si el programa evaluado es o no focalizado de acuerdo a su objetivo; en una segunda etapa se debe reconocer el tipo de focalización (individual, grupal o autoselección)<sup>11</sup> así como los instrumentos y procedimientos mediante los cuales se implementa.

Respecto de la aplicación, se debe identificar si los instrumentos y procedimientos funcionan según lo diseñado.

Finalmente, en relación a los resultados, se debe evaluar si el programa ha logrado llegar a la población que deseaba atender, para lo cual se debe determinar si existe o no errores de inclusión o exclusión en la focalización (Cuadro 1). Para esto es necesario utilizar indicadores consistentes con los criterios definidos. Ejemplo de éstos pueden ser indicadores tales como, el porcentaje de beneficiarios pertenecientes al primer quintil de ingreso, porcentaje de beneficiarios bajo la línea de pobreza, etc. cuando el criterio es socioeconómico; e indicadores como el porcentaje de beneficiarios pertenecientes a cierta etnia, o el porcentaje de beneficiarios hombres, cuando el criterio de focalización es étnico o de género, respectivamente.

**Cuadro 1: Inclusión y Exclusión por Focalización**

Seleccionado	Cumple con criterio focalización	No cumple con criterio de focalización
<b>Si</b>	Éxito	Error Inclusión
<b>No</b>	Error Exclusión	Éxito

Fuente: Elaboración propia.

Muchos programas con modalidad de fondos concursables están dirigidos a una determinada población objetivo que es quien puede participar de dichos concursos. En estos casos además es necesario aplicar criterios para seleccionar los beneficiarios efectivos del fondo, los que también deben ser considerados en el proceso de evaluación.

### III. Ámbitos de la Evaluación ex - post de Programas e Instituciones

Dipres pone el foco en evaluar ciertos ámbitos de los programas e instituciones, de tal forma que sea posible emitir juicios evaluativos de cada uno de ellos y, a su vez, éstos conformen un juicio global de desempeño. Los ámbitos analizados son los siguientes:

#### III.1 Diseño

Corresponde a la articulación entre los elementos que componen un programa o institución y a través de la cual se pretende lograr los objetivos planteados.

<sup>11</sup>Se entiende por: i) evaluación individual, cuando se considera criterios como ingreso o riqueza, o una entrevista directa de un asistente social u otro profesional; ii) evaluación de grupos, donde se consideran criterios agregados de un determinado grupo, por ejemplo: áreas geográficas, zonas rurales y/o urbanas o establecimientos, y iii) autoselección, donde el individuo debe tomar ciertas decisiones para acceder al beneficio, ej: programas que otorguen beneficios de sala cuna sólo en el caso en que la madre trabaje.

Cuando se evalúa el diseño de un programa se analiza las relaciones de causalidad y si éstas están debidamente identificadas a través de la definición de sus distintos niveles de objetivos para dar respuesta al problema identificado y sus causas.

En el caso de las instituciones, se analiza la consistencia de sus definiciones estratégicas (misión, objetivos y productos estratégicos) y como su estructura organizacional responde a dichas definiciones. Para ello, se evalúa si los bienes y/o servicios (productos estratégicos) que entrega la institución son los necesarios y suficientes para cumplir con los objetivos institucionales y si los objetivos definidos permiten cumplir la misión. También considera el análisis del marco legal en que se basa el accionar de la institución.

Por último, se evalúa la calidad y pertinencia de los indicadores definidos para medir los objetivos planteados. Es importante asegurar que el diseño de los indicadores sea consistente con los objetivos que se quiere medir.

### **III.2 Organización y Gestión**

Corresponde a los mecanismos de operación y producción de los bienes y servicios provistos por el programa o institución, y cómo éstos se implementan dentro de la organización.

Al evaluar, se analizan los principales procesos, instancias de coordinación e instrumentos para desarrollar las actividades y cumplir los objetivos. Deben observarse aspectos tales como:

- Procesos de producción de los bienes y servicios entregados por el programa o institución.
- Mecanismos de coordinación al interior de la institución (en sus distintos niveles funcionales y/o territoriales) y con otras instituciones relacionadas con la ejecución.
- Gestión directa de los recursos, evaluar los criterios de asignación de recursos a los distintos bienes y servicios que entrega el programa, y cómo estos criterios afectan la cantidad y combinación de bienes y servicios que el programa entrega para resolver el problema o necesidad al cual está orientado. Evaluar los mecanismos de transferencia y modalidades de pago.
- Mecanismos de seguimiento, monitoreo y evaluación, considerando la disponibilidad de información necesaria y suficiente para la gestión.

### **III.3 Dimensiones o Foco del Desempeño**

- a. Eficacia: Se refiere al grado de cumplimiento de los objetivos planteados, es decir, en qué medida el área, el programa o la institución como un todo, está cumpliendo con sus objetivos en sus distintos niveles, sin considerar necesariamente los recursos asignados para ello.

### *Niveles de Resultados*

Muchas veces la relación causa-efecto entre los componentes del programa y los beneficios esperados puede ser distinta dependiendo del período que se considere; en este sentido, se rescata la posibilidad de que existan diferenciadamente resultados de corto, mediano y largo plazo. Por lo tanto, en el proceso de evaluación se debe incluir, si corresponde, los distintos tipos de resultados y establecer el momento en que deben ser medidos y analizados (horizonte temporal para los resultados). De esta manera, se pueden identificar, resultados a nivel de producto, nivel intermedio y final (impacto).

- A nivel de producto, (componentes) los resultados están relacionados con los bienes o servicios que el programa/institución provee directamente (nivel y procesos de producción).
- A nivel de propósito, los resultados están relacionados con los cambios en la situación de los beneficiarios como efecto directo de la entrega de los bienes o servicios producidos por el programa/institución. Esto puede verse como un resultado intermedio para el logro del fin.
- A nivel de resultado final para identificar el impacto o los beneficios “netos” de un programa sobre su población beneficiaria, se deben identificar las relaciones de causa-efecto entre los componentes que produce el programa y los resultados o variables de interés sobre las cuales están definidos los objetivos del programa. Se intenta responder la pregunta ¿Los bienes y servicios que entrega el programa contribuyen a lograr los efectos esperados en la población beneficiaria? Para esto, es necesario aislar de los beneficios observados todos aquellos efectos sobre la población beneficiaria derivados de factores externos al programa y que de todas maneras hubiesen ocurrido sin la ejecución de éste, como consecuencia de la evolución normal de las condiciones del entorno o la acción de otros programas públicos o privados.

b. Eficiencia: Describe la relación entre dos magnitudes: la producción física de un producto (bien o servicio) y los insumos o recursos que se utilizan para alcanzar ese nivel de producción (para más información de este último punto ver Anexo II).

Se utilizan indicadores que relacionen gastos con productos y/o resultados, analizando su evolución o variación porcentual a lo largo del período evaluado, comparando estos indicadores. Además, se pueden comparar estos indicadores con los de otros programas (que entregan bienes y/o servicios similares) del sector público o privado, con una meta o un compromiso de desempeño o con un referente internacional. Asimismo se evalúa lo siguiente:

- Si las acciones o intervenciones para generar bienes y servicios, se ejecutaron incurriendo en el menor costo posible.
- La productividad media de los recursos humanos y otros factores de producción relevantes que se utilicen para generar bienes y servicios.

- Actividades o procesos que se llevan a cabo en la actualidad y que podrían ser prescindibles o posibles de sustituir por mecanismos de menor costo o posibles de externalizar al sector privado.

Por último, se evalúan los gastos administrativos, siendo lo más usual comparar el porcentaje que éstos representan respecto del gasto total del programa, o sea cuantos recursos se destinan a los servicios de apoyo al programa por cada peso que gasta el programa. Se evalúa su evolución, composición, y se compara con otros programas o instituciones similares del sector público y/o privado.

c. Economía<sup>12</sup>: Este concepto se relaciona con la capacidad para generar y movilizar adecuadamente los recursos financieros en pos de los objetivos definidos por una institución y/o programa.

Se evalúa la capacidad del programa de ejecutar las actividades programadas, en función de los recursos contemplados en el presupuesto, comparando el gasto devengado y/o ejecutado del programa en relación al presupuesto inicial.

En los casos que corresponda, se evalúa la capacidad del programa para allegar recursos de terceros, sean estos aportes de otras instituciones (públicas o privadas) o de los beneficiarios. Un indicador relevante a estimar y evaluar en estos casos será el porcentaje de los aportes realizados por instituciones (públicas o privadas) o por beneficiarios en relación al total del gasto.

Por último, y en los casos que corresponda, se evalúa la política de recuperación de gastos y se presenta un análisis de las tarifas aplicadas y su relación con los gastos. Por otro lado, si el programa no contempla estos mecanismos, se analiza si existen las condiciones para una eventual recuperación de gastos por los servicios prestados.

d. Calidad: Es una dimensión específica del concepto de eficacia que se refiere a la capacidad para responder a las necesidades de sus clientes, usuarios o beneficiarios. Se refiere a atributos de los bienes o servicios entregados, tales como: oportunidad, accesibilidad, precisión y continuidad en la entrega del servicio, comodidad y cortesía en la atención.

Al evaluar estas dimensiones en su conjunto y en los distintos niveles de resultado, se debe emitir un juicio evaluativo global de desempeño, que permita concluir si el programa o institución, logra sus objetivos de manera eficiente, si se justifica la continuidad de las intervenciones y si existen espacios de mejora en cualquiera de los ámbitos evaluados.

---

<sup>12</sup> El Anexo II presenta en detalle los conceptos utilizados en materia de uso de recursos.

### **III.4 Sostenibilidad**

La sostenibilidad de un programa dice relación con todas aquellas condiciones que permiten que el programa logre continuidad operativa y cumplimiento de sus objetivos en el mediano y/o largo plazo. Estos plazos deben entenderse como el período de vida útil del programa y, por lo tanto, la duración en años de ese periodo dependerá del diseño y los objetivos de cada programa.

En cuanto a las condiciones que permiten que el programa logre continuidad y cumplimiento de objetivos, éstas pueden ser de diversa naturaleza. En general, se relacionan con (i) características organizacionales tales como, compromiso de las autoridades, acceso a los centros de decisión, espacios de coordinación inter-institucional, formalización de su estructura operativa; (ii) capacidades existentes en el programa tales como, capacidades profesionales, dotación de recursos, disponibilidad de financiamiento para un período relevante; y (iii) valoraciones de actores externos a la institución o programa, tales como legitimidad ante los beneficiarios y los ciudadanos y credibilidad por parte de actores políticos y sociales.

La evaluación de este aspecto debería entregar un juicio global acerca de la factibilidad de que el programa siga funcionando y cumpliendo de manera adecuada sus objetivos, dadas las condiciones actuales de operación, en términos de los factores antes mencionados.

### **III.5 Enfoque de Género**

La incorporación del enfoque de género en la evaluación de programas genera información a las autoridades para llegar de mejor manera a los ciudadanos, con las distintas políticas y programas (Anexo V).

En la evaluación de programas se incorporan las siguientes preguntas en términos de la pertinencia de la incorporación de la variable de género en los programas, para que estas sean respondidas en el marco de la evaluación:

¿Corresponde la incorporación del enfoque de género en el programa según la evaluación?, de ser así ¿se debió incorporar en la definición de la población objetivo?, si corresponde incorporarla ¿se incorporó satisfactoriamente?

En caso de corresponder la incorporación del enfoque de género, ¿éste fue incorporado en la definición del objetivo principal, y/o en los bienes o servicios entregados por el programa o en los procesos involucrados en su producción o provisión? ¿se incorporó indicadores que midan los efectos que ha tenido el programa en la situación de hombres y mujeres y en las relaciones que entre ellos se establecen y si existe algún grado de desigualdad entre hombres y mujeres en el acceso a los bienes y servicios que entrega el programa?

## **IV. Metodologías e Instrumentos**

La DIPRES en el marco del sistema de evaluación ex post ha desarrollado cuatro líneas de evaluación para las cuales se ha utilizado distintas metodologías e instrumentos.



La línea de Evaluación de Programas Gubernamentales (EPG) ha utilizado desde sus inicios la metodología de Marco Lógico para analizar e identificar oportunamente problemas de diseño y para identificar y luego cuantificar los principales indicadores que dan cuenta de su desempeño en los distintos niveles de objetivos, en particular a nivel de propósito y producto (componentes).

En cambio, la línea de Evaluación de Impacto (EI), dispone de distintas metodologías para estimar el impacto de una intervención, de manera que permita aislar la influencia de factores externos al programa que pudieran estar influyendo en sus resultados.

En la línea de Evaluación del Gasto Institucional (EGI) si bien no existe una metodología específica predeterminada, se aplica una metodología similar a la que se aplica en la línea EPG, pero adaptada a los requerimientos de una evaluación de carácter más amplio, centrada en la institución como un todo y no en un programa específico.

En la línea de Evaluación de Programas Nuevos (EPN) la metodología utilizada ha sido desarrollada ad hoc, con el objeto de identificar y fundamentar las eventuales brechas entre el diseño propuesto y el efectivo (implementación). Por otra parte, revisa el estado de avance del programa en cuanto a nivel de producción y resultados de corto plazo, así como cuantificar ciertos indicadores de gasto relevante (gasto por beneficiario, gasto por unidad de componente, etc.). De esta manera, una vez conocido en profundidad el diseño y la operación del programa, analiza la factibilidad y pertinencia de diseñar y eventualmente, realizar una evaluación de impacto.

A continuación se desarrollan con más detalle diferentes metodologías de uso frecuente en la evaluación de programas e instituciones.

#### **IV.1 Metodología de Marco Lógico**

La Metodología de marco Lógico (MML) es una herramienta utilizada en la línea de Evaluación de Programas Gubernamentales (EPG). Este instrumento facilita el proceso de conceptualización, diseño, ejecución y evaluación de proyectos o programas. Su énfasis está centrado en la orientación por objetivos, hacia grupos de beneficiarios y en facilitar la participación y comunicación entre las partes interesadas. (Ortegón, Pacheco y Prieto, 2005).

Presenta ventajas respecto a otros enfoques menos estructurados, entre los que se identifican:

- Aporte de una terminología uniforme que facilita la comunicación entre las partes interesadas y reduce ambigüedades.
- Aporte de un formato para consensuar los objetivos, metas y riesgos de un proyecto o programa.
- Suministro de información para organizar y preparar en forma lógica el plan de ejecución del programa.
- Suministro de información necesaria para la ejecución, monitoreo y evaluación.

Esta metodología es considerada una “ayuda para pensar” en el problema, agentes involucrados, jerarquía de objetivos y selección de una estrategia de implementación óptima. Su principal producto es la Matriz de Marco Lógico, que resume y refleja los elementos más importantes del programa (qué hace, cómo lo hace, para qué, los supuestos involucrados, y la forma en que el desempeño del programa será evaluado).

