Anexo 6. Informe de análisis de los registros institucionales

Evaluación del efecto de la campaña #NoEsDeHombres

Con el objetivo de disminuir las prácticas de abuso sexual que sufren a diario miles de mujeres en el transporte público (ONU Mujeres, 2017), a finales de marzo de 2017 ONU Mujeres, en colaboración con el Gobierno de la Ciudad de México, anunció la campaña #NoEsDeHombres. El proyecto constó de la instalación de materiales gráficos en diferentes espacios de transporte público. En total se instalaron 657 carteles en 73 estaciones distintas. El principal objetivo de la campaña era disuadir a los hombres de ser perpetradores de dichos abusos.

En el presente trabajo, Data Cívica realiza una estimación del impacto de la campaña #NoEsDeHombres. Se utilizaron los datos de los reportes mensuales de detenidos, disentimientos y quejas por abuso sexual dentro del Sistema de Transporte Colectivo (STC), reportes obtenidos por el Instituto Nacional de Mujeres Ciudad de México mediante solicitudes de información.

Para el análisis utilizamos dos métodos cuasi-experimentales distintos; el método de Diferencias en Diferencias, usado generalmente para analizar el efecto que tuvo un tratamiento o experimento a través del tiempo. Intuitivamente, este método nos permiten comparar las diferencias entre dos grupos, uno que recibió el tratamiento y uno que no en dos periodos distintos: antes y después de la intervención o tratamiento (Imai, 2017). También corroboramos la robustez de los resultados haciendo un Coarsened Exact Matching (King, 2010) que busca distinguir el efecto de un tratamiento cuando la asignación de este puede depender y ser modificado sistemáticamente por características de la población tratada.

Los resultados derivados del análisis de Diferencia en Diferencias que realizamos con los datos proporcionados apuntan a que no hay evidencia alguna que nos permita sostener que la campaña #NoEsDeHombres tuvo un en efecto significativo en el número de denuncias promedio por abuso sexual reportadas al interior del STC. Estos resultados se corroborran con el CEM matching, que secunda los resultados de que el efecto de la campaña no es estadísticamente distinto a cero.

El trabajo se divide de la siguiente manera: en la primera sección se dará una breve descripción de la variable dependiente y la información que se puede extraer de los reportes mensuales de abuso; en la segunda sección damos paso al análisis descriptivo de los datos; por último, en la tercera sección realizáremos la estimación de impacto de la campaña y de las conclusiones puntuales del proyecto.

1. Variable dependiente y potenciales mecanismos causales

El objetivo central de la campaña **#NoEsDeHombres** fue disminuir el acoso del que son víctimas las mujeres al usar el metro en la Ciudad de México. La campaña estaba específicamente dirigia a hombres; es decir, buscaba explícitamente que hubiera menos acoso no necesariamente que las mujeres identificaran mejor cuando eran víctimas de este. Es importante hacer una distinción entre la magnitud del acoso y el total de casos *reportados* de acoso para entender el mecanismo causal que ataría teorícamente a la campaña con una disminución de acoso. Los datos con los que contamos son sobre los segundos, es decir *reportes* no sobre los primeros, es decir *instancias de acoso* y es importante pensar bajo que circunstancias un acoso queda registrado.

Para poder hacer una evaluación de la campaña usando datos de *reportes de acoso* tenemos que hacer un supuesto fundamental: que la magnitud de reportes está directamente relacionada con la magnitud de acoso que se vive. Esto no implica que para poder evaluar la política un requisito indispensable es que cada acoso quede registrado, simplemente quiere decir que estamos asumiendo que variaciones en el primero implican variaciones en el segundo. En todo caso, lo que es observable es el efecto de una campaña sobre el número de reportes.

Habiendo dicho eso, podemos pensar en un mecanismo causal que conecte la implementación de esta campaña con cambios en la candidad de denuncias de acoso registradas. Se esperaría que hombres que en circunstancias normales serían perpetradores, al ser expuestos a esta publicidad, cayeran en cuenta que sus acciones son dañinas, que su conducta está mal o que es

ilegal y se disuadieran de cometer un acto de acoso. Esto resultaría en una disminucíon significativa de reportes de acoso.

