

Dirección de Desarrollo Digital

Unidad de Científicos
de Datos



**El futuro
es de todos**

DNP
Departamento
Nacional de Planeación

EFFECTO DEL SISTEMA GENERAL DE PARTICIPACIONES SOBRE EL DESARROLLO TERRITORIAL

Entidad

- Dirección de Desarrollo Digital
- Dirección de Descentralización y Desarrollo Regional

Sector

Planeación

Lenguaje

R.

Fuente de datos

1. Terridata
2. TIPOLOGIAS DDT - Departamentales
3. TIPOLOGIAS DDT - Municipales

Presentación

Terridata es un proyecto multipropósito motivado por la buena disponibilidad de información a nivel municipal y de alcance nacional, contenida en una estructura de datos panel que centraliza información diversa y completa sobre las principales estadísticas asociadas a la situación de las entidades territoriales. Este proyecto busca capitalizar la oportunidad de establecer comparaciones y generar analítica a partir de información de diferentes categorías y fuentes. Se propusieron un conjunto de metodologías para explotar desde varias perspectivas la información estructurada y no estructurada contenida en Terridata, buscando pronosticar las variables, establecer relaciones causales entre niveles de inversión y desarrollo municipal y aplicar técnicas de minería de texto sobre los Planes de Desarrollo Territorial.

Terridata is a multipurpose project motivated by the availability of rich Municipality-level information gathered in a panel data structure that centralizes diverse and complete information relating to the status of all territorial entities. This project aims to exploit the opportunity to make comparisons and gain analytical insights from data belonging to different categories and sources. A set of methodologies are proposed to exploit the data contained in Terridata, seeking to predict variable series, to find causal relationships between investment levels and municipal development, and to find patterns in the goals set forth in local "Territorial Development Plans".

Objetivo general

Generalizar la metodología desarrollada durante la primera fase de Terridata, para facilitar su aplicación a numerosos ejercicios de evaluación de impacto.

Objetivos específicos

1. Desarrollar una herramienta que permita presentar la evolución de las variables de Terridata asociadas al desarrollo territorial.
2. Evaluar el grado de relación entre los indicadores base de las tipologías DDT con el sistema general de participaciones.
3. Proponer una metodología estadística para estimar el impacto de las principales inversiones sobre los indicadores de desarrollo municipal.

Metodología

Terridata contiene información de los diferentes montos de inversión asignados a los municipios por las principales fuentes que constituyen el Sistema Nacional de Transferencias, en particular el Presupuesto General de la Nación (PGN) y el Sistema General de Regalías (SGR). Estas inversiones tienen una asignación sectorial específica (Educación, Salud, Saneamiento de Agua Potable, etc.). Desde un punto de vista económico, social y administrativo, es relevante indagar por la eficiencia de las inversiones y el impacto causal que las mismas representan para la evolución de los indicadores en las cuáles debe reflejarse el Desarrollo Municipal.

La asignación de diferentes montos de inversión a las entidades territoriales se inscribe dentro de lo que podríamos denominar una situación “no experimental”, debido a que la cantidad que se invierte en los municipios es no aleatoria (*non-random assignment*) y por el contrario responde a una lógica institucional implícita. Por ejemplo, el monto de las inversiones puede estar condicionado por la capacidad que tienen los municipios de señalar eficiencia administrativa. En ese sentido, los municipios que tienen niveles de gasto relativamente bajos pueden ser sistemáticamente diferentes y no constituir un contrafactual válido de los municipios que presentan niveles de inversión relativamente altos, lo cual dificulta la estimación del efecto causal del nivel de inversión sobre el desarrollo municipal. Esto se conoce en la literatura como el problema fundamental de la Evaluación del Impacto y para solventarlo existe un conjunto de metodologías de análisis, entre las cuales contamos el emparejamiento basado en medidas observables (*Matching*).

Teniendo en cuenta la gran cantidad de información contenida en Terridata, el primer paso es una preselección de variables pertinentes a la operación estadística. Para ello, se determinó un conjunto reducido de indicadores que fuesen representativos del Desarrollo Municipal, las cuáles fueron seleccionadas de las categorías de Servicios Públicos, Educación, Salud, Empleo y Vivienda (variables actualmente empleadas para la estimación de diferentes medidas de pobreza (ICV, IPM, entre otras). Por otra parte, se consideró un conjunto de rubros de inversión que deberían afectar (en el sentido de estricta causalidad) estos indicadores, resultando en la conformación de una matriz de efectos teóricos que determinan los casos o instancias interesantes de evaluación de impacto. Nótese que, en un contexto con indicadores de desarrollo municipal y rubros de inversión, tenemos múltiples instancias de evaluación de impacto, de los cuáles solo un subconjunto tiene relevancia teórica.

