

EVALUACIÓN DE PROGRAMAS GUBERNAMENTALES: ANÁLISIS 2011-2019

- En el contexto de los procesos presupuestarios, la Dirección de Presupuestos publica los resultados de las evaluaciones *ex post* realizadas a programas sociales, de fomento productivo y de desarrollo institucional.
- Con la información oficial disponible y bajo algunos supuestos se puede realizar un seguimiento a 131 de los 157 programas evaluados bajo EPG en el período 2011–2019.
- Considerando que a cada programa evaluado se le asigna una calificación final de desempeño -“bueno”, “suficiente”, “insuficiente” y “no demostrado”- llama la atención que sólo el 5% de estos 131 programas obtuvieran una clasificación “buena”; 34% “suficiente” y el 60% restante “insuficiente”.

En el marco de cada proceso presupuestario, la Dirección de Presupuestos (Dipres) del Ministerio de Hacienda publica los resultados de la Evaluación de Programas Gubernamentales (EPG). EPG representa una de las cuatro líneas de evaluación *ex post* de programas sociales, de fomento productivo y de desarrollo institucional, a través de la cual se evalúa la consistencia de los objetivos y del diseño del programa, aspectos de su organización y gestión y resultados a nivel de producto.

Desde el 2003ⁱ la Dipres cuenta con la facultad de efectuar evaluaciones *ex post* de programas. La selección de estos surge de propuestas de los distintos actores del Comité Interministerial, compuesto por Dipres, Secretaría General de la Presidencia (Segpres) y el Ministerio de Desarrollo Social (MDS). Adicionalmente, en el marco de la tramitación de la Ley de Presupuestos, el Congreso puede complementar o sugerir nuevas evaluaciones, plasmándose en algunos casos a través de protocolos de acuerdo entre el Ministerio de Hacienda y el Congreso, o por una resolución exenta del Ministerio de Hacienda.

No hay duda que un sistema de evaluación de programas gubernamentales es una herramienta fundamental para una asignación eficiente y efectiva de los recursos involucrados en el proceso presupuestario. Sin embargo, un análisis de serie de tiempo más detallado de este proceso muestra que el sistema vigente presenta numerosos espacios de mejora.

RESULTADOS EVALUACIÓN *EX POST* 2011 A 2019

Desde 1997 a la fecha, se han evaluado en torno a 420 programas e instituciones bajo la modalidad EPG. Algunos cambios metodológicos implementados a lo largo de estos años en las evaluaciones no permiten realizar un análisis comparativo de serie de tiempo del período completo. Dada la información oficial disponibleⁱⁱ, y a pesar de la última modificación observada en 2018, bajo algunos supuestos se puede realizar un seguimiento a 131 de los 157 programas evaluados bajo EPG en el período 2011–2019.

Considerando que a cada programa evaluado se le asigna una calificación final de desempeñoⁱⁱⁱ -“bueno”, “suficiente”, “insuficiente” y “no demostrado”- llama la atención que sólo el 5% de estos 131 programas obtuvieran una clasificación “buena”; 34% “suficiente” y el 60% restante “insuficiente”^{iv}. En monto de dinero destinado a cada uno de los programas en estas categorías, los resultados encontrados se mantienen, mientras que tomando la antigüedad al momento de ser evaluados, se ve que en promedio se ha ubicado en torno a los 12 años, destacando que los programas calificados como “buenos” presentan menor antigüedad promedio (5 años), mientras que los evaluados como “insuficientes” presentan una antigüedad superior. Por otro lado, también destaca que los evaluados como “suficientes” por tener debilidad en algunas de las dimensiones evaluadas, presentaron 16 años de antigüedad promedio. Estos resultados contrastan con lo que se esperaría en cuanto a que a mayor antigüedad mejor funcionamiento.