Alternativamente, los datos podrían no estar siendo un reflejo fiel del fenómeno subyacente de acoso y simplemente estar midiendo cambios en la propensión de las mujeres a denunciar. La campaña podría tener un efecto sobre las *víctimas*, que al ser expuestas a los anuncios, se enteran de que lo que vivieron es un delito o se sienten más respaldadas para ir a denunciar el acoso ante una autoridad. Si este fuera el caso, lo que esperaríamos observar es un aumento en los reportes registrados.

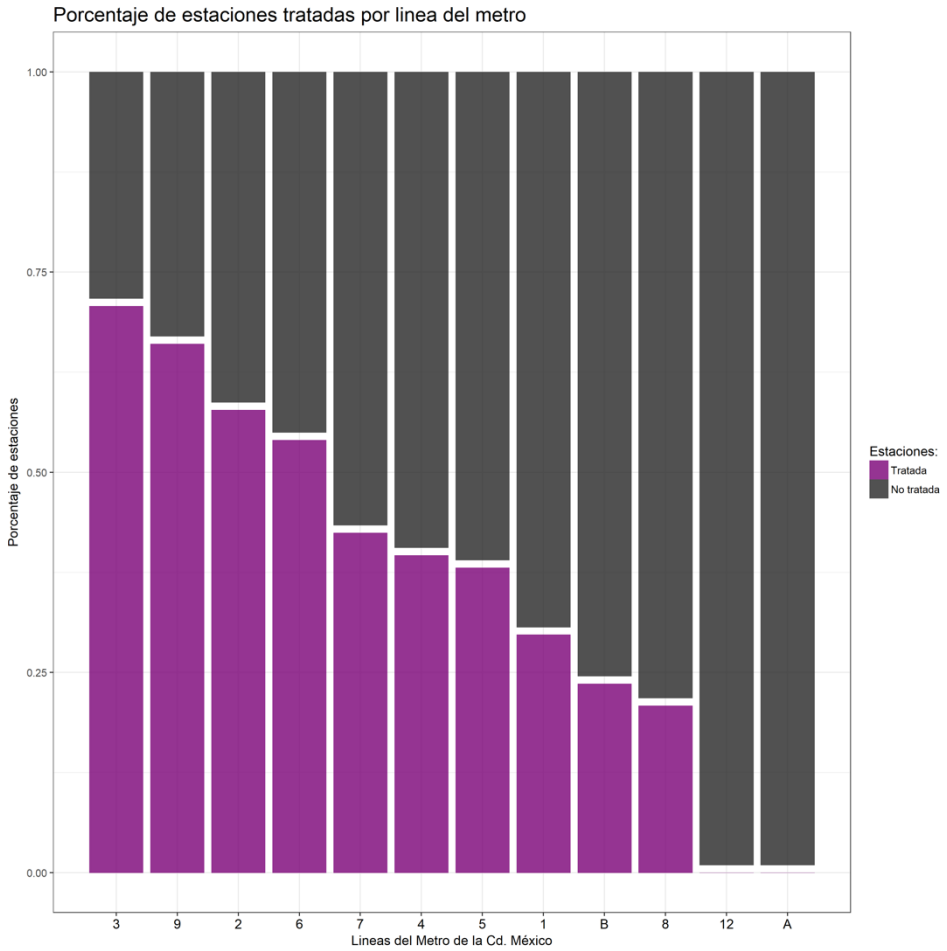
2. Descripción de los datos

Los datos de los reportes de abuso los extrajimos de documentos en formato PDF proporcionados por ONU Mujeres y los colocamos en una base de datos que enumeraba los destinos de los reportes de abuso (queja, juzgado cívico, carpeta de investigación o disentimiento) por mes y por estación. A dicha base de datos le agregamos también la fecha, estación y lugar (barandal, columna, gran muro, anden, banca o trabe) en la que se instalaron los carteles.

A continuación incluimos una descripción gráfica de los datos.