En este proyecto se implementó una metodología de Evaluación de Impacto basada en el emparejamiento (*Matching*), la cual permite establecer relaciones aproximadamente causales entre los niveles de inversión en las entidades territoriales y el Desarrollo Municipal observado a través de un conjunto de indicadores. En la Figura 1 se presenta un resumen de la relación entre algunos indicadores de inversión y el sector que impactan de forma directa.

Indicador	Inversión	Educación	Salud	Agua Potable	Propósito general	Alimentación escolar	Ribereños	Resguardos indígenas
Cobertura de acueducto (Censo)				■			■	
Cobertura de alcantarillado (Censo)				■			■	
Penetración de banda ancha				■	■			
Déficit cuantitativo de vivienda (Censo)				■	■			
Déficit cualitativo de vivienda (Censo)				■	■			
Tasa de cobertura neta en educación preescolar	■				■	■		
Tasa de cobertura neta en educación primaria	■				■	■		
Tasa de cobertura neta en educación secundaria	■				■	■		
Tasa de cobertura neta en educación media	■				■	■		
Tasa de Mortalidad			■	■	■	■		■
Tasa de Mortalidad Materna			■	■	■	■		■
Tasa de Fecundidad	■	■	■	■	■	■		■
Tasa de Mortalidad Infantil	■	■	■	■	■	■		■
Homicidios * 10000 hab	■	■	■	■	■	■		■
Hurtos * 10000 hab	■	■	■	■	■	■		■

Figura 1: Relación entre algunos indicadores de inversión y algunos sectores económicos.

Bajo esta consideración, la Figura 2 presenta una descripción de los pasos para determinar el impacto de las inversiones en los indicadores.

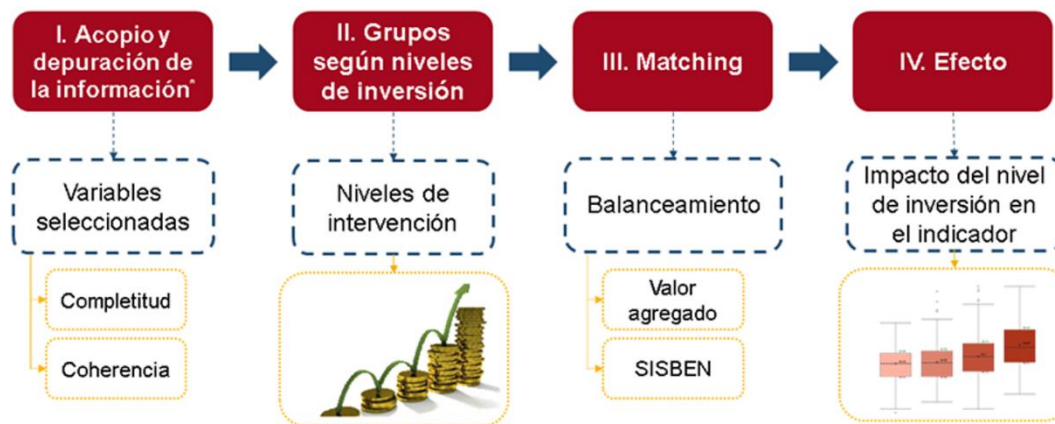


Figura 2: Pasos en la evaluación de impacto.

Una vez seleccionado un caso particular de evaluación de impacto, se procede a un acopio de toda la información pertinente que consiste en la conformación de un objeto tabular con la información asociada al *c-ésimo* ejercicio de evaluación. Por lo tanto, se construye una base de datos depurada que contiene información asociada al monto de la inversión, el indicador de desarrollo municipal y la lista de controles pertinentes para un corte transversal de todos los municipios de Colombia.

El objeto tabular conformado contiene la información mínima suficiente para cierta evaluación de impacto. No obstante, en su formato original la variable de inversión requiere un procesamiento específico. La teoría de evaluación de impacto está diseñada para problemas que admitan una representación a través de un tratamiento binario, es decir, donde se pueden identificar claramente los grupos de control y tratamiento.