PROGRAMAS EVALUADOS EN EL PERÍODO 2011-2019 COMO “SUFICIENTES” TIENEN ANTIGÜEDAD PROMEDIO DE 16 AÑOS

Tabla N°1: Información de los programas evaluados *ex post* por la Dipres

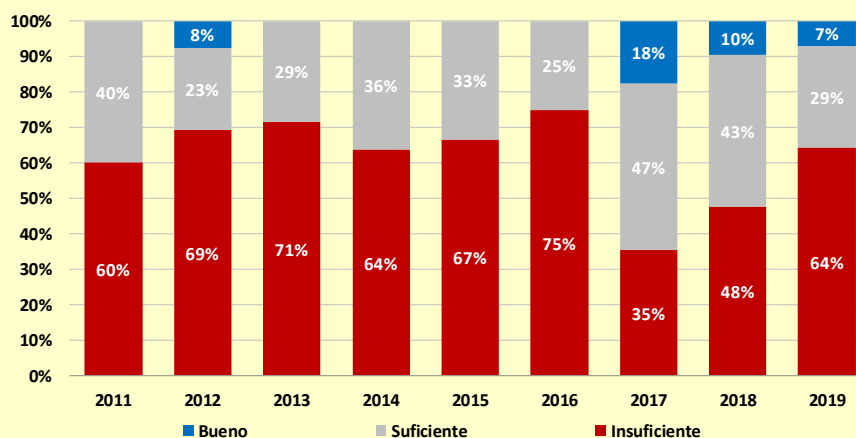
Nivel desempeño	Programas		Programas		Antigüedad promedio al ser evaluados (años)
	número	participación	Mill. US\$	participación	
Bueno	7	5%	343	5%	5
Suficiente	45	34%	2.646	37%	16
Insuficiente/ND	79	60%	4.225	59%	10
	131	100%	7.214	100%	12

Fuente: LyD con información obtenida de la Dirección de Presupuestos, Ministerio de Hacienda.

Una mirada temporal de estas evaluaciones muestra que, aun cuando se observan resultados disímiles en prácticamente todos los años, los recursos destinados a los programas calificados como “insuficientes” representan la mayoría (Gráfico N° 1).

RECURSOS DESTINADOS A PROGRAMAS “INSUFICIENTES” REPRESENTAN LA MAYORÍA

Gráfico N° 1: Participación en el monto total evaluado cada año

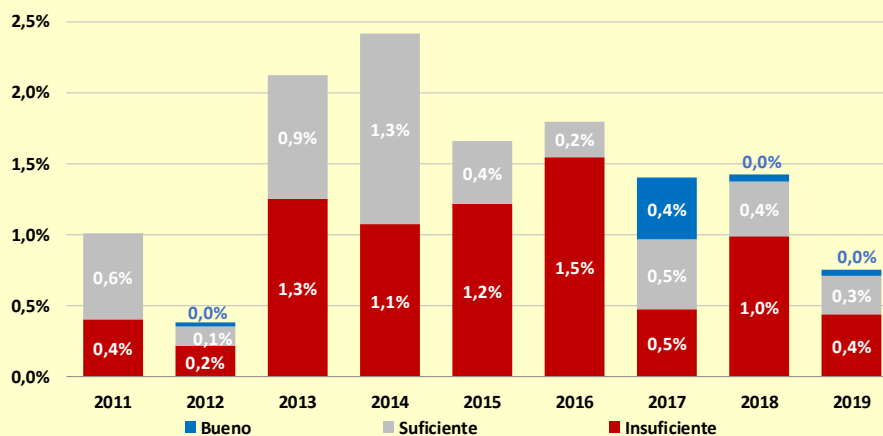


Fuente: LyD con información obtenida de la Dirección de Presupuestos, Ministerio de Hacienda.

Otra mirada interesante, especialmente en el contexto de la necesidad urgente de avanzar en la modernización del Estado para gestionar más eficientemente el gasto público, muestra que los programas evaluados bajo esta metodología (EPG) representan un porcentaje reducido del presupuesto fiscal total de cada año.