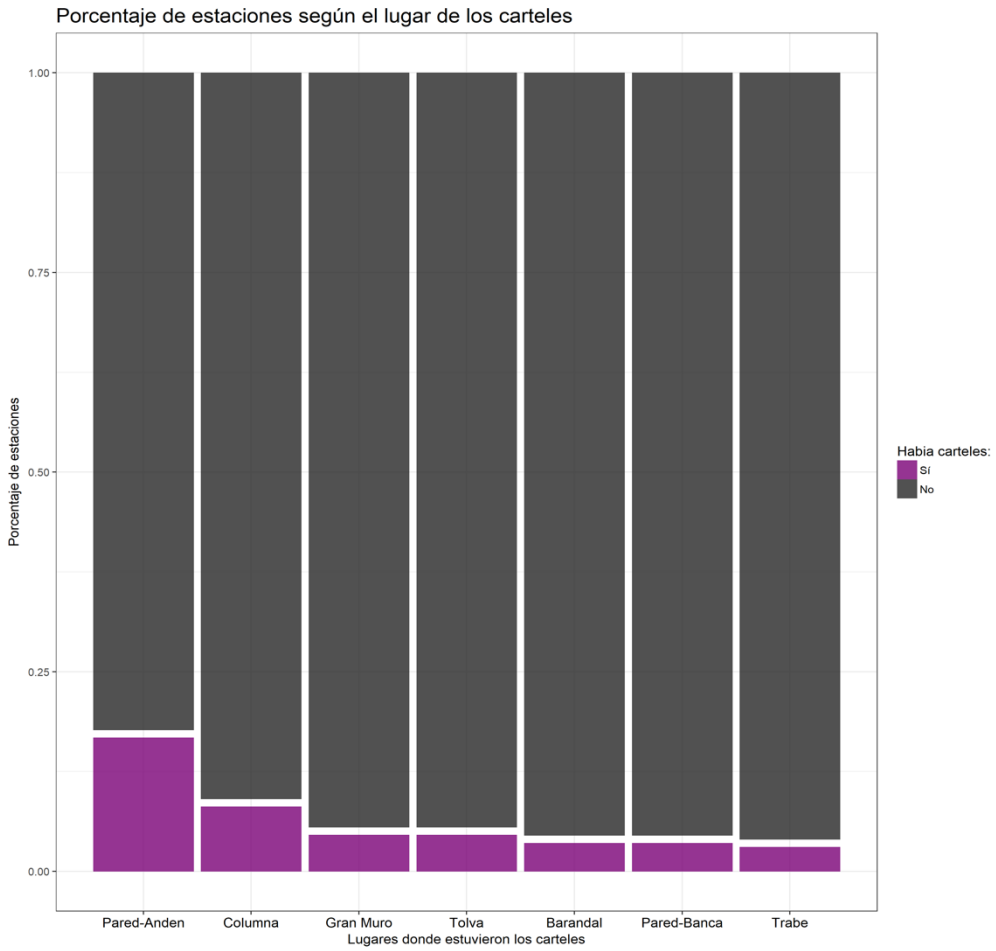
La campaña se aplicó en el 37.5% de las estaciones de metro de la CDMX. En la siguiente gráfica podemos ver su distribución por línea.

2.1 Distribución de estaciones tratadas por línea



Tres de las 10 líneas tratadas (líneas 3 ,9 y 2) cuentan en la mayoría de sus estaciones con al menos un cartel. La Línea 8 es la línea con menor porentaje de carteles, mientras que la Línea 12 y la Línea A no contaron con ningún cartel.