Teniendo en cuenta que el monto de inversión es un valor real continuo, se aplicó una discretización que permite identificar varios estratos de inversión. Por lo tanto, se considera que las entidades territoriales pueden ser expuestas a diferentes niveles de inversión, lo cual permite plantear relaciones dinámicas de control/tratamiento entre sucesivos niveles de inversión. Esta discretización satisface un criterio de optimalidad estadística, que consistente en minimizar la varianza dentro de los grupos o clústeres conformados por cada nivel de tratamiento.

El emparejamiento busca que la distribución empírica de las características de los municipios tratados y no tratados sea lo suficientemente similar, como para justificar que las diferencias observadas en el indicador de desarrollo se atribuya únicamente a los diferentes niveles de inversión. En este caso se aplicó un algoritmo de emparejamiento sobre la toda muestra, procurando que las distribuciones de las características de los municipios sean similares en todos los niveles de tratamiento. A continuación, se presenta una descripción del método de emparejamiento empleada en el estudio.

Propensity score matching (PSM)

En el análisis estadístico de los estudios observacionales, el emparejamiento por puntaje de propensión o Propensity score matching (PSM) en inglés, es una técnica estadística de coincidencia que intenta estimar el efecto de un tratamiento, una política, u otra intervención por cuenta de las covariables que predicen que recibe el tratamiento. PSM intenta reducir el sesgo debido a la confusión de las variables que se pueden encontrar en una estimación del efecto del tratamiento obtenido de la simple comparación de los resultados entre unidades que recibieron el tratamiento frente a los que no lo hicieron. La técnica fue publicada por primera vez por Paul Rosenbaum y Donald Rubin en 1983, y aplica el modelo causal de Rubin para los estudios observacionales.

La posibilidad de sesgo surge porque la aparente diferencia en los resultados de estos dos grupos de unidades puede depender de las características que afectaron en la decisión de administrar un cierto tratamiento a una unidad en lugar de por el efecto del tratamiento en sí. En experimentos aleatorios, la aleatorización permite la estimación objetiva de los efectos del tratamiento, para cada covariable aleatorización implica que los grupos de tratamiento se equilibrarán en promedio, por la ley de los grandes números. Desafortunadamente, para los estudios de observación, la asignación de tratamientos a los sujetos de investigación es, por definición, no aleatorio. Este método trata de imitar al azar mediante la creación de una muestra de unidades que recibió el tratamiento que sea comparable en todas las covariables observadas a una muestra de unidades que no recibieron el tratamiento.

Por ejemplo, uno puede estar interesado en saber las consecuencias de fumar o las consecuencias de ir a la universidad. El conjunto 'tratado' son simplemente aquellos- los fumadores, o graduados de la universidad- que se someten en el curso de la vida diaria a aquello que está siendo estudiado por el investigador. En ambos casos no es factible (y quizás poco ético) asignar al azar a las personas a fumar o a la educación universitaria. El efecto estimado del tratamiento mediante la simple comparación de un resultado, por ejemplo, la tasa de cáncer o los ingresos a lo largo de la vida, entre los que fumaban y no fumaban o entre los que asistieron o no asistieron a la universidad estaría sesgado por factores que predicen el tabaquismo o la asistencia a la universidad, respectivamente. PSM intenta controlar estas diferencias para que los grupos que recibieron tratamiento y no tratamiento se aproximen a una investigación aleatoria y ver que estos sean más comparables.

La metodología empleada consiste en la siguiente secuencia de pasos:

1. Realizar una regresión logística teniendo en cuenta las siguientes consideraciones:

- Variable dependiente: Y = 1, si el municipio cumple alguna característica de interés; Y = 0, de lo contrario.
 - Emplear las variables de confusión adecuadas (variables hipotéticamente asociadas con el tratamiento y el resultado) para obtener la puntuación de propensión: probabilidad predicha (p) o $\log [p / (1 - p)]$.
2. Emparejar cada municipio participante a uno o más de los no participantes según el puntaje de propensión. Las metodologías consideradas fueron:
 - Vecino más cercano similar
 - Emparejamiento por calibre (caliper)
 - Emparejamiento según el valor Mahalanobis junto con PSM
 - Emparejamiento por estratificación
 3. Verificar que las covariables está equilibradas entre el grupo de tratamiento y el de comparación de la nueva muestra emparejada o ponderada.
 4. Realizar un análisis multivariado sobre la base de la nueva muestra utilizando los análisis apropiados para diferentes muestras no independientes.