LOS RECURSOS EVALUADOS REPRESENTAN UN PORCENTAJE MUY MENOR DEL PRESUPUESTO PÚBLICO DE CADA AÑO

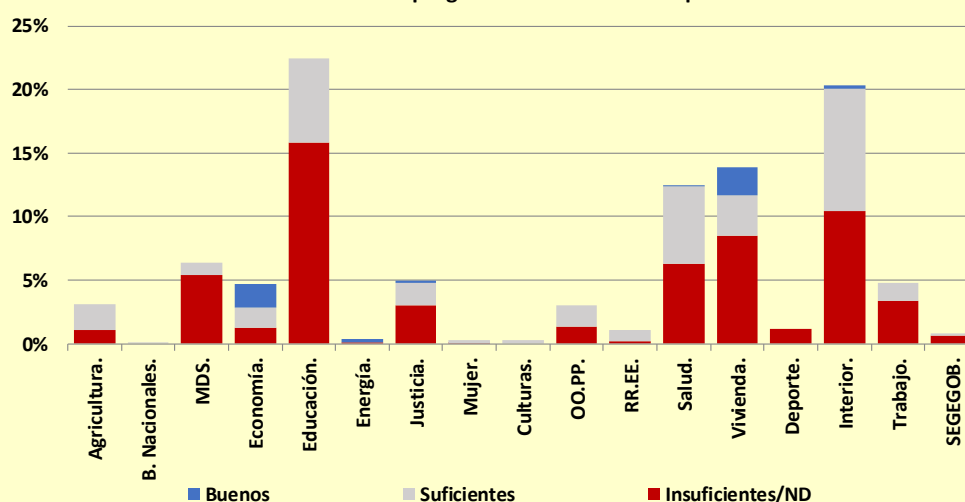
Gráfico N° 2: Participación en el presupuesto anual de cada año



Fuente: LyD con información obtenida de la Dirección de Presupuestos, Ministerio de Hacienda.

AUNQUE EL PROBLEMA DE PROGRAMAS MAL EVALUADOS ES GENERALIZADO, UN NÚMERO ACOTADO DE MINISTERIOS DESTACAN POR SU IMPACTO SOCIAL

Gráfico N° 3: Participación del monto gastado por cada ministerio, en pesos del 2019, en el total de recursos destinados a los programas evaluados en el período 2011-2019



Fuente: LyD con información obtenida de la Dirección de Presupuestos, Ministerio de Hacienda.

Aun cuando prácticamente todos los ministerios tienen algún programa con evaluación “insuficiente”, en términos de los recursos involucrados - y por tanto de su importancia en la política pública- sólo destaca un grupo acotado: Educación, Interior, Vivienda, Salud y Desarrollo Social (Gráfico N° 3).

ALERTAS

Reconociendo la enorme relevancia que presenta para el proceso presupuestario la evaluación de los programas públicos, de un análisis más detallado del actual proceso de evaluación surgen elementos en los cuales es necesario avanzar con miras a lograr una asignación realmente eficiente y eficaz de los recursos involucrados.

Como parte del proceso, y con el objeto que los resultados de aquellas evaluaciones cuya calificación sea inferior a “buen desempeño” se traduzcan en acciones concretas de mejoramiento -ya sea en la gestión de los programas o en la asignación y/o uso de sus recursos- se definen compromisos por parte de los servicios a cargo de los programas evaluados, en conjunto con la Dipres. El proceso de evaluación de una iniciativa se da por finalizado una vez que se incorporan los cambios comprometidos, no habiendo, necesariamente, una nueva evaluación que determine el impacto que la

reformulación o cambio tuvo en los beneficiarios o en la forma de ejecutar los programas. De este modo, no existe un mecanismo concreto que permita determinar si las falencias fueron superadas.

En consecuencia, aun cuando el objetivo de este sistema de evaluación es generar información adecuada para promover la eficiencia y la eficacia en la asignación y uso de los recursos públicos en cada proceso presupuestario, los resultados encontrados en estas evaluaciones no son vinculantes, ni en el corto ni en el mediano plazo, no existiendo una relación directa y explícita entre recursos y desempeño.

Además de la inexistencia del vínculo entre evaluación y asignación de recursos, los programas evaluados anualmente representan un porcentaje menor del gasto total del sector público, limitando aún más el impacto de estas evaluaciones.

Adicionalmente, hay dificultades que surgen en el intento de asignar una calificación a cada aspecto de cada programa a evaluar, particularmente en el último bienio, lo que deja de manifiesto la necesidad de vincular más la puesta en operación de un programa con su método de evaluación. Es decir, en la etapa de diseño de un programa se debiera definir cómo se evaluarán sus resultados. Mejoras en este vínculo permitirían avanzar en un sistema más permanente de evaluación para cada programa y no esporádico como es en la actualidad, permitiendo adicionalmente, evaluar montos más significativos destinados a estos programas.

El análisis muestra que aproximadamente el 60% de los recursos evaluados entre los años 2011 y 2019 se han destinado a programas con calificación “insuficiente”, levantando la legítima preocupación de si este resultado se explica por un sesgo en la selección de los programas a evaluar anualmente o a que son más bien una muestra representativa de todos los programas ejecutados por el Gobierno. Por otro lado, sólo el 5% de los recursos correspondientes a programas que fueron objeto de evaluación han estado destinados a programas “bien evaluados”.