La metodología contempla dos etapas (Ortegón, Pacheco y Prieto, 2005):

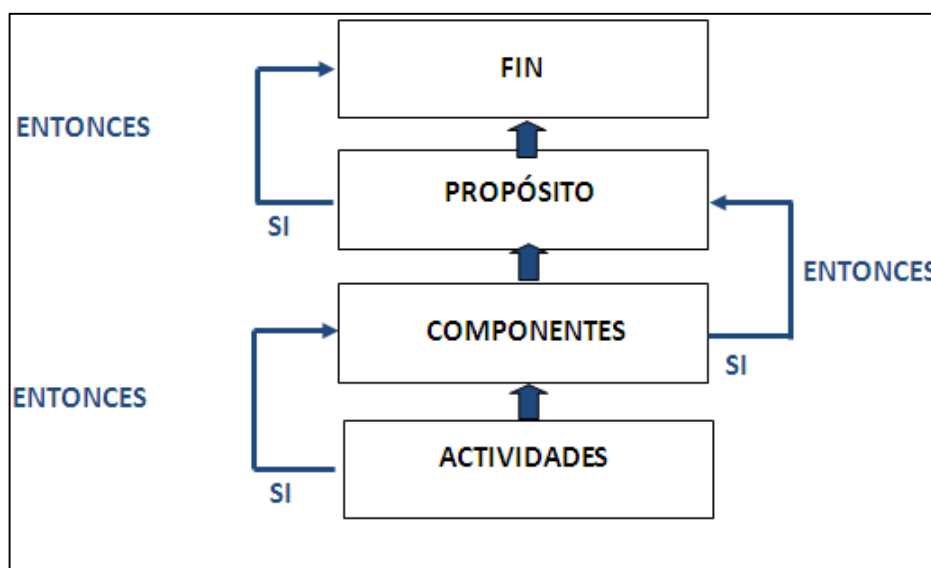
- a. Diagnóstico: Que consiste en la identificación del problema que abordará la intervención. Se debe tener presente que los programas son diseñados para resolver problemas o necesidades de grupos meta u objetivos particulares.
- b. Etapa de planificación, en la cual se hace operativa la estrategia definida a través de un plan de intervención. Se definen los objetivos, los servicios, las actividades y recursos necesarios a ser ejecutados en un determinado plazo o tiempo.

Para el desarrollo de la primera etapa se debe definir lo siguiente:

1. Identificar los agentes involucrados en la situación problema que el programa va a abordar. Lo relevante es tener en consideración a todos aquellos que tengan alguna relación con dicho problema. Pueden ser grupos determinados, entidades u organizaciones y las personas que pueden tener interés o verse beneficiados o perjudicados directa o indirectamente con el programa.
2. Identificar de manera participativa, con los involucrados, **el problema** a intervenir, sus causas y sus efectos. Esto se puede materializar en un “árbol de problemas” que es la imagen/representación de la situación negativa existente, que da cuenta de un problema central con sus causas y sus efectos.
3. Convertir los estados negativos del árbol de problemas en soluciones, lo que se traduce en un árbol de medios/fines (árbol de objetivos).
4. Identificar las diferentes alternativas de solución y seleccionar aquella alternativa que represente la mejor solución al problema detectado. En esta etapa, se pueden utilizar diferentes metodologías y criterios para seleccionar la mejor estrategia (análisis de costo-beneficio, costo-eficiencia, multicriterio, entre otros).

En la segunda etapa, y previo a representar el programa en una matriz de marco lógico, se debe construir la Estructura Analítica del Proyecto (EAP) o en otras palabras, desarrollar la etapa de planificación. Ello consiste en jerarquizar los objetivos del programa: fin, propósito (objetivo central del programa), identificar los componentes (bienes o servicios a proporcionar) y definir las actividades como un proceso que permite obtener los componentes. De esta manera, y tal como muestra la figura 2, se establece una relación vertical de causalidad, de abajo hacia arriba, donde el cumplimiento de las actividades permite obtener los componentes, los que, una vez generados, permiten lograr el propósito, y finalmente, contribuir al logro del fin.

**Figura 2: Estructura Analítica del Proyecto**



Fuente: Elaboración propia.

Aún cuando las iniciativas/programas no hayan sido diseñadas bajo la MML, siempre es posible su construcción. Lo importante es que la matriz refleje lo que el programa es en la actualidad.

Según se presenta en la Cuadro 2, el marco lógico es una matriz de columnas y filas donde las columnas representan los objetivos del programa, los indicadores mediante los cuales se van a medir dichos objetivos, los medios de verificación que dan cuenta de las fuentes de información para la medición de los indicadores y los supuestos que deben dar cuenta de las condiciones externas y necesarias para la implementación del programa. Por su parte, en las filas se identifican los distintos niveles de objetivos (fin, propósito, componentes, actividades).

**Cuadro 2: Estructura de Matriz de Marco Lógico**

Resumen Narrativo de los Objetivos	Indicadores	Medios de Verificación	Supuestos
Fin			
Propósito			
Componentes			
Actividades			

Fuente: Elaboración propia.

#### IV.1.1 Objetivos

- **Fin:** debe responder a la siguiente pregunta: *¿Por qué el programa es importante para los beneficiarios y la sociedad? Describe la solución a problemas de nivel superior e importancia nacional, sectorial o regional que se*

han diagnosticado. El fin representa un objetivo a nivel estratégico y describe el impacto al cual el programa contribuye.

El programa en sí mismo no es suficiente para lograr el fin, pero sí contribuye a su logro. Luego, diversos programas pueden contribuir a un mismo fin y, a su vez, un mismo programa puede contribuir a diversos fines.

- **Propósito:** debe responder a la siguiente pregunta: *¿Por qué el programa es necesario para los beneficiarios?* Describe el efecto directo o resultado esperado del programa sobre la población objetivo definida. Es la hipótesis sobre la cual se construye la intervención y a la cual la producción y uso de los componentes se encuentra dirigida. Se exige en la matriz la existencia de un solo propósito, para no producir ambigüedades en el análisis de efectividad del programa.
- **Componentes:** debe responder a la siguiente pregunta: *¿Qué debe ser producido por el programa para lograr su propósito?* Son los bienes y servicios que el programa entrega a la población beneficiaria y se definen como resultados, vale decir, como obras terminadas, estudios terminados, capacitación terminada, etc. Un programa puede presentar varios componentes, los que deben ser necesarios y complementarios entre sí para el cumplimiento del Propósito.
- **Actividades:** deben responder a la siguiente pregunta: *¿Qué insumos o tareas son llevadas a cabo para producir los bienes y/o servicios del programa?* Son aquellas que el ejecutor tiene que llevar a cabo para producir cada Componente (bien o servicio); deben dar cuenta de un proceso de producción del bien o servicio, en un periodo de tiempo y espacio definido, e implican la utilización de recursos (humanos, materiales, financieros, de infraestructura, entre otros).

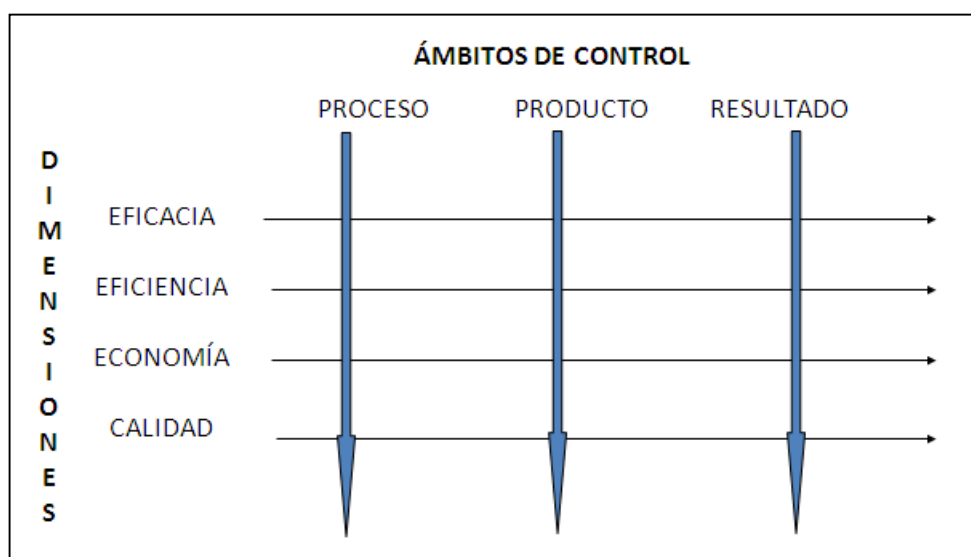
#### IV.1.2 Indicadores

Los indicadores de desempeño son una herramienta que entrega información cuantitativa respecto al logro o resultado en la entrega de los productos (bienes o servicios) generados por la institución, pudiendo cubrir aspectos cuantitativos o cualitativos de este logro. Es una expresión que establece una relación entre dos o más variables, la que comparada con períodos anteriores, productos similares o una meta o compromiso, permite evaluar desempeño a nivel de actividades, componentes, propósito y fin.

El número de indicadores no puede exceder la capacidad de análisis de quienes los van a usar. Por esto es necesario diferenciar entre indicadores de gestión o de monitoreo interno de los programas o instituciones (a nivel de los diferentes centros de responsabilidad o unidades de trabajo), de aquellos necesarios para los niveles superiores de toma de decisiones o con fines de *accountability*. En este último caso los indicadores deben cubrir los aspectos más significativos de la gestión, privilegiando los principales objetivos del programa o de la organización.

Los indicadores de desempeño se construyen en función del ámbito de control y de la dimensión que miden (Figura 3).

**Figura 3: Diagrama de Indicadores de Desempeño**



Fuente: Elaboración propia.

Los ámbitos de control tienen relación con el momento en que se realiza la medición:

- **Procesos:** está relacionado con las actividades que se realizan, con la ejecución o forma en que el trabajo es realizado para producir los productos (bienes y/o servicios); incluyen actividades o prácticas de trabajo tales como procedimientos de compra, procesos tecnológicos y de administración financiera.
- **Producto:** está relacionado con los resultados del programa a nivel de componentes (bienes y/o servicios). En una línea de tiempo corresponden al primer tipo de resultados que debe generar un programa y al efecto más inmediato que tiene la producción de los componentes sobre la población beneficiaria.
- **Resultados:** se identifican dos tipos de resultados, intermedios y finales (impacto). Los primeros corresponden, principalmente, a nivel de propósito. Se refieren a cambios en el comportamiento, actitud o estado de los beneficiarios, una vez que han recibido los bienes y servicios que entrega el programa. Los segundos implican un mejoramiento significativo y perdurable en el tiempo de alguna de las condiciones o características de la población objetivo que se plantearon como esenciales en la definición del problema. Un resultado final suele expresarse como un beneficio de mediano y largo plazo obtenido por la población atendida.

Las dimensiones o focos de desempeño tal como se describe en el capítulo 3 son las siguientes: Eficacia, Eficiencia, Economía y Calidad.

### IV.1.3 Medios de Verificación

Corresponde a las fuentes de información que se utilizarán para la construcción y medición de los indicadores planteados. Pueden ser fuentes primarias, producidas por el propio programa; o fuentes secundarias independientes del programa, como por ejemplo, estadísticas del INE, estadísticas sectoriales, censo de población y vivienda, encuesta CASEN, entre otros.

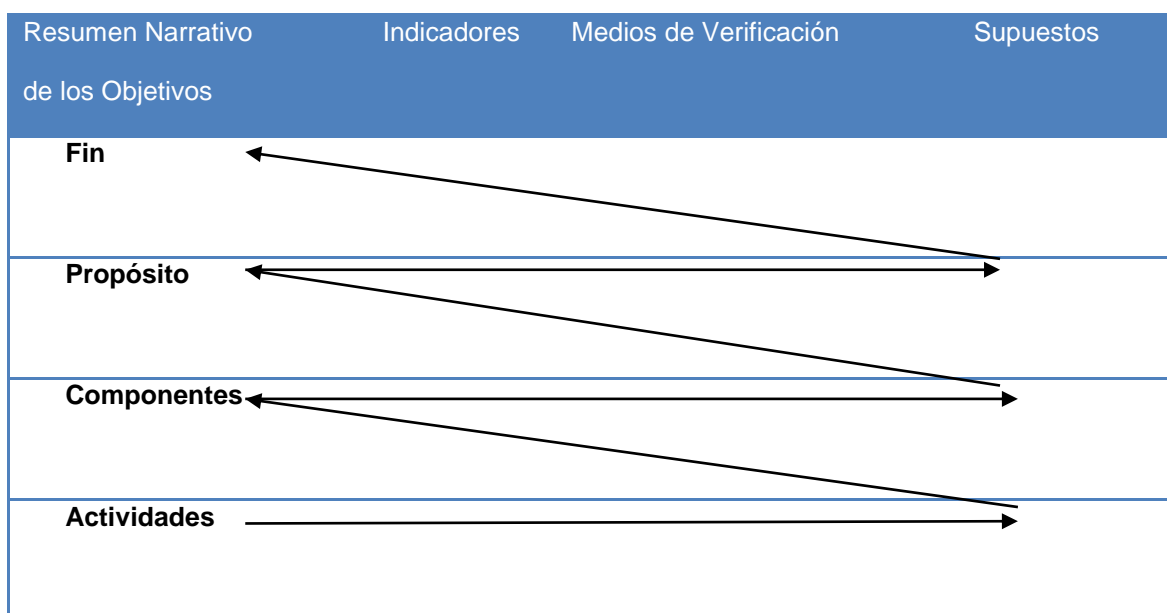
### IV.1.4 Supuestos

Los supuestos son condiciones externas, fuera del control de la administración del programa, que son necesarias para el éxito de la intervención; si estos supuestos no se cumplen se transforman en riesgos. Su identificación y monitoreo juega un rol importante, tanto en la planificación como en la ejecución del programa, dado que el programa debe enfrentar la situación en caso que un supuesto no se cumpla.

### IV.1.5 Lógica Vertical y Horizontal

La Figura 4 muestra la relación entre los distintos elementos que constituyen el programa y las relaciones que deben existir entre ellos, dando cuenta de una lógica vertical y horizontal de la matriz.

**Figura 4: Relación entre Supuestos y Objetivos en la Matriz de Marco Lógico**



Fuente: Elaboración propia.

La lógica vertical corresponde al análisis de jerarquía de los objetivos planteados y de los vínculos causales, de abajo hacia arriba, entre los niveles de objetivos. El análisis se inicia en la parte inferior de la matriz, a nivel de actividades, considerando que el

cumplimiento de ellas, más el cumplimiento de los supuestos identificados a ese nivel, permiten obtener los componentes del programa. Es decir, las actividades, junto con los supuestos a este nivel, describen las condiciones necesarias y suficientes para producir los componentes.

Si se producen los componentes y se cumplen los supuestos de este nivel, se logrará, entonces, el propósito del programa. Es decir, los componentes, junto con los supuestos a este nivel, describen las condiciones necesarias y suficientes para lograr el propósito.

Si se logra el propósito, y se confirman los supuestos de este nivel, se habrá contribuido de manera significativa a alcanzar el fin. El propósito, junto con los supuestos a este nivel, describen condiciones necesarias, aún cuando no suficientes, para lograr el fin.

Por su parte, la lógica horizontal del programa analiza, para cada nivel de objetivo, si los indicadores diseñados permiten hacer un buen seguimiento del programa y evaluar adecuadamente su logro; y si los medios de verificación son los necesarios y suficientes para obtener los datos requeridos para la medición de los indicadores.

## **IV.2 Metodologías de Evaluación de Impacto**

La evaluación de los resultados de un programa público requiere desarrollar un proceso de trabajo sistemático orientado a responder preguntas tales como: ¿Se ha contribuido significativamente a resolver el problema que dio origen al programa? ¿Los receptores de los bienes y servicios que genera el programa han resultado realmente beneficiados? ¿Los beneficios recibidos por los participantes son los que se había propuesto lograr el programa o son otros no previstos? ¿Los beneficios están llegando a la población que se pretende atender? ¿Los usuarios del programa están satisfechos con los bienes y servicios recibidos? ¿Los beneficiarios han experimentado un mejoramiento significativo en su condición-problema inicial, como consecuencia de su participación en el programa? De las preguntas anteriores se desprende que la evaluación de impacto comprende la “efectividad” de las medidas tomadas o los servicios entregados por el programa.