Los resultados obtenidos a partir de estos emparejamientos se presentan en la siguiente sección.

Resultados

En la Tabla 1 se presenta el conjunto de variables considerada para el seguimiento del desarrollo territorial en Colombia a partir de un conjunto de indicadores estratégicos de los sectores de Educación y Salud, y Saneamiento de Agua Potable y Alcantarillado, entre otras según su disponibilidad en el tiempo. La disponibilidad en Terridata de las variables es un proxy de la disponibilidad de información municipal en Colombia.

Variable	Unidad de medida	Concepto Asociado	Sector	Disponibilidad	Fuente
Cobertura de acueducto (REC)	Porcentaje	Agua potable	Vivienda y acceso a servicios públicos	2008 - 2016	DNP a partir de información de la Superintendencia de Servicios Públicos Domiciliarios
Cobertura de alcantarillado (REC)	Porcentaje	Agua potable	Vivienda y acceso a servicios públicos	2008 - 2016	DNP a partir de información de la Superintendencia de Servicios Públicos Domiciliarios
Cobertura neta en educación media	Porcentaje	Calidad gratuidad	Educación	2005 - 2017	Ministerio de Educación Nacional
Cobertura neta en educación preescolar	Porcentaje	Calidad gratuidad	Educación	2005 - 2017	Ministerio de Educación Nacional
Cobertura neta en educación primaria	Porcentaje	Calidad gratuidad	Educación	2005 - 2017	Ministerio de Educación Nacional
Cobertura neta en educación secundaria	Porcentaje	Calidad gratuidad	Educación	2005 - 2017	Ministerio de Educación Nacional

Educación	Puntos	Calidad matrícula	Medición de desempeño municipal	de 2016 - 2016	DNP - DDDR
Pruebas saber 11	Puntos de 100	Calidad matrícula	Educación	2010 - 2017	ICFES
Pruebas saber 11	Puntos de 100	Calidad matrícula	Educación	2010 - 2017	ICFES

Tabla 1: Variables consideradas en el seguimiento del desarrollo territorial en Colombia.

Las transferencias nacionales constituyen un margen de intervención de primer orden para afectar el desarrollo territorial. La Tabla 2 presenta la disponibilidad de información sobre los montos de transferencias nacionales en el sistema SICODIS del DNP por cada concepto del Sistema General de Participaciones

Concepto	Municipios observados
Educación	1101 (100%)
Calidad gratuidad	1101 (100%)
Calidad matrícula	1101 (100%)
Prestación servicios	63 (5.72%)
Salud	1101 (100%)
Régimen subsidiado	1101 (100%)
Salud pública	1101 (100%)
Prestación de servicios y subsidio a la oferta	497 (45.14%)
Agua potable	1101 (100%)
Total SGP	1101 (100%)

Tabla 2: Disponibilidad de información sobre los montos de transferencias nacionales en el sistema SICODIS.

Con el objetivo de presentar de manera ordenada los resultados del presente estudio, se desarrolló utilizando el paquete Shiny de R un tablero de control el cual facilita la consulta de información e interacción con los resultados, el despliegue y acceso a la aplicación se puede realizar utilizando RStudio.

La primera sección del aplicativo entregado se denomina “Reconocimiento de las variables relacionadas”. En esta sección se presenta el comportamiento histórico del indicador desarrollo territorial experimentado por una muestra de las entidades territoriales municipales colombianas. La relación de cambio se aproxima como la suavización lineal del cambio en el indicador sobre la disponibilidad actual en la base Terridata.

El eje de las ordenadas se refiere a los rangos de variación histórica normales del indicador dados por la naturaleza técnica del mismo. Las trayectorias del indicador aluden a los diferentes procesos de cambio atravesados por los municipios en el tiempo. La Figura 4 presenta la evolución temporal de un indicador que el usuario puede escoger en el menú de interacción del aplicativo (ver Figura 3). El menú interactivo permite al usuario del aplicativo apreciar la heterogeneidad en el desarrollo territorial a través de la variabilidad en intercepto y pendiente.