2.2 Por tipo de tratamiento

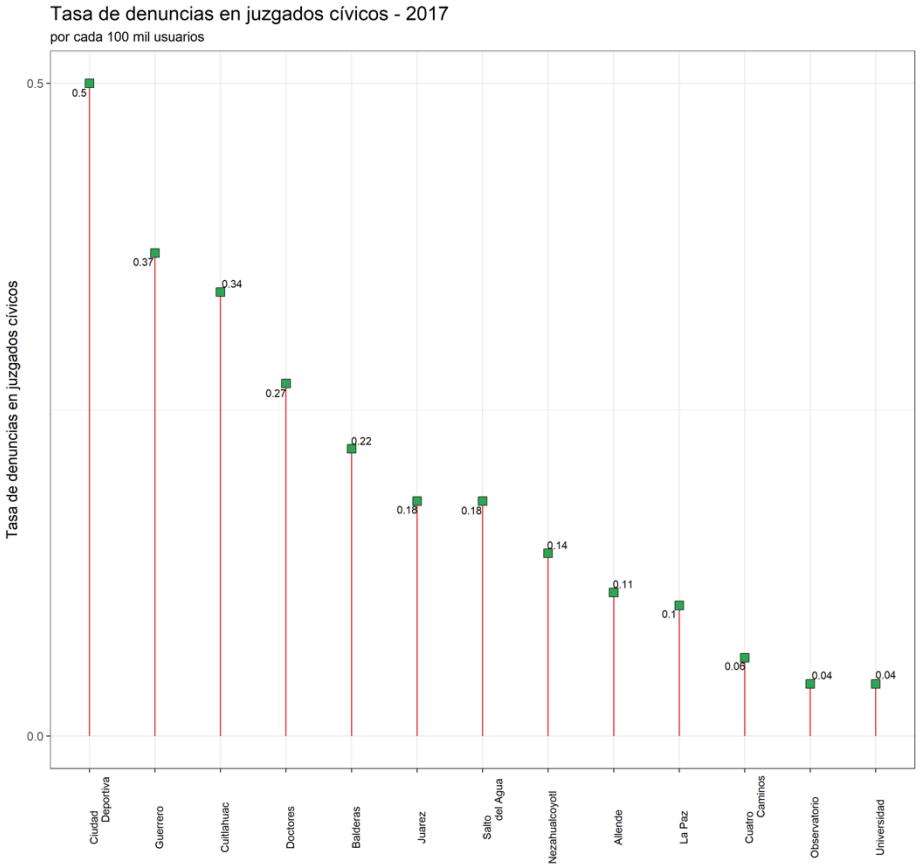


2.3 Estaciones con mayor tasa de abuso

Un punto de arranque para analizar el comportamiento de los datos es revisar cuáles son las estaciones en las que se presentan más denuncias, ya sea que terminen en el juzgado cívico, queja, carpeta de investigación o desistimiento. Esto con la intención de verificar la intensidad con la que suceden las denuncias en cada estación y no dejarnos llevar únicamente por el tamaño de la afluencia, ya que es lógico pensar que mientras mayor afluencia haya en una estación, mayor será el número de denuncias que ahí se reporten, no necesariamente porque se trate de estaciones más peligrosas. Al usar tasas, nos aseguramos de estar comparando la intensidad de un fenómeno entre estaciones, no el total de personas que por ahí transiten.

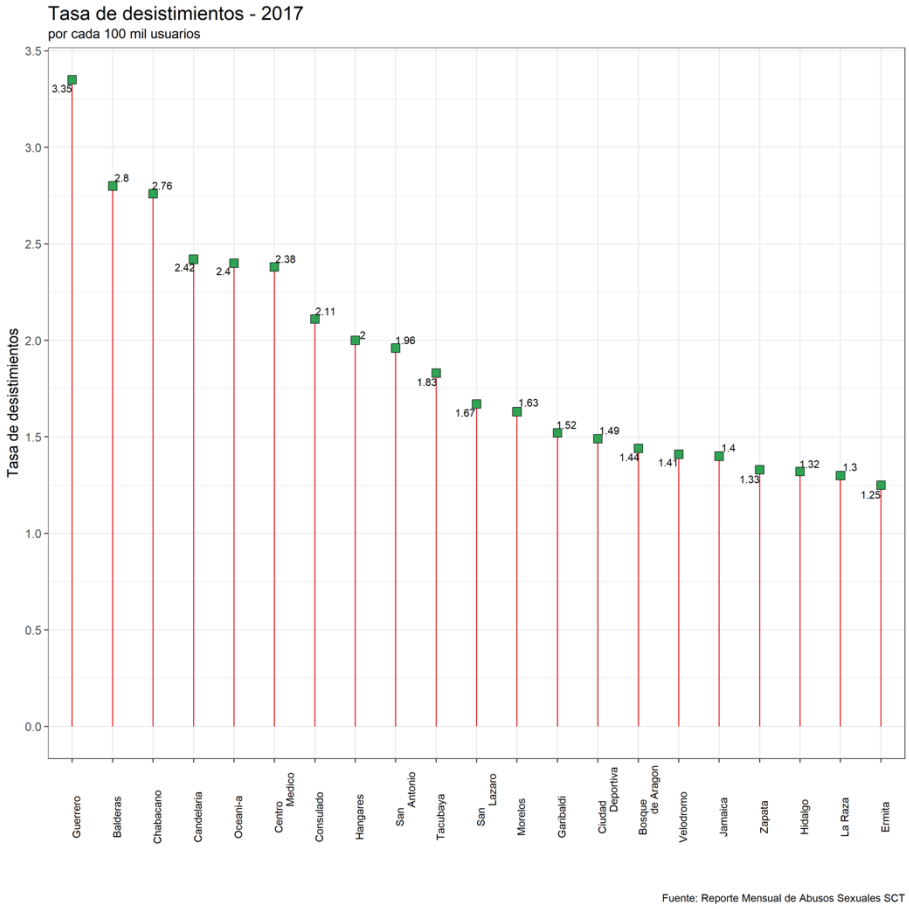
Las estaciones con tasas más altas de eventos denunciados ante juzgado cívico son Ciudad Deportiva, Guerrero y Cuitláhuac.