Un elemento fundamental a evaluar en todo programa debiera ser el aspecto de “resultado”. De acuerdo a la información entregada en el bienio 2018-2019, sólo 9 de los 17 programas presentaron calificación para el aspecto “resultados”, el cual considera el cumplimiento de sus propósitos, condición clave a evaluar para determinar la necesidad y utilidad del programa. En términos generales, la falta de calificación de este aspecto se debería a que la evidencia no permite tener un juicio evaluativo, ya sea por falta de información relevante, por el estado de avance en la implementación del

programa o porque, incluso teniendo información sobre resultados, estos no permiten sacar conclusiones.

Llama la atención que los programas con debilidades en sus evaluaciones (“suficiente” e “insuficiente”) sean justamente aquellos que presentan una mayor antigüedad promedio al momento de ser evaluados. A modo de ejemplo, en la evaluación realizada en el 2019 destacaron los siguientes programas:

- i. Rehabilitación y Reinserción Social (Gendarmería), iniciado en 1996 y con desempeño “insuficiente”^v.
- ii. Fondo de Desarrollo Indígena (CONADI), iniciado en 1993 y evaluado con “insuficiente”^{vi}.
- iii. Centro de Recursos para el Aprendizaje (Bibliotecas CRA), iniciado en 1994 y con evaluación “insuficiente”^{vii}.

A la luz de esta información vale la pena preguntarse por qué estos programas, luego de todo el tiempo transcurrido, presentan mala evaluación, con el agravante de reiteradas evaluaciones negativas de algunos de ellos, y continúan recibiendo financiamiento sin que se propongan mejoras sustanciales a los mismos o se evalúe la continuidad del programa. Tal es el caso del programa Fondo de Desarrollo Indígena de CONADI, iniciado en 1993, evaluado negativamente también en 1999 y 2009, con cumplimiento completo de los compromisos asumidos luego de ambas evaluaciones negativas, para luego volver a ser evaluado con mal desempeño en 2019.

CONCLUSIÓN

No hay duda que un sistema de evaluación de programas gubernamentales es una herramienta fundamental para una asignación eficiente y efectiva de los recursos involucrados en el proceso presupuestario. Sin embargo, observamos que el sistema vigente presenta numerosos espacios de mejora como, por ejemplo, vincular el diseño e implementación del programa a indicadores objetivos de medición de los cuatro aspectos a evaluar; implementar algún sistema de vinculación entre las evaluaciones y la continuidad del programa; potenciar la rigurosidad de las evaluaciones; y ampliarlas hacia más programas.

ⁱ 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 N° 1.177 del 11 de diciembre del 2003. No obstante, ya desde fines de los años 90 que Dipres evaluaba determinados programas sociales.

ⁱⁱ Información contenida en la Dipres. Evaluación de Programas e Instituciones.

ⁱⁱⁱ Se evalúa con: i) “bueno” (B) si el programa presenta resultados positivos y significativos a nivel intermedio y final, así como también en las dimensiones de diseño, gestión y producto; ii) “suficiente” (S) si el programa tiene resultados positivos que dan cuenta de sus objetivos, sin embargo, presenta debilidades en algunas de las dimensiones evaluadas; iii) “insuficiente” (I) si presenta resultados insuficientes y/o muestra debilidades en las dimensiones evaluadas que no le permiten dar cuenta de sus objetivos; iv) “no demostrados” (ND) en los casos en que la evaluación no entrega información suficiente que permita obtener resultados concluyentes respecto del desempeño del programa. Esta clasificación es la oficial para el período 2011-2017. Para el bienio 2018- 2019 la clasificación oficial considera las categorías “bueno”, “medio”, “bajo” y “malo”. Para efectos de este análisis, y considerando la similitud de las definiciones de cada calificación, “medio” se compatibiliza con “suficiente”, y “bajo” y “malo” con “insuficiente”.

^{iv} En esta clasificación se incluyeron también aquellos programas cuya evaluación fue “no demostrado”.

^v En la metodología de esa evaluación la clasificación es “bajo desempeño”.

^{vi} En la metodología de esa evaluación la clasificación es “mal desempeño”.

^{vii} En la metodología de esa evaluación la clasificación es “mal desempeño”.