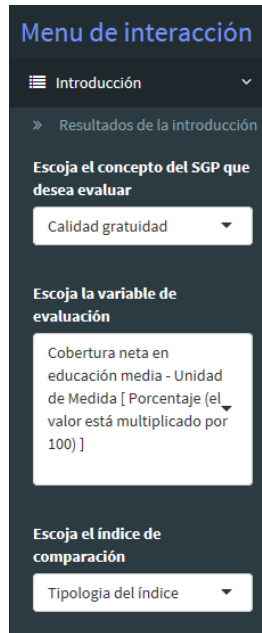


Figura 3: Menú interactivo de la primera sección del aplicativo.

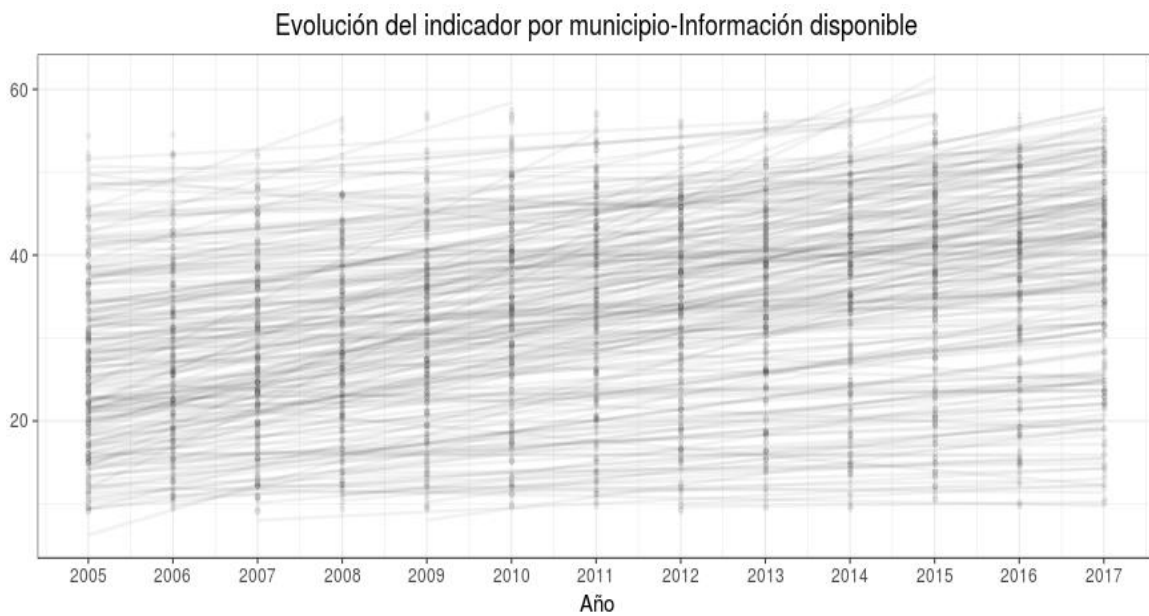


Figura 4: Evolución histórica del indicador con suavización temporal.

La Figura 5 presenta la relación global entre la asignación municipal per cápita y los indicadores asociados que el usuario puede escoger en el menú de interacción del aplicativo (ver Figura 3). Se considera que el nivel observado en los indicadores se atribuye fundamentalmente a una combinación desconocida de las

características más estructurales de los municipios, su desempeño institucional y los niveles de inversión aplicados.