Para determinar la efectividad de una intervención, en primer lugar se requiere identificar los efectos esperados de mediano y largo plazo sobre la población beneficiaria que espera obtener el programa. ¿Cuáles son los beneficios con que quedará la población beneficiaria una vez que concluya su participación en el programa?, (el impacto) por ejemplo: incremento en el nivel de remuneraciones de trabajadores capacitados, reducción de la tasa de morbilidad de menores beneficiados con un programa de vacunación, aumento del rendimiento escolar de los alumnos de escuelas participantes en un programa de reforzamiento educativo, etc.

Lo anterior implica identificar y separar los beneficios logrados como consecuencia de la intervención del programa, de aquellos que de todas maneras hubieran ocurrido sin la ejecución de éste, por efecto de la evolución normal de las condiciones del entorno o la acción de otros programas públicos o privados. Para ello, corresponderá definir la(s) variable(s) que deben ser medidas para determinar sí el programa ha generado

realmente los beneficios o resultados finales o impacto esperados.

Para identificar la existencia, estadística y robusta, de una relación causal, y por lo tanto el impacto, entre la intervención pública y las variables de resultado que se han definido, existen metodologías específicas, las que serán descritas de forma simple y concisa en este documento. No obstante, es una literatura que se encuentra en permanente desarrollo, tanto modificando y perfeccionando estrategias actuales como desarrollando nuevas. Es por lo anterior que el contenido de esta sección debe ser interpretado como una guía o simple referencia de las metodologías existentes en este ámbito<sup>13</sup>.

#### IV.2.1 Estimación del Impacto de una Intervención

En la práctica quisiéramos observar al mismo tiempo al individuo afectado por la política o intervención, que en esta sección llamaremos tratamiento, y no afectado a la intervención para identificar el impacto de la política en él.

Sea  $Y_i^T$  sea el resultado del individuo "i" expuesto al tratamiento T, y  $Y_i^{NT}$  el resultado del mismo individuo sin el tratamiento. Luego, el impacto de la intervención es la simple diferencia entre los resultados:  $Y_i^T - Y_i^{NT}$ .

No obstante, lo anterior es sólo un ejercicio teórico ya que es imposible observar a la misma persona con y sin tratamiento en el mismo momento del tiempo. Luego, las metodologías diseñadas para la identificación de impacto buscan solucionar este problema de factibilidad práctica, pero los resultados de las mismas varían según la naturaleza y metodología utilizada<sup>14</sup>.

Los tipos de impactos posibles de encontrar por medio de las diversas metodologías son el efecto promedio del tratamiento o con su nombre en inglés, *Average Treatment Effect* (ATE), el efecto promedio sobre los tratados o *Average Treatment on the Treated* (ATT), efecto promedio local sobre los tratados o *Local Average Treatment Effect* (LATE) y efecto marginal sobre los tratados o *Marginal Treatment Effect* (MTE)<sup>15</sup>.

El ATE, corresponde al efecto promedio o esperado en la población objetivo de la intervención y es el estimador más cercano del efecto tratamiento. Tal como refleja (1), el ATE está compuesto por el ATT y una medida de sesgo. Este último, es un reflejo de las diferencias intrínsecas que inciden en la selección de los individuos que participan en un programa y se espera que este sesgo sea en general distinto de cero, con lo cual el impacto de la intervención se encontrará sobre o subestimado.

Es por lo anterior, que la literatura tanto teórica como empírica ha buscado y diseñado metodologías que permiten la creación de un grupo control o sin tratamiento o contrafactual, que debiesen ser en promedio igual al grupo de personas tratadas,

---

<sup>13</sup> La presente sección fue construida en base a Wooldridge (2002), Angrist y Pischke (2008).

<sup>14</sup> La elección de una (o más) metodologías deberá siempre fundarse, principalmente en la forma de la implementación del tratamiento como de la información cuantitativa que se puede utilizar para ello. Lo anterior, permite reducir ambigüedades y cuestionamiento de los resultados encontrados.

<sup>15</sup> Se hará mención a los mismos con sus respectivos nombres en inglés en vista de que no existe una traducción adecuada al español.



aislando a su vez otros factores que pueden afectar o inducir el resultado que se está analizando, para así eliminar dicho sesgo y encontrar el ATT.

$$ATE = E(Y_i^T | T_i = 1) - E(Y_i^{NT} | T_i = 0) \quad (1)$$

$$ATE = \underbrace{E(Y_i^T - Y_i^{NT} | T_i = 1)}_{ATT} + \underbrace{E(Y_i^{NT} | T_i = 1) - E(Y_i^{NT} | T_i = 0)}_{Sesgo} \quad (2)$$

Las metodologías que permiten en teoría la obtención de un ATE=ATT son<sup>16</sup>: i) Experimento Aleatorio; ii) Mínimos Cuadrados Ordinarios; iii) *Matching*; iv) Diferencias en Diferencias; v) Datos de Panel y; vi) Diseño Regresión Discontinua. Por su parte es posible la obtención de un LATE por medio de Variables instrumentales y Diseño Regresión Discontinua. Por último, la obtención de efectos heterogéneos o MTE se pueden obtener a través de *Quantil Regression* y Estrategia de Efectos Marginales.

### **Experimento Aleatorio**

Tal como se ha señalado, el problema fundamental de la estimación del efecto de una intervención es el sesgo de selección que “contamina” la estimación del impacto real de una determinada intervención. Una forma de remover este sesgo es por medio de la realización de un experimento aleatorio, el que corresponde a una implementación aleatoria del tratamiento o selección al azar de los individuos a tratar.

La aleatorización de la intervención implica que un individuo que recibe tratamiento es, en promedio, igual a alguien que no lo recibe. Luego, por construcción se cumple que el sesgo de selección es cero (3):

$$E(Y_i^{NT} | T_i = 1) = E(Y_i^{NT} | T_i = 0) \quad (3)$$

Este tipo de metodología presenta la ventaja de una fuerte validez interna<sup>17</sup> y luego la estimación del impacto del programa es efectivamente un ATE=ATT. No obstante, existen algunas amenazas a la misma, tales como *attrition*<sup>18</sup> (en términos simples pérdida de la muestra a seguir), *Imperfect Compliance*<sup>19</sup> (contaminación en los

<sup>16</sup> Se indica en teoría, ya que el resultado dependerá del cumplimiento de los supuestos que avalan una determinada metodología.

<sup>17</sup> Éste término se refiere a la capacidad de una estrategia empírica de resolver el problema del sesgo de selección y por ende de encontrar el efecto del tratamiento.

<sup>18</sup> Ocurre cuando individuos en la muestra (control o tratamiento) no pueden ser encontrados después que el programa se implemento. Esto es un problema en la medida que el *attrition* no es aleatorio.

<sup>19</sup> Ocurre cuando hay individuos en el grupo de tratamiento que no reciben el tratamiento o cuando individuos en el grupo de control, reciben el tratamiento. En este escenario la estimación del efecto del programa, la simple diferencia en los promedios de la variable de resultado del grupo tratado con el de control, ya no es un ATE sino que un *intention to treat* (ITT), que corresponde al tratamiento promedio,

respectivos grupos de tratamiento y control), y no cumplimiento de un supuesto básico llamado *Standard Unit Treatment Value* o SUTVA<sup>20</sup>, que deben ser consideradas por los evaluadores del programa.

Otro aspecto a considerar en la aplicación de esta metodología es la determinación del tamaño muestral del experimento. Se requiere de un tamaño que permita detectar el efecto de la intervención, dicho tamaño se llama *Minimum Detectable Effect* (MDE) o, en su traducción al español mínimo tamaño detectable, y asegurar la confiabilidad del mismo (poder y significancia)<sup>21</sup>.

En la práctica, el diseño experimental es una técnica con baja viabilidad política, en base a sus razones éticas (selección aleatoria de individuos), muy intensiva en tiempo y costosa. Por todo lo anterior, en general, su aplicación se da en programas de carácter piloto o en aquellos en los cuales su implementación, en términos de cobertura, es gradual en el tiempo.

### **Mínimos Cuadrados Ordinarios (MCO)**

MCO es un método de estimación de los parámetros poblacionales de una regresión lineal. Tal como explica Greene (2003) la regresión lineal es una de las herramientas más utilizadas por la econometría ya que es una forma simple de representar (linealmente) la relación entre variables dependientes con independientes (4). Por su parte, la metodología de MCO lo que trata es estimar los parámetros poblacionales y de interés de dicha regresión, a través de la minimización de la esperanza de sus errores al cuadrado.

$$Y_i = \alpha + \beta T_i + u_i \quad (4),$$

donde Y es la variable dependiente, T variable explicativa,  $\alpha$  y  $\beta$  parámetros de interés, u una media de error.

Con este proceso de minimización y si se cumplen los supuestos básicos de MCO<sup>22</sup> es posible obtener estimadores de  $\alpha$  y  $\beta$ , insesgados, consistentes y eficientes (de mínima varianza).

En la representación en (4) la variable explicativa es T, correspondiente al tratamiento o intervención y por lo tanto el interés radica en la estimación de  $\beta$ . Luego, para poder efectivamente identificar  $\beta$  se deben cumplir los siguientes supuestos, los que en

---

comparando los grupos originalmente asignados, sin importar si recibieron o no el tratamiento. Sin embargo, con el ITT es posible recuperar el ATE.

<sup>20</sup> Corresponde a un supuesto de independencia entre los resultados potenciales de cualquier unidad de tratamiento o control. En otras palabras, no hay externalidades entre los agentes. En caso de temer el cumplimiento de SUTVA, una alternativa de solución, es modificar la escala del grupo de tratados y control.

<sup>21</sup> Para mayor información del cálculo de la muestra y de la implementación de un experimento aleatorio ver Duflo, Glennerster y Kremer (2007).

<sup>22</sup> Homocedasticidad, ausencia de autocorrelación y el error tiene una distribución normal, entre otras. Para mayor información ver Greene (2003).

términos sencillos nos quieren decir que la asignación del tratamiento debe ser tan buena como si se hubiese realizado un experimento aleatorio:

Ignorabilidad Fuerte<sup>23</sup>: Condicionado en el set de características observables del individuo ( $X_i$ ) la probabilidad de recibir el tratamiento no se encuentra correlacionado con los resultados potenciales con o sin tratamiento. En otras palabras existe selección por observables.

$$\{Y_i^{NT}, Y_i^T\} \perp T_i | X_i \quad (5)$$

Ignorabilidad Débil: El contrafactual esperado, condicional en características, es el mismo independiente si el individuo recibió o no el tratamiento.

$$E(Y_i^{NT} | T_i, X_i) = E(Y_i^{NT} | X_i) \quad (6)$$

Como es posible inferir, la estimación por MCO es posible siempre que la selección de los individuos (generación del grupo de tratamiento y control) se deba a las diferencias existentes en sus características observables y que la especificación (relación de causalidad) sea lineal. Aún cuando MCO no sea la mejor estrategia de identificación, dado por sus supuestos altamente restrictivos, siempre se considera como punto de comparación ya que permite evaluar la magnitud del sesgo.

### **Matching**

Uno de los objetivos centrales de esta metodología es crear un grupo de control o contrafactual que presente covarianzas similares en sus características observables a la muestra de tratados. Luego, se asume que no existen diferencias en las características no observables. Por su parte, Rubin (1976) establece el concepto de *equal percent bias reducing* (EPBR), que permite verificar si efectivamente la metodología seleccionada establece la “igualdad” entre los individuos de ambos grupos.

La metodología de *matching* presenta varias alternativas de aplicación, siendo una de las más antiguas, el *Matching* Exacto. En ésta se busca un par que sea idéntico en todas las características observables al individuo tratado. Luego, es de esperar que surja el problema de reducción de la muestra a análisis por causa de celdas vacías o, en otras palabras, “la maldición de la dimensionalidad”<sup>24</sup>.

El *matching* exacto se basa en los mismos supuestos que MCO. Sin embargo, al ser

---

<sup>23</sup> La literatura se refiere a este supuesto indistintamente como *unconfoundedness*, selección en observables o *conditional independence assumption* (CIA).

<sup>24</sup> Celdas vacías hace referencia a la imposibilidad de encontrar un match que cumpla con todos los atributos observables definidos.

una técnica no paramétrica reduce los problemas de especificación del *outcome* o resultado del grupo de control y permite cierta heterogeneidad en el efecto del tratamiento. No obstante, es posible que MCO sea más eficiente que *Matching* en el escenario de una correcta especificación del modelo de causalidad (no existen variables relevantes omitidas) y sea posible “romper” el supuesto de linealidad (forma funcional) incorporando interacciones entre las variables.

Cabe señalar que el impacto estimado por medio del *matching* corresponde al igual que MCO un promedio ponderado y, por lo tanto, en caso de cumplirse todos los supuestos de rigor se estaría estimando un  $ATE=ATT$ . Sin embargo, la metodología de *matching* le otorga un mayor peso a aquellas celdas que tienen un mayor número de tratados mientras que MCO le otorga una mayor ponderación a aquellas celdas que tienen igual número de tratados y de controles.

Considerando el problema de dimensionalidad del *matching* exacto han surgido dos alternativas o líneas de acción, que se describen a continuación, que se hacen cargo de dicho problema. No obstante, la literatura ha avanzado en el perfeccionamiento de la metodología de *matching* modificando éstas alternativas y realizando combinaciones entre las mismas.

*Covariate Matching (CVM)*: Esta metodología elimina la diferencia entre los grupos y por lo tanto, siguiendo el principio de la metodología de *matching*, el sesgo en la estimación por causa de características observables asumiendo que no existen diferencias en atributos no observables. La idea central es establecer una medida de distancia entre las covarianzas de los individuos ( $X$ , vector de características observables del individuo  $i$ ), de forma tal de encontrar el *match* más adecuado<sup>25</sup>. Tal como indica Zao (2004), la idea básica de esta metodología es:

$$X_i = X_j \Rightarrow f_t(X_i) = f_t(X_j), \quad t = 0,1 \quad (7)$$

$$d(X_i, X_j) < \varepsilon \Rightarrow d'(f_t(X_i), f_t(X_j)) < \delta, t = 0,1 \quad (8)$$

donde  $d$  y  $d'$  son métricas.

La ecuación (7) justifica la existencia de un *match* exacto y (8) nos indica que  $f_t(*)$  es continua en  $X$ , lo que permite generar una medida de distancia entre los atributos permitiendo subsanar el problema de dimensionalidad.

Una de las métricas más utilizadas en la de Mahanobis, siendo la medida de distancia utilizada la señalada en (9)<sup>26</sup>. Además, existe una métrica similar a ésta, desarrollada por Abadie e Imbens (2002), donde lo que cambia es ponderador entre las observaciones.

<sup>25</sup> Una covarianza es un valor que indica el grado de variación conjunta de dos variables aleatorias.

<sup>26</sup> Mayor detalle de ésta metodología de *matching* en Carpenter (1977) y Cochran and Rubin (1973).

$$d_M = (X_i - X_j)D^{-1}(X_i - X_j)' \quad (9)$$

siendo D, la matriz de varianzas y covarianzas de X.

Zhao (2004) muestra que la métrica de Mahanobis es robusta frente distintas especificaciones y funciona relativamente bien en muestras pequeñas ( $N < 500$ ). Además, Rubin (1980) destaca que es una metodología que cumple con EPBR (por construcción) y es una extensión natural del caso univariable del vecino más cercano. No obstante, en la medida que aumenta el número de covarianzas, la viabilidad de encontrar un “par” cercano disminuye (Rubin y Thomas, 2000).