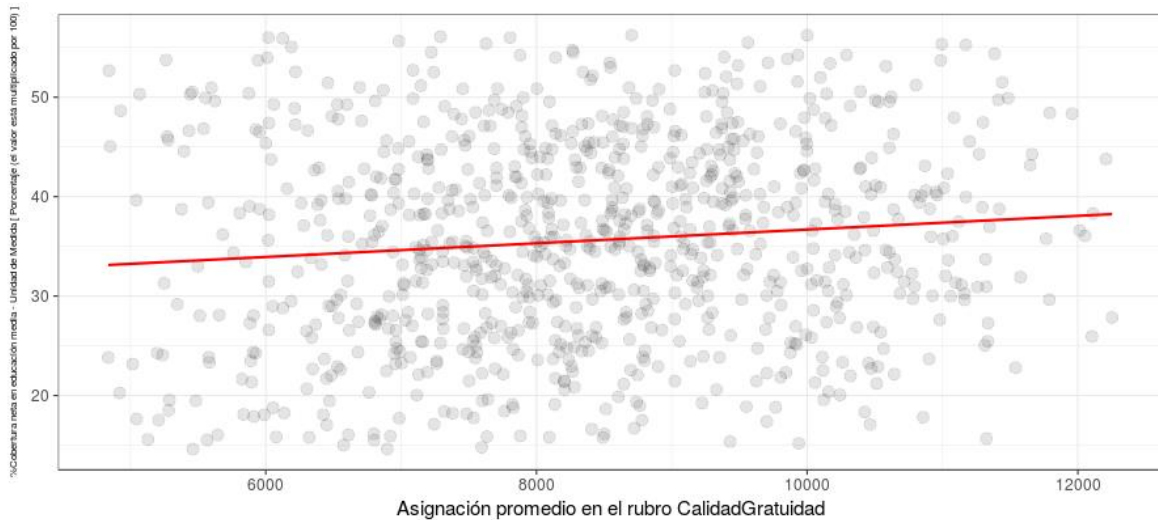


Figura 5: Asignación municipal per cápita del indicador seleccionado.

La segunda sección del aplicativo “Segmentación de la población de estudio” presenta la estratificación de la población de municipios bajo estudio a partir de los resultados de un modelo de segmentación. Se desarrolla la técnica utilizada y la conformación de los clústeres resultantes

La Figura 6 presenta el funcionamiento de las transferencias del SGP de participaciones como un sistema de asignaciones condicionales. En particular, la analítica presentada es un diagrama de cajas y bigotes entre el monto de la inversión per cápita y el índice de tipología de la DDT. Para una interpretación más intuitiva, se construyeron clústeres municipales por cuartiles en los niveles de inversión. Los diagramas de cajas informan sobre el desplazamiento de la distribución condicional de las inversiones con las características de los municipios. La relación observada tiene una interpretación directa en términos de la progresividad en la focalización del gasto.

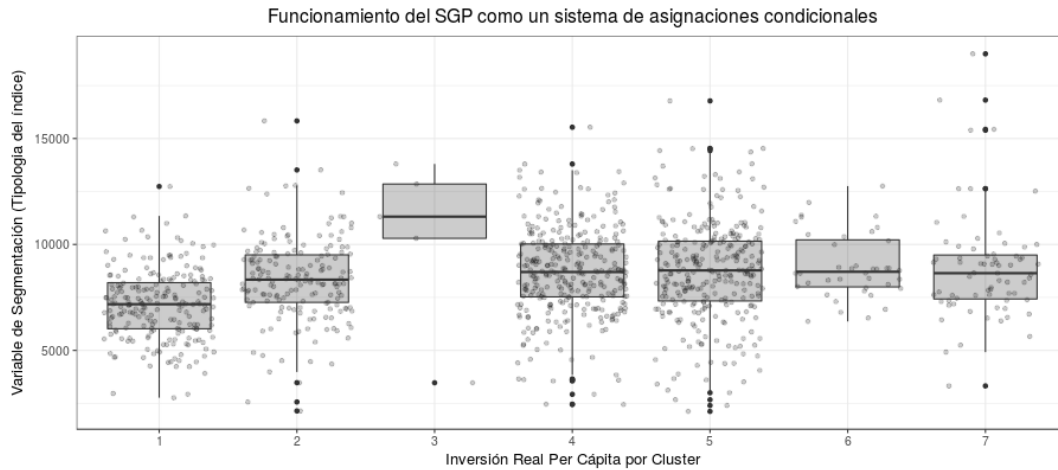


Figura 6: Diagrama de cajas y bigotes entre el monto de la inversión per cápita y el índice de tipología de la DDT.

La Figura 7 presenta el menú de interacción de la segunda sección del aplicativo, que consta de un botón (óvalo rojo) de que permite desplegar el porcentaje de municipios a ser reportado en la tabla clasificación (Ver óvalo Morado).

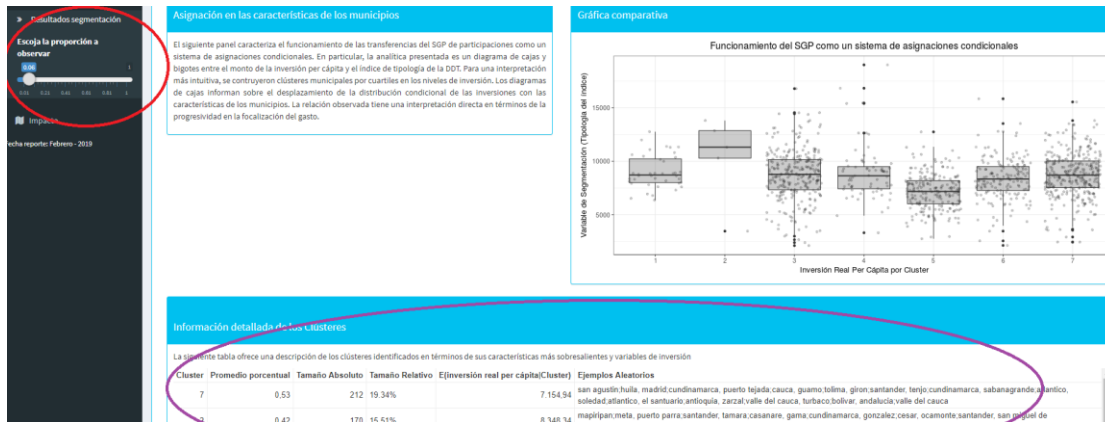


Figura 7: Sistema de consulta - Sección 2 aplicativo.

En la tercera sección del aplicativo, “estimación del impacto”, se reportan las estimaciones del impacto de la política de transferencias del Estado por medio de regresiones por clúster. En la Figura 8 se reportan los valores asociados a la contribución marginal de la financiación real sobre el comportamiento del indicador estimados dentro de un esquema de regresión lineal.

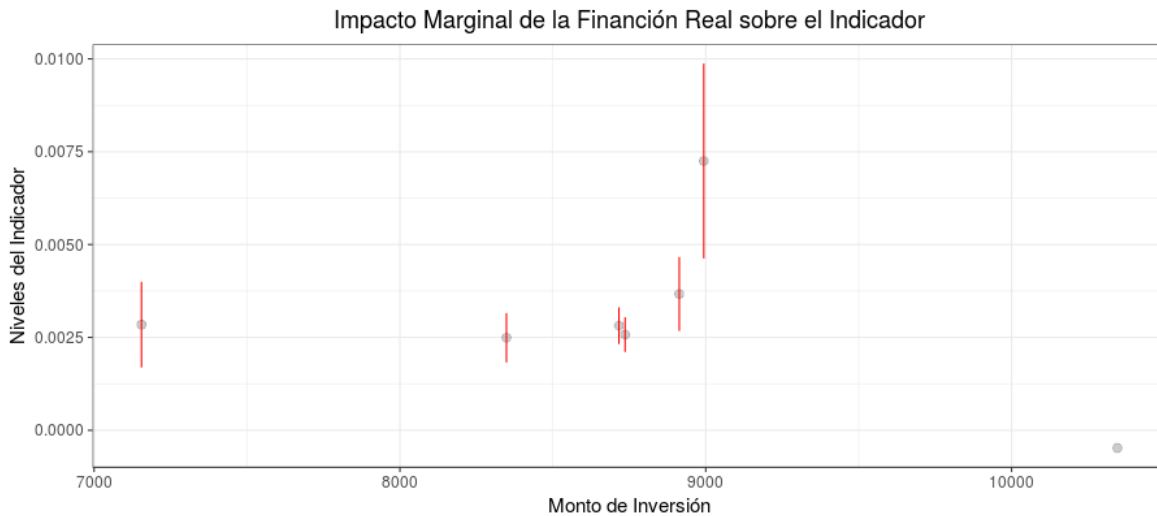


Figura 8: Contribución marginal de la financiación real sobre el comportamiento del indicador estimado.

Todos los efectos marginales reportados están contralados por las principales características de los municipios.

Conclusiones y recomendaciones

1. Los métodos de emparejamiento (PSM) pueden ser empleados para realizar el balanceo de las transferencias del SGP de participaciones como un sistema de asignaciones condicionales de los municipios según sus niveles de inversión en los sectores económicos considerados en el presente estudio.
2. Actualmente los valores asociados a la contribución marginal de la financiación real sobre el comportamiento del indicador estimados dentro de un esquema de regresión lineal permiten caracterizar de forma aproximada el comportamiento general entre estas medidas que funciona adecuadamente si la clasificación de municipios es conveniente.

Socialización

Este tablero de control ha sido socializado con la subdirectora de la Subdirección de Ordenamiento y Desarrollo Territorial con el objetivo de ayudar en la formulación y seguimiento de políticas, planes, programas, proyectos, orientados a fortalecer el ordenamiento, el desarrollo y la planificación territorial, según las disposiciones legales, normativas, institucionales y los criterios técnicos relacionados.