*Matching Propensity Score (MPS)*<sup>27</sup>: Otra forma de romper el problema de dimensionalidad del *matching* exacto es la desarrollada por Rosenbaum y Rubin (1983), quienes plantean como alternativa utilizar una variable escalar que resume la información del individuo, la que a su vez determina la posibilidad de ser tratado. Esta variable escalar es el Propensity Score (P-Score) (10) que es la probabilidad de recibir el tratamiento condicional a los atributos observables del individuo.

$$P(T = 1|X_i) = p(X_i) \quad (10)$$

Cumplíéndose el supuesto de Ignorabilidad Fuerte<sup>28</sup>, los supuestos que validan la metodología de *matching propensity score* son las siguientes:

Ignorabilidad Fuerte:

$$\{Y_i^{NT}, Y_i^T\} \perp T_i | X_i \Rightarrow \{Y_i^{NT}, Y_i^T\} \perp T_i | p(X_i) \quad (11)$$

Soporte Común: Permite comparar individuos que tengan una probabilidad (y no certeza) de ser tratado.

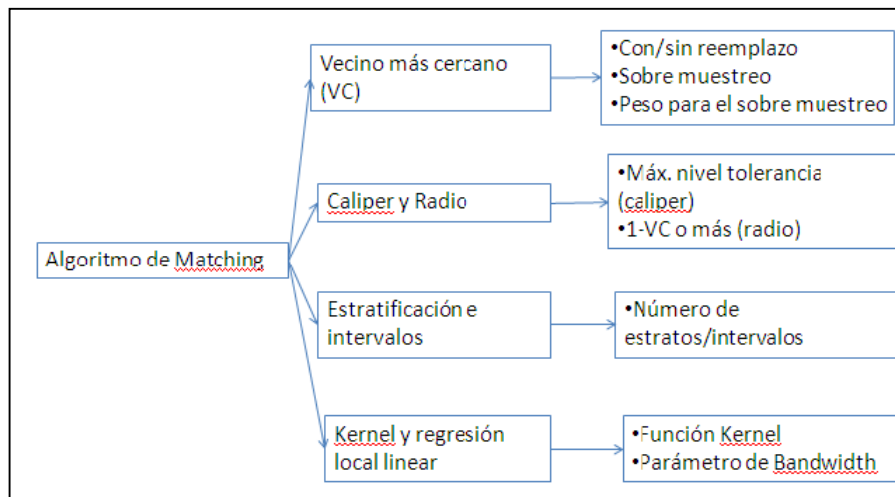
$$0 < P(T_i = 1|X_i) < 1 \quad (12)$$

Por su parte, existen varios algoritmos de *matching*, los que sólo serán señalados en este documento (Figura 5), y lo que se busca con ellos es lograr un *match* entre individuos con similar o cercano *p-score*, para posteriormente calcular la diferencia de *outcome* entre los pares seleccionados y así estimar el efecto promedio del tratamiento en la población objetivo.

<sup>27</sup> Se sugiere ver Caliendo y Kopeinig (2008), los que desarrollan un guía práctica, con gran detalle de la metodología, para la implementación de *matching propensity score*.

<sup>28</sup> Este concepto fue definido en la sección referente a MCO.

**Figura 5: Algoritmos de Matching**



Fuente: Traducido de Caliendo y Kopeinig (2008).

Se debe considerar que cada elección tanto del algoritmo de *matching* como de sus formas de implementación, tiene sus ventajas y desventajas (Cuadro 3), incidiendo en el sesgo y eficiencia de la estimación. Es por ello que las decisiones deben encontrarse siempre fundadas y dependerán del diseño/implementación del programa y de la naturaleza de los datos con los que se cuenta para la realización de la evaluación.

**Cuadro 3: Disyuntiva entre sesgo y eficiencia**

Decisión	Sesgo	Varianza
<b>Vecino más cercano:</b>		
Múltiples vecinos/un vecino	(+)/(-)	(-)/(+)
Con caliper/sin caliper	(-)/(+)	(+)/(-)
<b>Control de individuos:</b>		
Con reemplazo/sin reemplazo	(-)/(+)	(+/-)
<b>Elección del método:</b>		
Vecino más cercano/Radio	(-)/(+)	(+/-)
Kernel o regresión local lineal/ Vecino más cercano	(+)/(-)	(-)/(+)
<b>Elección de Bandwith:</b>		
chico/grande	(-)/(+)	(+/-)
<b>Orden del polinomio:</b>		
chico/grande	(+)/(-)	(-)/(+)

Fuente: Caliendo y Kopeinig (2008).

### **Variables Instrumentales (VI)**

Las metodologías descritas anteriormente asumen que el sesgo en la estimación se debe a las diferencias observables entre los individuos y no considerara la posibilidad de diferencias en atributos no observables, como por ejemplo la motivación de los agentes para postular o ser elegidos en un determinado programa.

Luego, aún cuando se controle por los atributos observables persistirá el sesgo en las estimaciones, lo que regularmente se explica por el sesgo de selección de los agentes.

Para controlar por éste sesgo, frente a intervenciones que ya sucedieron y que por construcción no consideraron una selección de los agentes vía un experimento aleatorio surgió como alternativa metodológica: la utilización de variables instrumentales, las que tienen por finalidad hacer que el tratamiento sea exógeno para los individuos.

Para representar mejor el problema del sesgo por autoselección y la no identificación del efecto del tratamiento de un programa, consideremos la siguiente regresión lineal a estimar por MCO:

$$y = x\beta + \mu \quad (13)$$

donde  $y$  es escalar y  $x$  representa una observación de dimensión  $1 \times K$ ,  $\beta$  es un vector de parámetros de  $K \times 1$  y  $\mu$  es un término de error no observable.

Solo sí se cumple el supuesto de identificación,  $E(x'\mu) = 0$ ,  $\beta$  (parámetro poblacional) se encontrará identificado en sus momentos poblacionales (14) y luego por el principio de la analogía será posible de representar lo anterior en sus momentos muestrales, tal como se representa en (15).

$$\beta = (x'x)^{-1}E(x'y) \quad (14)$$

$$\hat{\beta} = \left[\frac{1}{N} \sum x_i'x_i\right]^{-1} \left[\frac{1}{N} \sum x_i'y_i\right] \quad (15)$$

Frente al no cumplimiento del supuesto de identificación, producto del sesgo de selección o también llamado endogeneidad, no es posible encontrar una expresión del parámetro poblacional en función de momentos poblacionales de las variables observables del modelo y luego no es factible estimar el efecto (en éste caso) del tratamiento de un programa o intervención.

La metodología de variables instrumentales soluciona el problema anterior utilizando un set de variables ( $Z$ ) como instrumentos de la variable explicativa que produce la endogeneidad en el modelo. Estas variables deben cumplir con las siguientes propiedades: i) se encuentran correlacionadas con la variable endógena pero no con la dependiente, ii) generan un efecto constante en la población y iii) se cumple monotonicidad del tratamiento en función del instrumento. En términos formales, los supuestos antes expuestos se traduce en las siguientes dos expresiones:

$E(z'\mu) = 0$ ; El error no se encuentra correlacionado con el instrumento.

$E(z'x) \neq 0$ ; El instrumento esta correlacionado con la variable explicativa de interés.

Ahora bien, tal como postulan y explican Imbens y Angrist (1994) el coeficiente encontrado por medio de la metodología de VI ya no corresponderá al efecto tratamiento promedio (ATE) sino a un efecto promedio local (LATE). El cual a diferencia del ATE, corresponde al efecto promedio sobre aquellos individuos cuyo tratamiento está influenciado por los instrumentos seleccionados. Formalmente, lo anterior es posible de evidenciarlo en las ecuaciones (16) y (17), siendo la primera la identificación poblacional del parámetro y la segunda la estimación muestral del mismo<sup>29</sup>.

$$\beta = (z'x)^{-1}E(z'y) \quad (16)$$

$$\hat{\beta} = \left[\frac{1}{N} \sum z_i'x_i\right]^{-1} \left[\frac{1}{N} \sum z_i'y_i\right] \quad (17)$$

El proceso de estimación se visualiza en dos etapas, aún cuando la estimación se realiza en una sola por medio de un proceso de Máxima Verosimilitud<sup>30</sup>. La primera etapa consiste en la estimación de variable endógena en función de los instrumentos y otras variables exógenas (18) y la segunda sería la estimación de la ecuación principal siendo incluida la estimación de la variable endógena (19).

$$T = Z\pi + x\gamma + \varepsilon \quad (18)$$

$$y = \hat{T}\beta + x\delta + \mu \quad (19)$$

y es la variable resultado, T la endógena, x las exógenas,  $\varepsilon$  y  $\mu$  errores aleatorios, y el resto de los parámetros coeficientes a estimar.

En el caso de considerar la aplicación de esta metodología de estimación se debe justificar tanto la utilización de la misma como de los instrumentos seleccionados, enfatizando en el cumplimiento de los supuestos necesarios para la identificación.

Existen algunos métodos formales para sustentar el análisis por medio de VI. Para evaluar el cumplimiento del supuesto 1 se puede utilizar el test de sobre identificación o test de Hausman, pero se requiere contar con más de un instrumento por variable endógena. Para el supuesto 2, se puede utilizar el estadístico de Cragg y Donald (1993) junto con las tablas de Stock y Yogo (2005). No obstante, este último test (diagnóstico de instrumentos débiles) requiere que la variable endógena sea continua.

<sup>29</sup> Para mayor detalle de la metodología ver por ejemplo: Imbens y Angrist (1994), Angrist y Pischke (2009) o Heckman y Navarro-Lozano (2004).

<sup>30</sup> Se obtiene mayor eficiencia en los estimadores por medio de Máxima Verosimilitud (MV) respecto de una estimación en etapas. A su vez, se requiere la estimación por medio de MV frente a variables endógenas discretas ya que permite la estimación de la primera etapa por medio de un modelo de elección discreto manteniendo a su vez las propiedades de consistencia y eficiencia de las estimaciones.



Luego, en caso de estar en presencia de variables endógenas dicotómicas (por ejemplo el caso ser tratado o no) se requiere de una robusta justificación teórica de los instrumentos.

### **Diferencias en Diferencias (DD)**

El presente método busca eliminar tanto la incidencia de factores no observables del individuo que afectan la selección del mismo para ser beneficiario de un programa o tratamiento como a la vez controlar por aquellos factores exógenos al individuo (que afectan al grupo de tratados y controles), por ejemplo factores macroeconómicos u otras políticas actuando de forma simultánea, que no son posibles de incluir en la estimación y que pueden estar influyendo o sesgando el resultado.

La esencia de esta forma de estimación radica en la “explotación” de variaciones de la política en el tiempo y espacio, asimilándolo a un experimento natural. No obstante, para su aplicación se requiere de un amplio conjunto de información de los individuos (tratados y controles) antes y después de la ejecución o inicio de la política, programa o regulación, tal como se resume en el Cuadro 4.

**Cuadro 4: Información Básica para la Estimación de DD**

Tiempo	Grupo afectado por la política (tratados)	Grupo no afectado por la política (control)
<b>Antes del programa</b>	$Y_{i0} T = 1$	$Y_{i0} T = 0$
<b>Después del programa</b>	$Y_{i1} T = 1$	$Y_{i1} T = 0$

Fuente: Elaboración propia.

La metodología DD se sustenta en el cambio de las variables de resultado y, en especial, en los de tendencia del grupo de tratamiento a partir del momento (tiempo) de la intervención, tal como se expone en la Figura 6. Luego un supuesto clave es que el grupo de control mantiene su tendencia pre-tratamiento.



- Usar otros grupos de control
- Verificar que los grupos de control y tratamiento se encuentren balanceados
- Experimentar con otra variable de resultado que no debiese haber cambiado por el tratamiento.

Existen dos extensiones de la metodología clásica de DD. La primera, Diferencias en Diferencias (DDD), que consiste en una modificación natural de DD, donde en el caso de no contar con un grupo de control de “calidad” se procede (en caso de que exista) a incorporar un segundo grupo de control. Por ejemplo, supongamos que se establece una política para un grupo específico de la población en el área 1 (A1) y el área 2 (A2) no es beneficiaria de la política. Luego el método nos indicaría considerar como grupos de control a aquellos con similares características a los tratados pertenecientes al área 1 y a su vez a los posibles tratados del área 2 con su respectivo grupo de control<sup>32</sup>. Así este método permite controlar por factores propios de cada área o zona afectada por la política. Por su parte el estimador DDD es el que se presenta a continuación:

$$DDD = [[(\bar{Y}_1|T = 1) - (\bar{Y}_0|T = 1)] - [(\bar{Y}_1|T = 0) - (\bar{Y}_0|T = 0)]]^{A1} - [[(\bar{Y}_1|T = 1) - (\bar{Y}_0|T = 1)] - [(\bar{Y}_1|T = 0) - (\bar{Y}_0|T = 0)]]^{A2} \quad (22)$$

La segunda, es la metodología desarrollada por Athey e Imbens (2006), los que plantean un modelo general y no paramétrico de estimación, obteniéndose de él el estimador *Changes-in-Changes* (CIC). Para la aplicación de esta última metodología se requiere la determinación de las funciones de distribución acumulada de las variables de resultado del grupo de tratados y de control.

### **Datos de Panel**

Otra forma plausible de eliminar el sesgo de selección de los agentes por causa de variables no observables al investigador es por medio de la utilización de datos de panel o longitudinales. Éstas se caracterizan por seguir a los individuos en el tiempo, proporcionando varias observaciones por individuo en diversos momentos del tiempo.

Una de las ventajas de utilizar datos de panel es que permite controlar tanto por variables observables como no observables invariantes en el tiempo pero heterogéneas entre individuos o aleatorios pero constantes entre los agentes (efectos temporales).

Sea el modelo a estimar, bajo este contexto, (23) donde  $\mu_{it}$  corresponde al error. A su vez, éste error puede ser descompuesto entre aquellas variables que no son observables ( $z$ ) y el error aleatorio ( $u$ ), quedando la regresión a estimar por MCO la

---

<sup>32</sup> Un clásico ejemplo de DDD es el expuesto por Gruber (1994).

descrita en (24).

$$y_{it} = \alpha_i + \beta' x_{it} + \mu_{it} \quad (23)$$

$$y_{it} = \alpha_i + \beta' x_{it} + \gamma' z_{it} + u_{it} \quad (24)$$

Luego, en el caso de que estemos en presencia de factores invariantes en el tiempo pero heterogéneas entre individuos ( $z_{it} = z_i$ ) la estimación consistente y eficiente del parámetro de interés ( $\beta$ ) se puede obtener tomando la diferencia entre las observaciones en el tiempo, tal como se indica en la siguiente expresión (25).

$$y_{it} - y_{it-1} = \beta' (x_{it} - x_{it-1}) + (u_{it} - u_{it-1}) \quad (25)$$

En el caso de estar en presencia de efectos aleatorios pero constantes entre los agentes ( $z_{it} = z_t$ ) se procede a tomar la diferencia entre las observaciones individuales respecto del promedio de todos los individuos en  $t$ , tal como se muestra en la siguiente expresión (26):

$$y_{it} - \bar{y}_t = \beta' (x_{it} - \bar{x}_t) + (u_{it} - \bar{u}_t) \quad (26)$$

Cabe señalar, que respecto de esta metodología se ha explicado su esencia pero no su forma de estimación ya que en general lo que interesa es más bien controlar por estos factores no observables y no estimarlos<sup>33</sup>.

Otras ventajas de esta metodología son la mayor disponibilidad de observaciones, lo que incrementa los grados de libertad y eficiencia de las estimaciones, permite discriminar entre diversas hipótesis respecto al fenómeno de causalidad. A su vez, dependiendo del largo del panel estos nos pueden dar información sobre la velocidad de ajuste a shocks exógenos.

Entre las desventajas asociadas a la metodología se encuentran la recolección de la información, tanto por su costo como por el problema de pérdida de muestra en el tiempo (*attrition*) y al error de medición, que puede exacerbarse.

### ***Regression Discontinuity Design (RDD)***

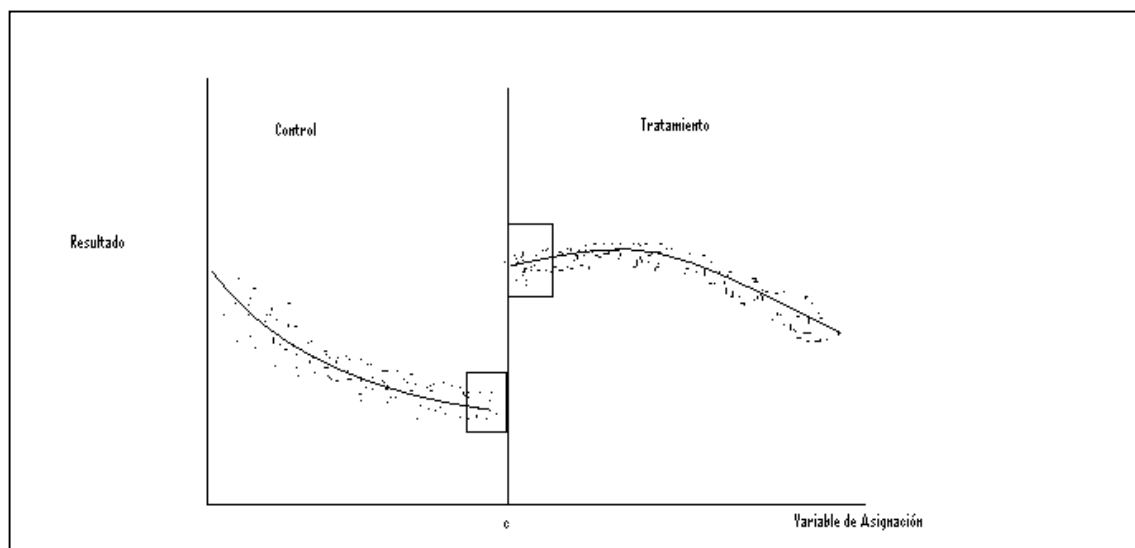
Es un método muy usado en evaluación de programas cuando la regla de selección

---

<sup>33</sup> Mayores referencias de esta metodología y su forma de estimación en Angrist y Pischke (2008), Cameron y Trivedi (2005), entre otros autores.

del tratamiento es una función discontinua de una variable observable<sup>34</sup>. En otras palabras, los seleccionados a ser tratados son aquellos que presentan un valor numérico de la variable de asignación ( $c$ ) por sobre o debajo un determinado punto de corte ( $\bar{c}$ ). Luego, el efecto del tratamiento o parámetro de interés es la discontinuidad en sí que se genera en dicho punto de corte en la variable de resultado (Figura 7).

**Figura 7: Caracterización de *Regression Discontinuity Design***



Fuente: Elaboración propia.

Al utilizar este método de estimación, tal como destaca Lee (2008), no es necesario asumir que RDD es tan bueno como un experimento aleatorio; sino, que la aleatorización que se da en el punto de corte es consecuencia de la incapacidad de los agentes de controlar la variable de asignación del tratamiento. Por lo tanto, el efecto o impacto estimado es de carácter local (en la vecindad del punto de corte) y no refleja a toda la población tratada.

Existen, hasta el momento, tres diseños de RDD los que se encuentran en función de la naturaleza de los datos, la forma de focalización y la implementación de la intervención. A su vez, el tipo de impacto estimado dependerá del diseño de RDD<sup>35</sup>.

Para la aplicación de esta metodología se requiere<sup>36</sup>:

- La existencia de la discontinuidad en el punto de corte (tanto en la probabilidad de recibir el tratamiento como en la variable de resultado).
- Continuidad de la esperanza condicional de la variable de resultado.
- Independencia condicional del tratamiento respecto de la variable de asignación.

<sup>34</sup> Esta metodología fue introducida en la literatura por Thistlethwaite y Campbell (1960).

<sup>35</sup> En Anexo 4 mayor detalle de los tipos de diseño de RDD.

<sup>36</sup> Para un completo desarrollo teórico y empírico de RDD se sugiere ver Lee y Lemieux (2010).

Notar, que no es necesario controlar por otras variables observables para estimar el efecto causal de una determinada intervención. Ahora bien, el supuesto clave para la validez de la identificación radica en la no manipulación completa por parte de los agentes de la variable de asignación (independencia condicional) o en otras palabras los agentes no presentan un comportamiento *foward looking*. McCrary (2008) desarrolla un test indirecto para verificar la existencia de manipulación completa y monótonica<sup>37</sup>. No obstante, es de esperar que la manipulación no sea completa y por tanto, tal como demuestra Lee (2008), ello es suficiente para garantizar la asignación aleatoria del tratamiento en el punto de corte.

Para estimar el efecto tratamiento por medio de RDD se requiere, en términos sencillos, la estimación por la derecha e izquierda del punto de corte de la esperanza condicional en la variable de asignación de la variable de resultado y luego tomar la diferencia entre ambos. Siendo esto último nuestro parámetro de interés.

La estimación del impacto puede ser de carácter paramétrica o no paramétrica. En la primera, la clave es determinación de la forma funcional (y no en sí otros controles) y en la segunda la elección del *bandwidth*<sup>38</sup>. Por su parte, Lee y Lemieux (2010) plantean que ambas metodologías actúan de forma complementaria para el análisis.

## IV.2.2 Efectos Heterogéneos

Hasta el momento hemos descrito metodologías que intentan identificar el efecto promedio de una política en una determinada población objetivo. No obstante, es claro que la intervención puede afectar de forma muy distinta, en dirección y/o magnitud, a sub-grupos de dicha población. Luego sería necesario, dependiendo del tipo de programa y de las preguntas de interés de la evaluación, identificar la heterogeneidad de los efectos causales. Para esto último existen dos metodologías, las que serán enunciadas a continuación.

### **Quantil Regression**

Es un método semi-paramétrico que estima el impacto de una intervención sobre cuantiles de la distribución de la población, en lugar de la media condicional, permitiendo el análisis de efectos heterogéneos. Al igual que en el resto de las metodologías antes descritas se está comprando resultados entre tratados y no tratados, siendo la diferencia la comparación a nivel de percentiles de la distribución entre los grupos.

El punto de inicio de esta metodología es la determinación de *Conditional Quantile Function* (CQF) (27). En el caso de que la función de cuenta de una variable continua la CQF queda representada por la función inversa de la acumulada de la variable de resultado (Y), en el cuantil  $\lambda$ , condicional a las variables observables (X), (28)<sup>39</sup>. Luego, la determinación de la CQF proviene de un proceso de minimización “similar” al que se realiza en la metodología de MCO, tal como se muestra en (29).

---

<sup>37</sup> El propósito del test es analizar si existe discontinuidad en la densidad de la variable de asignación.

<sup>38</sup> Imbens y Kalyanaraman (2009) desarrollan un método de estimación del *bandwidth* óptimo. Éste es calculado automáticamente al estimar una regresión discontinua en STATA.

<sup>39</sup> Función de Distribución acumulada de la variable de resultado.

$$Q_\lambda(Y_i|X_i) = \inf\{y: F_y(y|X_i) \geq \lambda\} \quad (27)$$

$$Q_\lambda(Y_i|X_i) = F_y^{-1}(\lambda|X_i) \quad (28)$$

$$Q_\lambda(Y_i|X_i) = \operatorname{argmin}_q E[\rho_\lambda(Y_i - q(X_i))] \quad (29)$$

$\rho$  corresponde a la *check function*, forma funcional de penalización de los errores.

Algunos aspectos que se deben tener en cuenta al momento de utilizar esta metodología son:

- Al estimar efectos causales para pequeños grupos de la muestra total, el número de datos de la estimación se reduce especialmente en las colas de la distribución. A su vez, la posibilidad de analizar el impacto en diversos subgrupos dependerá del número de observaciones que se dispongan. Lo anterior afectando la estimación de errores estándares, que deberán ser estimados vía *bootstrap*.
- No es claro si es que las personas, por ejemplo, que estaban en el décimo percentil originalmente son los mismos que están en el décimo percentil después de la intervención, depende de si es que el orden (*rank*) dentro de la distribución se mantiene. En caso de no cumplirse lo anterior se altera y complejiza la interpretación de los resultados e incluso del impacto de la intervención.

### ***Estrategia de Efectos Marginales***

La estrategia de efectos marginales enunciada por Bjorklund y Moffitt (1987) y desarrollada por Heckman y Vytlacil (1999, 2005, 2007) presenta una sinergia entre el enfoque estructural (desarrollo de modelos estructurales) con el tradicional de evaluación de impacto.

Consideremos el siguiente modelo, donde (30) representa el resultado potencial de ser tratado y (31) el de no serlo. Al igual que en los desarrollos anteriores,  $x$  representa las características observables,  $u$  las no observables y  $\beta$  el parámetro de interés. A su vez, se considera la ecuación de decisión de ser tratado, reflejando el proceso de autoselección, la cual se caracteriza por medio de una variable latente (32)<sup>40</sup>.

$$Y^1 = x\beta^1 + u^1 \quad (30)$$

$$Y^0 = x\beta^0 + u^0 \quad (31)$$

$$T^* = z\theta + u^T \quad (32)$$

---

<sup>40</sup> La variable latente, es aquella que genera la variable dicotómica de interés de acuerdo a la siguiente regla (siguiendo la nomenclatura de (28)):

$$T(z) = \mathbf{1}[T^* \geq 0] = \mathbf{1}[z\theta + u^T \geq 0]$$

El *Marginal Treatment Effect* (MTE) queda definido por la expresión (33), si  $(u^1, u^0, u^T)$  es independiente de  $x$ , y lo que representa es el impacto del tratamiento sobre aquellos individuos que se encuentran indiferentes entre estar o no afectados a la intervención y tal como señala Schroeder (2010), corresponde a una generalización de un LATE.

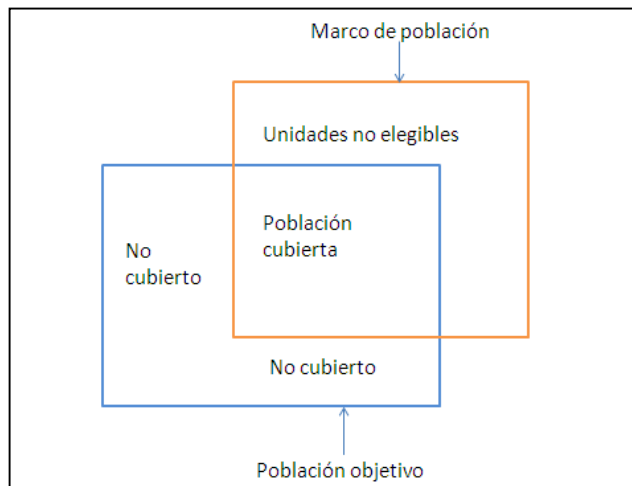
$$\begin{aligned}
 \text{MTE}(x, u^T) &= E(Y^1 - Y^0 | x, u^T) && (33) \\
 &= x(\beta^1 - \beta^0) + E(u^1 - u^0 | u^T, x) \\
 &= x(\beta^1 - \beta^0) + E(u^1 - u^0 | u^T)
 \end{aligned}$$

#### IV.2.3 Población Objetivo y Marco Muestral

La población objetivo consiste en la definición de las unidades a ser estudiadas a partir del universo de la población más extenso. Por su parte, el marco muestral es el registro de los miembros de la población objetivo, que tienen la posibilidad de ser elegidos en el muestra.

Los marcos muestrales pueden existir, por ejemplo por medio de registros administrativos o encuestas de mayor envergadura, o ser construidos en función de los objetivos del estudio. En la Figura 8, se da cuenta de la cobertura del marco muestral.

**Figura 8: Cobertura Marco Muestral**



Fuente: Elaboración propia.



## **Muestra y Entrevistados**

La muestra es un subconjunto del marco muestral desde el cual se realizará la medición. El diseño muestral, el que puede ser probabilístico o no, influirá en la forma y metodología de la selección de las unidades a encuestar. En el caso de que sea de carácter probabilístico, la unidad seleccionada tendrá asociada una probabilidad de selección, el que posteriormente es utilizado para la ponderación de los datos.<sup>41</sup>

Los entrevistados son los sujetos a quienes se les aplica la entrevista. Éste puede ser diferente a la unidad muestral, dependiendo del tipo de análisis.

## **Errores**

Existen distintos tipos de errores que se pueden producir durante el proceso de levantamiento de información, los que deben ser considerados tanto para la determinación del tamaño muestral como para el análisis de los datos.

- a. Error de medida: sus orígenes provienen ya sea por problemas de diseño del cuestionario, de respuesta de los entrevistados, de recolección de datos, de aplicación por parte del encuestador, entre otros. Los efectos de este error pueden ser sesgo de respuesta y varianza de la respuesta.
- b. Error de procesamiento: su origen puede ser de la digitación de los datos, codificación, ajustes en los factores de expansión, y otros.
- c. Error de cobertura: se produce cuando la población no tiene posibilidad de ser elegida pues no se encuentra en el marco muestral. Este error se reporta por medio de la tasa de cobertura.

$$Tasa\ de\ cobertura = \frac{Población\ objetivo\ no\ elegible}{Población\ total} \quad (34)$$

- d. Error de muestreo: depende del diseño muestral empleado y de la forma de selección de unidades muestrales dentro del marco muestral disponible. Este tipo de error se verá con detalle en el punto III del presente anexo metodológico.
- e. Error de no respuesta: la no respuesta puede ser total (el entrevistado no es localizado o no quiere participar) o parcial (un modulo o preguntas del cuestionario). El efecto de este error es el sesgo de las estimaciones de los parámetros poblacionales.
- f. Error total: corresponde a la suma de todos los errores antes señalados.

---

<sup>41</sup> Las muestras no probabilísticas son utilizadas en general para trabajos cualitativos y el problema de las mismas es que no es posible generar inferencia estadística robusta a partir de ella.

### **Medida de calidad de la herramienta estadística**

La medida tradicional es la estimación del error cuadrático medio (ECM), que considera los efectos de cada uno de los componentes del error total.

$$ECM = \text{Segso}^2 + \text{Varianza} \quad (35)$$

Para una sinopsis de los principales elementos a considerar en el diseño muestral de un instrumento estandarizado con representatividad estadística de una población, ver el Anexo III.

### **IV.3 Metodología Evaluación del Gasto Institucional (EGI)**

El foco principal de esta línea es evaluar la eficacia de la institución en la provisión de bienes y servicios, en concordancia con sus objetivos estratégicos, e identificar mejoras de eficiencia en los procesos de provisión de aquellos bienes y servicios más relevantes del quehacer institucional. Utiliza información existente que maneja la institución evaluada y levanta nueva información si se considera necesario.

Esta línea utiliza como antecedentes básicos las definiciones estratégicas, estructura organizacional y presupuesto de la institución a evaluar, así como también documentación de los procesos de provisión de los bienes y servicios, junto con los instrumentos y mecanismos mediante los cuales la institución mide sus resultados en las distintas dimensiones y ámbitos de control.

Estos antecedentes, junto con información que el equipo evaluador levante en cada caso, se utiliza para evaluar:

(i) Consistencia entre misión, objetivos estratégicos, oferta de bienes y servicios (productos estratégicos), y beneficiarios. Para ello, se evalúa si los bienes y/o servicios (productos estratégicos) que entrega la institución son los necesarios y suficientes para cumplir con los objetivos institucionales y si los objetivos definidos permiten cumplir la Misión. También considera el análisis del marco legal en que se basa el accionar de la institución.

Por último, se evalúa la calidad y pertinencia de los indicadores definidos para medir los objetivos planteados. Es importante asegurar que el diseño de los indicadores sea consistente con los objetivos que se quiere medir.

(ii) Resultados y uso de recursos, medidos a través de indicadores de desempeño de eficacia, calidad, eficiencia y economía, obtenidos de los productos estratégicos, vinculando estos resultados con los recursos asignados para la provisión de estos beneficios.

Es importante destacar que, si bien las EGI no tienen asociada una metodología específica, la mirada está siempre orientada a analizar y evaluar los resultados de la institución, alineando el análisis a la eficacia en la consecución de objetivos institucionales, en función de la hipótesis de causalidad que ésta sostenga que existe

entre dichos objetivos, los bienes y servicios que provee, y si los bienes y servicios que provee la institución tienen un costo razonable.

#### IV.4 Metodología Evaluación de Programas Nuevos

Tal como se indicó anteriormente, la metodología utilizada ha sido desarrollada ad hoc a cada evaluación, con el objeto de identificar y fundamentar las eventuales brechas entre el diseño propuesto y el efectivo (en implementación). Luego, una vez conocido en profundidad el diseño y la operación del programa, analiza la factibilidad y pertinencia de diseñar y eventualmente, realizar una evaluación de impacto, la cual se basa principalmente en metodologías experimentales, las que fueron revisadas en la sección anterior.

#### IV.5 Instrumentos de Recolección de Información

El estudio, análisis y evaluación de una intervención pública requiere de información lo más fidedigna que sea posible, tal que permita responder, de la forma más precisa posible, las preguntas planteadas en la evaluación a realizar. La información o datos a utilizar en una investigación pueden ser de naturaleza cualitativa o cuantitativa. La información cualitativa nos permite describir atributos de una intervención, no factibles de ser medidas de forma directa, sino que a través de la experiencia de los actores involucrados. En la medida de lo posible esta información puede ser categorizada bajo una escala ordinal o nominal. Por su parte, la información cuantitativa, es aquella de carácter numérica y nos permite conocer la magnitud o medida de un atributo. Ésta puede ser de carácter discreta (de conteo) o continua.

Para obtener información y, en especial datos, se puede recurrir a información existente de la intervención, usualmente proveniente de la operación de ésta, y que en adelante se le llamará **datos administrativos**; información proveniente de revisión de documentos asociados a la intervención, que en adelante se llamará **análisis documental**; información recabada en consultas o entrevistas a una población determinada, que en adelante se le llamará **información de campo**; entre otra.

De acuerdo al tipo de estudio que se esté realizando, las preguntas de investigación pueden ser respondidas a través de datos cualitativos y/o cuantitativos. Para ambos tipos de dato existen diversas herramientas que permiten recolectar información que no ha sido levantada y/o sistematizada por el programa o institución en estudio, siendo las más utilizadas en las evaluaciones desarrolladas por DIPRES las entrevistas (semiestructuradas), los focus group, y las encuestas, siendo posible el uso de otro instrumental si se estima pertinente.

*Entrevistas Semiestructuradas:* La entrevista es una técnica de investigación cualitativa<sup>42</sup>, en la cual el investigador solicita información a una o más personas, referida a un problema determinado. En una entrevista semiestructurada el investigador construye una pauta que guía la conversación con el/los entrevistado(s),

---

<sup>42</sup> En general, las técnicas cualitativas de investigación buscan comprender el fundamento y significado que las personas entregan a sus experiencias, buscando analizar casos concretos y particulares. Para mayor detalle, ver Flick (2004) y Corbetta (2007).

que le permite enfocarse en los temas atingentes al objeto de estudio, e ir formulando y ordenando las preguntas de una forma coherente.

*Focus Group*: Corresponde a una técnica de investigación cualitativa<sup>43</sup> que, si bien se origina en el contexto de los estudios de mercado, actualmente es utilizada en diversos ámbitos de la investigación social.

Consiste en una entrevista grupal, moderada por un entrevistador que orienta la conversación a partir de una pauta previamente diseñada, de tal modo de enfocarse en los temas relevantes para la investigación. Lo valioso de esta técnica es la interacción de los participantes, permitiendo que cada uno evalúe la opinión de los otros, generando una opinión coherente y compartida entre los miembros del grupo.

*Encuestas*: Es una técnica de investigación cuantitativa. Consiste en un método de recolección sistemática de información que permite analizar en detalle la problemática de una población en particular y debe utilizarse cuando no existe la información disponible o no se encuentra en el formato adecuado para la generación de análisis. Dentro de los puntos importantes a tener en consideración respecto a las encuestas, es su diseño y método de recolección<sup>44</sup>, tal que ésta sea de fácil comprensión para el entrevistado; que se pruebe previamente su funcionamiento a través de un pre-test, y que en la implementación de la encuesta se desarrollen todas las actividades pertinentes para evitar errores de medida<sup>45</sup>, que ésta represente correctamente a la población<sup>46</sup>, y que permita realizar los análisis estadísticos que sean necesarios.

Existen distintos factores que diferencian este tipo de instrumentos, destacándose la representatividad de éstos, debido a que usualmente<sup>47</sup> los *focus group* y entrevistas semiestructuradas no se aplican en forma tal que permitan realizar análisis estadísticamente representativos de la población; el uso que se les da, donde los *focus group* y entrevistas semiestructuradas mayoritariamente se utilizan para orientar las preguntas de la evaluación, entregar información cualitativa respecto a la evaluación, y complementar el análisis cuantitativo realizada en ésta. Por su parte las encuestas, siendo instrumentos que permiten la obtención de datos cuantitativos, se utilizan mayormente para obtener información estadísticamente representativa de una población determinada, y si bien permiten recabar información cualitativa, por su naturaleza estructurada, presenta limitaciones respecto a la amplitud y profundidad de la información que se pueda abordar en este ámbito. Las restricciones son menores para recolectar datos de carácter cuantitativo, si bien la naturaleza del programa, de la población, y de las restricciones cognitivas de quienes responden estos instrumentos también limita la calidad de la información que éstas recolectan.

---

<sup>43</sup> *Ibíd.* [Nota anterior]

<sup>44</sup> Para más detalles ver *Measurement Errors in Surveys* .Paul P. Biemer, Robert M. Groves, Lars E. Lyberg and Nancy A. Mathiowetz (1991)

<sup>45</sup> *Ibíd.* [Nota anterior].

<sup>46</sup> Para más detalles, consultar *Survey Sampling* .Leslie Kish (1995).

<sup>47</sup> Si bien es posible desarrollar *focus group* o entrevistas semiestructuradas que respondan a un criterio de representatividad poblacional, lo oneroso que ello significa, así como su uso complementario a otra información que sí cumpla con criterios de representatividad, llevan a que ambos instrumentos muchas veces no sean aplicados siguiendo criterios de representación estadística.

## Anexo I: Rol del Estado y Fallas de Mercado<sup>48</sup>

En sus distintas funciones - estabilización, asignación y distribución- el Estado puede justificar su intervención para lograr mayor eficiencia en los mercados, estabilidad en la economía, o bien una mayor equidad en la distribución de los recursos en la sociedad.

En la economía existen situaciones que hacen que los equilibrios alcanzados se contrapongan con algunos de los objetivos del Estado, por ello según su función podrá intervenir de diferentes formas.

### Asignación

En este contexto se utiliza el concepto de *fallas de mercado*, para referirse a las situaciones en que la economía no logra llegar a equilibrios eficientes. A continuación se detallan algunas de estas fallas de mercado.

- a) Competencia Imperfecta: Para que la economía logre un equilibrio eficiente en los mercados, es necesario que se den las condiciones de competencia perfecta. Cuando ello no se logra por diversos motivos, existe un espacio para que el Estado pueda intervenir a través de políticas de competencia (*antitrust*) y/o regulaciones en pos de lograr un equilibrio eficiente.
- b) Bienes públicos: Existen algunos bienes en la economía, los cuales presentan las características de no rivalidad y no exclusión en su consumo. Ello implica que no se puede cobrar por su consumo, precisamente por su característica de no exclusión, y una tarifa o precio implicaría la existencia de individuos polizones o *free-riders*<sup>49</sup>. Ante esta falla, por la naturaleza del bien de consumo, a través del mercado se proporciona una cantidad sub óptima del bien, por lo que se justifica la intervención del Estado cuya solución óptima en la mayoría de los casos corresponde a la provisión pública del bien.
- c) Externalidades: Existen casos en que las acciones de algún agente en la economía afectan a otros agentes, generando un costo o beneficio, para los cuales no existe compensación. Estos efectos se denominan externalidades, las cuales pueden ser positivas o negativas. Para ambos casos el equilibrio socialmente óptimo será diferente al equilibrio de mercado alcanzado (bajo el supuesto de competencia perfecta), lo que implicará una sub producción o sobre producción/consumo de un determinado bien o servicio. Ante esta situación, el Estado puede intervenir e implementar ciertos mecanismos que permitan acercarse a puntos más cercanos al equilibrio.
- d) Mercados Incompletos: Cuando el mercado no suministra un bien o servicio aún cuando el costo de hacerlo es inferior a la disposición a pagar de los consumidores, existe una falla de mercado, la cual se denomina como la existencia de *mercados incompletos*. Existen diversas causas por las que un mercado puede ser incompleto (por ejemplo problemas de información, existencia de costos de

---

<sup>48</sup> Basado en, Stiglitz (2000).

<sup>49</sup> Individuos que desean obtener los beneficios del consumo de un bien o servicio, no estando dispuestos a pagar por él.

transacción, innovación, entre otros), por lo que las soluciones o intervención del Estado deberá enfocarse en resolver las causas que originan el problema, en pos del correcto funcionamiento o creación de dicho mercado.

e) Asimetrías de Información: Uno de los supuestos clave para que el mercado alcance un equilibrio competitivo es que los agentes tienen igual conocimiento de la información del mercado y de los bienes y servicios que se transan. Lo anterior no siempre se cumple y por lo tanto surgen oportunidades de comportamiento estratégico por parte de los agentes que poseen mayor información. Según el tipo de asimetría de información se generan distintas consecuencias en la asignación de mercado. Los casos comúnmente estudiados son los de selección adversa (diferencias en la información respecto de las características de los agentes) y riesgo moral<sup>50</sup> (diferencias en la información respecto del comportamiento de los agentes). Las soluciones o posibles intervenciones del Estado dependerán del tipo de asimetría de información, por lo que en algunos casos a este tipo de problemas pueden resolverse entre los mismos agentes (por ejemplo generando contratos eficientes), o bien a través de la acción del Estado (por ejemplo transparentando y otorgando mayor información a los agentes).

### **Estabilización**

Usualmente la economía experimenta altibajos, lo cual se refleja en variaciones en el producto como en los precios de los bienes y servicios transados. Especialmente cuando estos cambios son desfavorables, como la desaceleración de la economía, una baja producción, alzas sostenidas en los precios, alto desempleo, entre otros, terminan por afectar el bienestar social.

De alguna forma, estos desajustes pueden ser interpretados como evidencia de la existencia de fallas de mercado. Dado lo anterior se justifica una intervención del Estado en su función de *estabilizador*, la cual permita mitigar los efectos adversos de estas perturbaciones en la economía. Los principales mecanismos utilizados para abordar esta tarea son la política fiscal y la política monetaria.

### **Distribución**

Los argumentos presentados hasta este punto se basan en fallas de mercado que impiden a la economía alcanzar sus puntos de equilibrio eficiente. Sin embargo, aún si la economía fuese eficiente, existen otros argumentos que justifican la intervención del Estado. El principal de ellos tiene relación con que un equilibrio eficiente en la economía no necesariamente coincide con la distribución de ingresos socialmente deseada. En este sentido, el Estado tiene como una de sus actividades más importantes la redistribución de la renta, la cual se concreta a través de los programas sociales implementados por los diferentes Ministerios.

---

<sup>50</sup> Para mayor estudio de estos casos ver Jehel y Reny (2011).

## **Anexo II: Definición de Conceptos para el Análisis de Uso de Recursos**

Los conceptos empleados para el análisis de uso de los recursos son aplicables a todas las líneas de evaluación de programas públicos.

### **Fuentes de financiamiento de un Programa**

Existen principalmente dos fuentes de financiamiento de los programas gubernamentales. La primera consiste en todos aquellos recursos que aparecen consignados en la Ley de Presupuestos del Sector Público, dentro de los cuales pueden distinguirse al menos dos situaciones: (i) los recursos que son asignados en la Ley de Presupuestos a la institución administrativamente responsable del programa; (ii) los recursos financieros que aparecen contemplados en el (los) presupuesto(s) de otros organismos públicos. Por ejemplo, si un proyecto es financiado con recursos del Fondo Nacional de Desarrollo Regional (FNDR), sus recursos aparecerán en el presupuesto de la Subsecretaría de Desarrollo Regional (SUBDERE) y en el presupuesto de la entidad que ejecuta el proyecto, por ejemplo: Ministerio de Obras Públicas, CORFO, etc.

La segunda fuente corresponde a recursos provenientes de fuentes extrapresupuestarias, que no aparecen en la Ley de Presupuestos; tales como: aportes de los municipios, de organizaciones comunitarias, de organismos internacionales, de donaciones de privados o de los propios beneficiarios. Ejemplo de lo anterior, son los programas en los cuales los recursos para gastos corrientes provienen del presupuesto público, mientras que la inversión se financia con recursos propios de los municipios participantes. Otro ejemplo, se encuentra en programas públicos en los cuales se exige al postulante que financie una parte de los proyectos.

### **Presupuesto del Programa**

El presupuesto del programa corresponde al presupuesto asignado cada año mediante la Ley de Presupuestos que es aprobada en el Congreso (Presupuesto Inicial), sin incorporar aquellas modificaciones que ocurren a lo largo del año por diversos motivos (reajustes de remuneraciones, leyes que incrementan la base de gastos del programa, situaciones imprevistas, etc.).

Al momento de identificar los recursos del programa en la Ley de Presupuestos, es importante tener presente al menos las tres siguientes situaciones:

- a) Todos los gastos del programa son financiados con cargo a una clasificación presupuestaria claramente identificable, razón por la cual los recursos con que cuenta el programa para el año son directamente obtenibles de la Ley de Presupuestos.
- b) Se asignan recursos al programa a través de una clasificación presupuestaria, pero éstos cubren sólo una parte del total de gastos que el programa ejecuta en el año. Por ejemplo, en muchos casos la Ley de Presupuestos identifica sólo los gastos directamente asociados a la producción de los componentes (gasto en subsidio para obras de riego, becas de perfeccionamiento docente,

pensiones, fondos concursables, etc.), mientras que los gastos administrativos son financiados con cargo a una clasificación presupuestaria más amplia y distinta a la del programa; por ejemplo gastos con cargo al Sub-título “Gastos en Personal” o materiales de oficina se realizan con cargo al Sub-título “Bienes y Servicios de Consumo”, Asignación 001, de la institución responsable.

- c) Existen programas cuyo presupuesto está contenido dentro de una clasificación presupuestaria más amplia, por lo que los recursos financieros con que cuenta deben ser identificados recurriendo a información que entregue la gerencia del programa.

Una adecuada estimación del presupuesto del programa debe tener en cuenta las situaciones descritas anteriormente e identificar todos los recursos del programa contemplados en la Ley de Presupuestos.

### **Gasto devengado del Programa**

El gasto devengado del programa corresponde al registro de todas las obligaciones financieras del programa, en el momento que se generan, independientemente de que éstas hayan sido o no pagadas<sup>51</sup>. La relación entre gasto devengado y presupuesto se denomina porcentaje de ejecución presupuestaria. Puede existir una sub ejecución, cuando el programa no gasta la totalidad de recursos asignados para el año, o una sobre ejecución, cuando el programa consigue incrementos presupuestarios durante el año, a través de modificaciones autorizadas por el Ministerio de Hacienda y, por lo tanto, gasta más de lo consignado en su presupuesto inicial.

### **Gastos de producción del Programa**

Corresponden a los gastos directamente asociados a la producción de los bienes y/o servicios (componentes) que entrega el programa a sus usuarios o beneficiarios (gasto becas, monto subsidios, costo de raciones alimenticias, remuneraciones del personal que labora directamente en la entrega de los servicios, etc.).

### **Costos o gastos monetarios**

Corresponden a los costos del programa que tienen una expresión en términos monetarios, es decir, lo que comúnmente se entiende como gasto devengado en la terminología de la Ley de Presupuestos. Se debe tener cuidado en computar todos los gastos considerados en el presupuesto público, puesto que en muchos casos parte de los gastos de un programa no aparecen en su asignación presupuestaria sino en otra asignación de la institución responsable de su administración<sup>52</sup>, o en el presupuesto de otro organismo relacionado. Ejemplo de esto último son aquellos programas que son coejecutados por más de una institución pública y/o programas que son administrados por el nivel central pero ejecutados por instancias regionales o locales, las que por lo general absorben parte de los gastos de funcionamiento.

---

<sup>51</sup> Fuente: Normativa del Sistema de Contabilidad General de la Nación - Oficio C.G.R. N° 60.820, de 2005.

<sup>52</sup> Otro concepto de gasto de acuerdo con la clasificación presupuestaria.



Lo mismo puede aplicarse respecto de aquellos gastos en que deben incurrir los beneficiarios para acceder a los bienes y servicios que genera el programa. Por ejemplo, un programa de capacitación de trabajadores debiera considerar entre sus costos financieros el gasto en movilización que realizan los beneficiarios para llegar al lugar donde se imparte la capacitación<sup>53</sup>.

### **Costos no financieros<sup>54</sup>**

Los costos no financieros son gastos que no tienen una expresión monetaria, debido a que no representan desembolsos y por ende no afectan el flujo de caja del programa, tal es el caso de la depreciación de los vehículos, de las oficinas y del resto de las instalaciones, y el valor del tiempo que los trabajadores ocupan en trasladarse al lugar donde se imparte la capacitación.

Otro ejemplo en este tipo de costos dice relación con el costo de oportunidad del patrimonio fiscal. Por lo general los análisis de eficiencia que se realizan en el sector público no toman en cuenta que los fondos fiscales tienen un costo de oportunidad financiero (la tasa de interés que podrían estar ganando si se mantuvieran en algún instrumento del mercado de capitales) o que los bienes inmuebles fiscales que se traspasan a particulares tienen un costo alternativo, puesto que podrían ser tasados y vendidos como cualquier otro inmueble.

### **Gastos de administración del Programa<sup>55</sup>**

Los gastos totales de un programa pueden desglosarse en dos grandes categorías. En primer término están los gastos directamente asociados a la producción de los componentes, como por ejemplo: monto gastado en subsidios, gastos de las becas, gastos de producción de prestaciones médicas, transferencias a entidades ejecutoras para el financiamiento de proyectos, etc. Por otra parte, están todos los gastos asociados a generar los servicios de apoyo que el programa necesita para funcionar y para producir los componentes.

Los servicios de apoyo se caracterizan por ser servicios transversales a todos los componentes del programa, por lo general, están vinculados a funciones tales como dirección, selección de personal, capacitación del personal, contabilidad, planificación, evaluación y supervisión, control de gestión, asesoría legal, etc. También es posible considerar la contratación de estudios y asesorías especiales como parte de los servicios de apoyo de un programa.

En consideración a lo anterior, los gastos de administración pueden definirse como todos aquellos desembolsos financieros que están relacionados con la generación de los servicios de apoyo a la producción de los componentes del programa.

Los gastos de administración de un programa pueden tener un componente fijo y otro variable. Ejemplo de los primeros son las inversiones necesarias para el

---

<sup>53</sup> Si bien estos gastos forman parte de los costos monetarios, no se exige su inclusión en las evaluaciones de programas gubernamentales debido a la complejidad metodológica y de información que involucra su estimación.

<sup>54</sup> Si bien estos gastos forman parte de los costos de producción de los componentes, no se exige su inclusión en las evaluaciones de programas gubernamentales (EPG) debido a la complejidad metodológica y de información que involucra su estimación.

<sup>55</sup> Los términos "costo" y "gasto" se usan indistintamente.

funcionamiento del programa: vehículos, oficinas, computadores, mobiliario y, en general, todos aquellos gastos que, en el corto plazo, ocurrirán independientemente del nivel de producción de los componentes. El componente variable de los gastos de administración puede estar asociado a la ejecución de actividades de supervisión, difusión y promoción, selección de beneficiarios y, en general, todos aquellos gastos que aumentan cuando se incrementa el nivel de producción de los componentes o viceversa.

En general dentro de un programa es más fácil identificar y calcular los gastos directamente asociados a la generación de los componentes. Una manera práctica de aproximarse a estimar los gastos administrativos consiste en hacer una estimación del gasto de producción de los componentes y considerar el gasto administrativo como el complemento del primero (el monto que falta para llegar a los gastos totales).

Los gastos administrativos por lo general tienen una expresión financiera, es decir, se materializan en un desembolso de carácter monetario. Lo importante es tener en cuenta que no necesariamente todos los gastos administrativos de un programa están incluidos en su presupuesto. Es común, por ejemplo, que parte de los gastos de administración se realicen con cargo a clasificaciones presupuestarias más amplias y distintas a la del programa. También es frecuente que se traspase una parte importante de los gastos de administración a las instancias locales y/o regionales, según la modalidad de ejecución.

Ante situaciones como las descritas anteriormente, es necesario realizar estimaciones de los gastos administrativos sobre la base de información entregada por la gerencia del programa. Lo anterior no siempre es fácil de realizar, en ocasiones, por ejemplo, el personal administrativo presta apoyo a más de un programa de la institución responsable. En tal caso, una estimación adecuada debe apuntar a determinar el porcentaje de la jornada que este personal destina al programa, y ponderar las remuneraciones respectivas para obtener una aproximación a los gastos en personal.

### **Recuperación de gastos**

Uno de los aspectos relevantes del análisis de economía o desempeño financiero de un programa es el relativo a la recuperación de gastos. Esto último se refiere a los mecanismos mediante los cuales se recuperan total o parcialmente los gastos que le genera al Estado la producción de los bienes y servicios que entrega el programa. En términos simples, la recuperación de gastos se manifiesta a través de mecanismos tales como: pago de derechos o tarifas por venta o acceso a servicios (certificaciones, peajes, peritajes, etc.); copagos por servicios (copago de beneficiarios FONASA en hospitales públicos, por ejemplo) y ventas de bases. El indicador más tradicional de recuperación de gastos es: ingresos propios / gasto total del programa

Una política de recuperación de gastos no es aplicable a la entrega de todos los bienes o servicios, y donde es aplicable puede serlo de forma e intensidad diferente. Al respecto es necesario tener presente las siguientes situaciones:

- i. En el caso de bienes o servicios públicos donde no es posible excluir su uso y para el cual no hay rivalidad en su consumo no es posible (o es muy costoso)

aplicar sistemas de precios, por lo tanto no es posible recuperar todo o parte del costo.

- ii. En el caso de bienes o servicios que, no obstante ser posible su exclusión en el consumo, generen externalidades positivas y/o sean considerados bienes meritorios, las señales de precios no reflejan adecuadamente el valor social del consumo, requiriéndose del financiamiento por parte del Estado de todo o parte del costo de estos bienes o servicios.

## Anexo III: Diseño Muestral

Como fue mencionado anteriormente existen los diseños probabilísticos y no probabilísticos, siendo de mayor interés el probabilístico ya que permite realizar inferencia estadística.

A modo de referencia, en el caso del no probabilístico existen los siguientes diseños posibles: muestreo por conveniencia, muestreo de comparación y muestreo por cuotas. No obstante, este tipo de diseño no es válido para estudios de impacto de programas.

El muestreo probabilístico es una estrategia de selección que se sustenta en el principio de aleatoriedad. La selección de la muestra se basa tanto en la probabilidad de selección de las unidades (iguales o desiguales) como en la mecánica de selección (con reemplazo o sin reemplazo).

A continuación se explicará brevemente los diseños probabilísticos existentes y el cálculo de los tamaños muestrales para aquellos estudios que tienen por finalidad analizar una población particular en un momento del tiempo. En los casos en que se quiera realizar comparaciones entre grupos poblacionales (muy utilizado para estudios biomédicos) o análisis del impacto de una política o tratamiento a un grupo de personas respecto de su contra factual (evaluaciones experimentales) se requieren de otras metodologías para el cálculo del tamaño muestral, los que no serán tratados en este documento.<sup>56</sup>

### 1. Muestreo Aleatorio Simple (MAS)

Es el procedimiento más simple y de referencia para el resto de los diseños muestrales. Consiste en la asignación de igual probabilidad de selección a cada unidad del marco muestral y posteriormente se realiza una selección al azar de las unidades a encuestar. Este diseño puede ser realizado con y sin reemplazo de unidades.

Para el cálculo del tamaño muestral, bajo este tipo de muestreo, se debe considerar el error muestral o máximo admisible ( $\delta$ )<sup>57</sup>, la probabilidad o error tipo I ( $\alpha$ )<sup>58</sup>, una distribución normal y el tamaño de la población (N).<sup>59</sup>

La fórmula de cálculo del tamaño muestral del MAS depende la naturaleza del parámetro de interés, el tipo de error establecido, entre otros. En el Cuadro 5 se presentan todas las formulas de cálculo posibles para este tipo de diseño.

A continuación, la ecuación (3) es una de las posibles formulas a utilizar para el cálculo del tamaño muestral, donde Z corresponde al valor tabla de una distribución

---

<sup>56</sup> A modo de referencia, para la realización de comparaciones simples entre grupos ver Duffau y Gastón (1999). Por su parte, para estudios de carácter experimental ver Duflo, Glennerster y Kremer (2007).

<sup>57</sup> El error muestral se encuentra en la misma unidad de medida que el indicador principal que se quiere estudiar y a su vez depende del tipo de diseño muestral a utilizar.

<sup>58</sup> El error tipo I, establecido por el investigador, es la probabilidad de rechazar las hipótesis nula cuando esta es verdadera.

<sup>59</sup> Para una revisión completa de las diferentes técnicas de selección de tamaño muestral, revisar Deaton (1997).

normal a un nivel de significancia  $\alpha$  y  $S$  corresponde a la desviación estándar de los datos.

$$n = \frac{Z_{\alpha}^2 S^2 N}{N\delta^2 + Z_{\alpha}^2 S^2} \quad (3)$$

A modo de ejemplo, supongamos que queremos estudiar el porcentaje de la población de la región del Maule que se alimenta de forma saludable ( $N=1.000.000$ ). A su vez, definimos como error máximo un 3% y de una encuesta anterior sabemos que el 25% de la población tiene una alimentación saludable<sup>60</sup>.

Con estos datos calculamos lo siguiente:

$$S^2 = 0,25(1 - 0,25) = 0,1875$$

$$Z_{\alpha}^2 S^2 = 1,96^2 * 0,1875 = 0,7203$$

$$n = 800$$

Luego, se deben encuestar a 800 personas de la región del Maule para alcanzar la representatividad de la población objetivo.

**Cuadro 5: Fórmulas para estimar tamaño de muestras en MAS**

Tipo de error\Parámetro	Error Absoluto	Error Relativo	Absoluto y coeficiente de confianza	Relativo y coeficiente de confianza
<b>Media</b>	$\frac{NS^2}{N\delta + S^2}$	$\frac{N(\frac{S}{\bar{x}})^2}{N\delta^2 + (\frac{S}{\bar{x}})^2}$	$\frac{Z_{\alpha}^2 S^2 N}{N\delta^2 + Z_{\alpha}^2 S^2}$	$\frac{Z_{\alpha}^2 N(\frac{S}{\bar{x}})^2}{N\delta^2 + Z_{\alpha}^2 (\frac{S}{\bar{x}})^2}$
<b>Proporción</b>	$\frac{Npq}{\delta^2(N-1)\delta^2 + q}$	$\frac{Nq}{p(N-1)\delta^2 + q}$	$\frac{NpqZ_{\alpha}^2}{\delta^2(N-1)p + Z_{\alpha}^2 q}$	$\frac{NqZ_{\alpha}^2}{\delta^2(N-1)p + Z_{\alpha}^2 q}$

Nota:

- Existen dos tipos de errores relativos,  $\delta_r = \frac{\delta}{\bar{y}}$  o  $\delta_r = \frac{\delta}{\hat{p}}$ , donde el primero da cuenta de las desviaciones respecto al promedio poblacional y el segundo mide las desviaciones en porcentaje sobre el valor estimado.
- $q$  corresponde a la proporción poblacional y  $p$  a la estimación de la misma.

## 2. Muestreo Sistemático Aleatorio

Es una manera simple de generar un muestreo estratificado. Consiste en el ordenamiento aleatorio del marco muestral y se enumeran los elementos de 1 a  $k$ , luego en función del tamaño muestral calculado se genera un intervalo  $k = \frac{N}{n}$ . A partir

<sup>60</sup> En caso de no contar con experiencias previas se debe realizar un ejercicio piloto o determinar una cota máxima.

de ese intervalo se selecciona al azar un número, el cual representa la primera unidad de la muestra, posteriormente se le suma a ese número el número del intervalo para obtener el siguiente elemento de la muestra y así sucesivamente.

### **3. Muestreo Aleatorio Estratificado (MAE)**

El propósito de este diseño es garantizar la representatividad de sub grupos (estratos) al interior de la muestra poblacional. Los beneficios de la estratificación dependen del tamaño de la muestra de cada estrato. En general, se opta por una estratificación proporcional, lo que significa que el tamaño de la muestra en los estratos se hace proporcional al tamaño de la población en el estrato.

Cabe señalar que para la aplicación de este diseño se requiere de cierta información poblacional de cada estrato, pero presenta la ventaja de reducir la varianza de los estimadores.

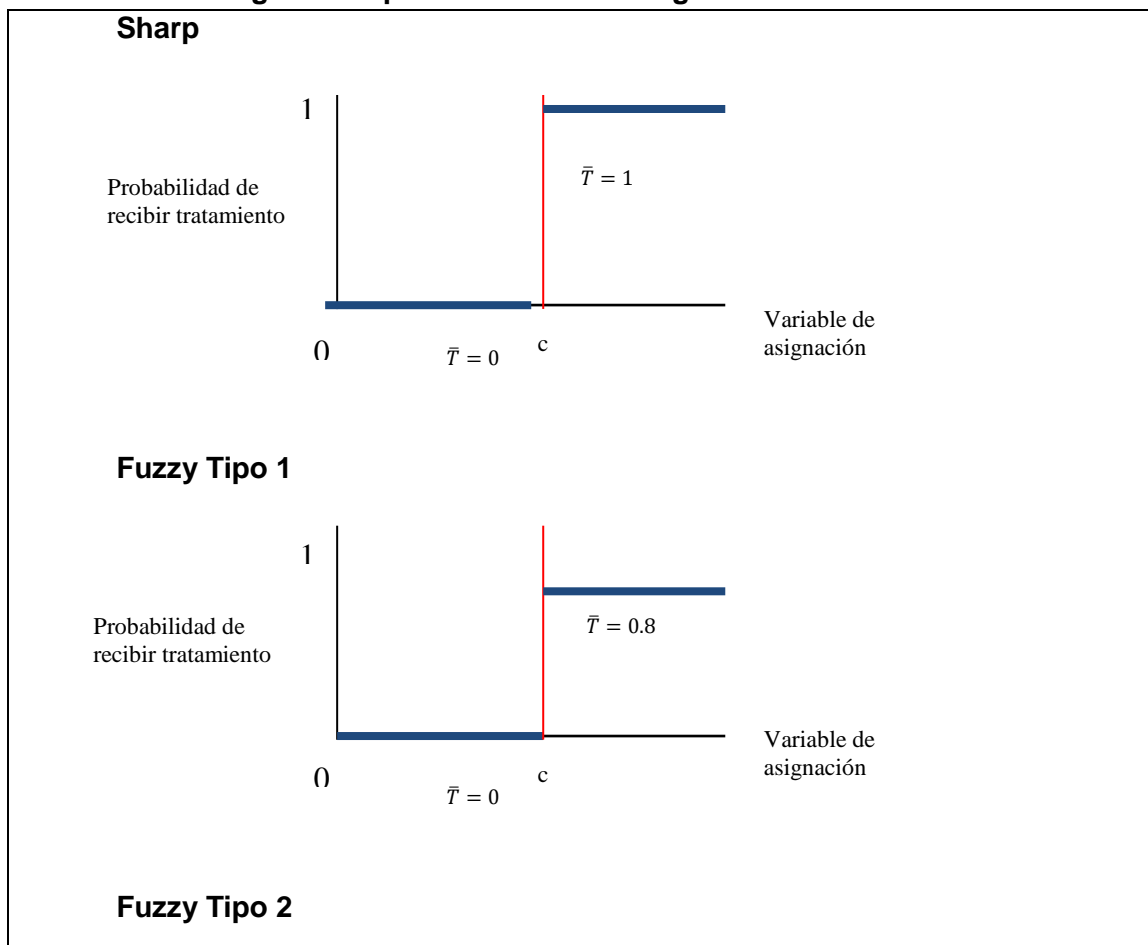
En el cuadro a continuación se indican las formulas asociadas al cálculo del tamaño muestra bajo MAE. Notar que el error muestral del muestreo aleatorio estratificado resulta mayor al error del muestreo aleatorio simple.

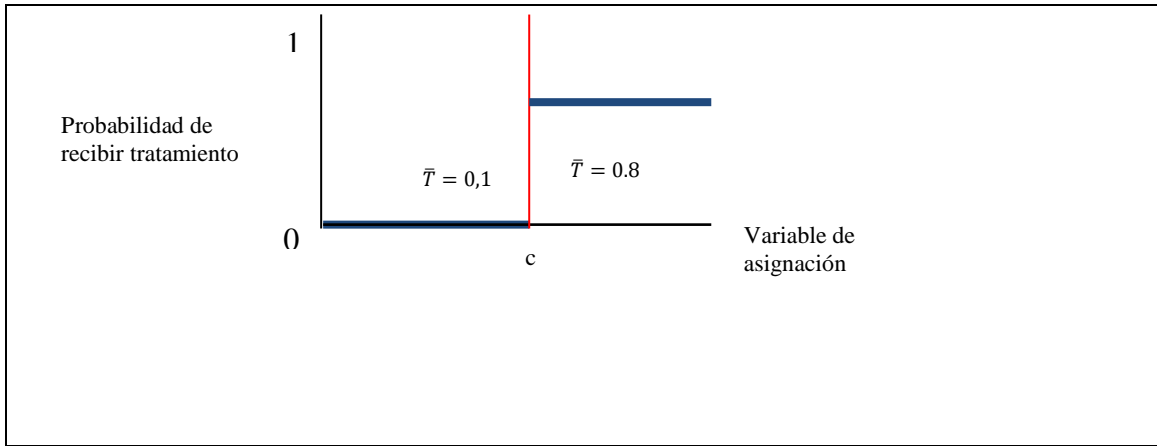
## Anexo IV: Diseños de Regresión Discontinua

Siguiendo a Battistin y Retorre (2008), y ejemplificando con una situación donde la persona es beneficiaria de la política si supera el punto de corte, existen tres tipos de diseño de regresión discontinua (Figura 9):

- Sharp Design**: La probabilidad de ser tratado es 1 si el individuo supera el punto de corte y es cero en caso contrario. Luego, el ser tratado es una función determinística de la variable de asignación y continua excepto en el punto de corte. El impacto estimado a través de este diseño corresponde a un ATE=ATT, pero de carácter local.
- Fuzzy Design Tipo 1**: la probabilidad de ser tratado es discontinua en el punto de corte, pero algunos agentes del grupo de tratamiento no reciben el tratamiento. El impacto estimado con este diseño es un *Intention to Treat on the Treated* (ITT).
- Fuzzy Design Tipo 2**: la probabilidad de ser tratado es discontinua en el punto de corte, pero algunos miembros del grupo de tratamiento no reciben el beneficio y algunos del grupo de control sí. Frente a este diseño el impacto del programa o intervención corresponde a un LATE (al igual que en el caso de aplicar la metodología de variables instrumentales).

Figura 9: Tipos de Diseño de Regresión Discontinua







## **Anexo V: Análisis de Género en la Formulación de Programas**

El Gobierno de Chile reconoce los avances obtenidos en pro de la igualdad de oportunidades entre mujeres y hombres, no obstante también reconoce la necesidad del Estado de evaluar los efectos diferenciados de sus productos (bienes y/o servicios) para hombres y mujeres.

Es por esto que se debe centrar aún más la atención en la integración de una perspectiva o enfoque de género en la determinación de las políticas y programas públicos y sus instrumentos. El proceso de preparación y formulación de los presupuestos públicos es un espacio de trabajo técnico necesario para esto.

### **Enfoque o Perspectiva de Género**

En todas las sociedades las características de mujeres y hombres han estado influidas por múltiples factores, entre otros se encuentran factores históricos, económicos, religiosos, culturales y étnicos. Estas características resultan en experiencias, oportunidades, perspectivas y prioridades diferentes entre mujeres y hombres.

El enfoque o perspectiva de género en políticas públicas es una orientación de la política destinada a considerar y modificar los roles históricos y subordinados en cada sexo, y otorgar igualdad de posibilidades de beneficiarse de las oportunidades del desarrollo y también de contribuir a éste. Se intenciona el enfoque a la mujer por ser el sexo históricamente más desmedrado.

La incorporación de un enfoque o perspectiva de género en los programas públicos debe entenderse como un proceso que permite aumentar la eficacia y eficiencia de las acciones y en que cada uno de los sectores, gradualmente, puede avanzar institucionalizando nuevas prácticas y procesos.

### **Análisis de Género**

El análisis de Género es un proceso de trabajo que permite que una política pública considere las diferencias que existen entre hombres y mujeres, entre otros aspectos, respecto de sus realidades sociales, roles que tienen en la sociedad, expectativas y circunstancias económicas.

A cualquier programa o proyecto se le puede hacer análisis de género, sean estos para hombres, mujeres o ambos. Este análisis es necesario para introducir en las políticas públicas el enfoque, perspectiva o consideraciones de género.

### **Análisis de Género y Ciclo de Programas:**

El análisis de género debe estar presente en todas las etapas de formulación del programa o proyecto.

**Diagnóstico:** El análisis de género en esta fase de la planificación debe abocarse a determinar cómo el problema que se identifica afecta diferenciadamente a hombres y mujeres y las características particulares que este toma en cada caso. Si el diagnóstico es en base a estadísticas la información se debe desagregar por sexo.

**Beneficiarios/as:** El análisis de género debe caracterizar a los beneficiarios respondiendo las siguientes preguntas ¿Quiénes son?; ¿Quiénes participan? ¿Mujeres, hombres, ambos?; ¿Cuáles son los obstáculos para la participación de mujeres y hombres?

**Componentes (bienes y/o servicios):** El análisis de género a nivel de componentes debe identificar con exactitud diferencias necesarias de considerar en la definición del componente específico, dependiendo de si los beneficiarios son hombres, mujeres o ambos, para que efectivamente se logre el propósito del programa. Estas diferencias pueden corresponder a las características de los bienes y/o servicios, al modo de acceso o provisión de éstos y/o los procesos vinculados.

**Indicadores:** Al aplicar análisis de género, los indicadores deben medir los efectos que han tenido las intervenciones realizadas en la situación de hombres y mujeres y en las relaciones que entre ellos se establecen. Entre otros resultados, deben medir el grado de desigualdad que pudiese existir entre hombres y mujeres.

## BIBLIOGRAFÍA

Abadie, A. y G. Imbens (2002). "Simple and Bias-Corrected Matching Estimators for Average". NBER Working Paper No. 283.

Angrist, Joshua D. y Jörn-Steffen Pischke (2009). Mostly Harmless Econometrics: An Empiricist's Companion, Princeton University Press.

Athey, S. y G. Imbens (2006). "Identification and Inference in Nonlinear Difference-in-Differences Models". *Econometrica* 74(2): 431-497.

Banco Mundial (2005). Evaluación de Impacto del Programa de Evaluación de Programas.

Battistin, E. y E. Retorre (2008). "Ineligibles and Eligible Non-Participants as a Double Comparison Group in Regression Discontinuity Designs." *Journal of Econometrics* 142: 715-730.

Bjorklund, A. y R. Moffitt (1987). "The Estimation of Wage and Welfare Gains in Self-Selection Models". *Review of Economics and Statistics* 69: 42-49.

Caliendo, M. y S. Kopeinig (2008) "Some Practical Guidance for the Implementation of Propensity Score Matching". *Journal of Economic Surveys* 22(1): 31-72.

Cameron, A. C. y P. K. Trivedi (2005). Microeconometrics: Methods and Applications, Cambridge University Press.

Carpenter, R. (1977). "Matching When Covariates are Normally Distributed". *Biometrika* 64: 299-307.

Cochran, W. G. y D. B. Rubin (1973). "Controlling Bias in Observational Studies: A Review." *Sankhyā: The Indian Journal of Statistics, Series A* 35(4): 417-446.

Corbetta, P. (2007). Metodología y Técnicas de Investigación Social. Mc Graw Hill.

Cragg, J. y S. Donald (1993). "Testing Identifiability and Specification in Instrumental Variable Models". *Econometric Theory* 9(2): 222-240.

Deaton, A. (1997). The Analysis of Household Surveys: A Microeconomic Approach to Development Policy. Johns Hopkins University Press.

Dipres (2009). Notas Técnicas. División de Control de Gestión, Dipres, Ministerio de Hacienda.

Duffau G. (1999). "Tamaño muestral en estudios biomédicos". *Revista Chilena de Pediatría* 70(4): 314-324.

Duflo E., R. Glennerster y M. Kremer (2008). "Using Randomization in Development Economics Research: A Toolkit", en T. P. Schultz y J. A. Strauss (Eds.). Handbook of Development Economics, Vol. 4, Elsevier.

Flick, U. (2004). Introducción a la Investigación Cualitativa. Fundación Paideia.

Franco, R. (1995). "Focalización y Pobreza". Serie Cuadernos de la CEPAL N° 71, CEPAL.

Greene, W. (2003). Econometric Analysis. Fifth Edition. Prentice Hall.

Groves R., Lyberg L., y Mathiowetz N. (1991). *Measurement Errors in Surveys*. Wiley-Interscience

Gruber, J. (1994). "The Incidence of Mandated Maternity Benefits". *American Economic Review* 84(3): 622-641.

Heckman, J., y E. Vytlacil. (2007). "Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Economic Estimators to Evaluate Social Programs and to Forecast Their Effects in New Environments", en J. Heckman and E. Leamer (Eds.). Handbook of Econometrics. Vol. 6B, Elsevier.

Heckman, J. y S. Navarro-Lozano (2004). "Using Matching, Instrumental Variables, and Control Functions to Estimate Economic Choice Models." *Review of Economics and Statistics* 86 (1): 30-57.

Heckman, J., y E. Vytlacil (1999). "Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects." *Proceedings of the National Academy of Sciences* 96(8): 4730-4734.

Heckman, J., y E. Vytlacil (2005). "Structural Equations, Treatment Effects and Econometric Policy Evaluation." *Econometrica* 73(3): 669-738.

Imbens, G. y J. Angrist (1994). "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62(2): 467-475.

Imbens, G. y K. Kalyanaraman (2009). "Optimal Bandwidth Choice for the Regression Discontinuity Estimator". NBER Working Papers 14726.

Jehel, G. A. y P. J. Reny (2011). Advanced Microeconomic Theory (3<sup>rd</sup> Edition), Prentice Hall.

Kish L. (1995) *Survey Sampling*. Wiley-Interscience.

Lee, D. y T. Lemieux (2010). "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48(2): 281-355.

Lee, D. (2008). "Randomized Experiments from Non-Random Selection in U.S. House Elections". *Journal of Econometrics* 142: 665-674.

McCrary. J. (2008). "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test". *Journal of Econometrics* 142: 698-714.

Ortegón, E., J. F. Pacheco y A. Prieto (2005). Metodología del marco lógico para la planificación, el seguimiento y la evaluación de proyectos y programas. Manuales Series N° 42, CEPAL.

Rosenbaum, P. y D. Rubin (1983). "The Central Role of the Propensity Score in Observational Studies for Causal Effects". *Biometrika* 70(1): 41-55.

Rubin, D. y N. Thomas (2000). "Combining Propensity Score Matching with Additional Adjustments for Prognostic Covariates". *Journal of the American Statistical Association* 95(450): 573-585.

Rubin, R. (1976). "Multivariate Matching Methods That are Equal Percent Bias Reducing, II: Maximums on Bias Reduction for Fixed Sample Sizes". *Biometrics* 32(1): 121-132.

Schroeder, D. (2010). "Marginal treatment effects". In Accounting and Causal Effects (pp. 275-300). Springer.

Stiglitz, J. E. (2000). La economía del sector público, Tercera Edición, Antoni Bosch.

Stock. J. y M. Yogo (2005). "Testing for Weak Instruments in Linear IV Regression" en D.W.K Andrews and J.H. Stock (Eds.) Identification and Inference for Econometric Models: A Festschrift in Honor of Thomas J. Rothenberg, Cambridge University Press.

Thistlethwaite, D. y D.Campbell (1960). "Regression-Discontinuity Analysis: An alternative to de the ExPost Facto Experiment". *Journal of Educational Psychology* 51: 309-317

Wooldridge, J. M. (2002), Econometric Analysis of Cross Section and Panel Data, MIT Press.