DATACÍVICA
Más datos para más personas

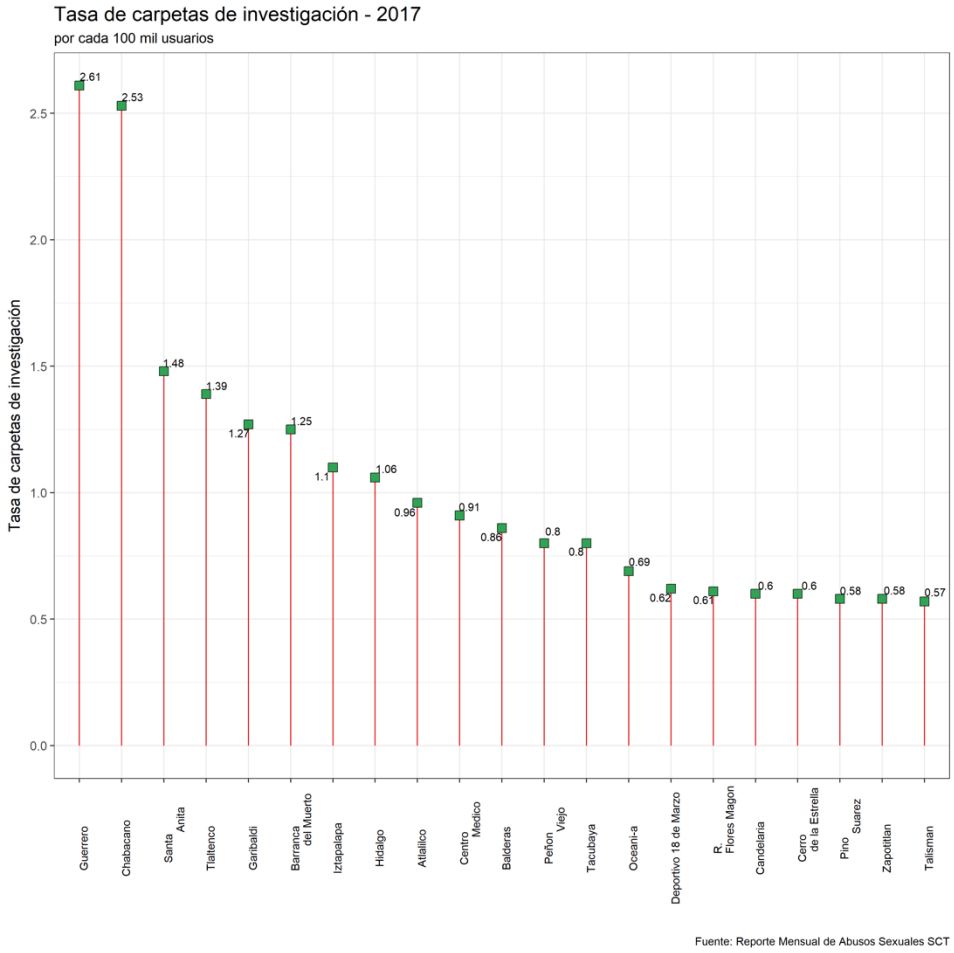


Fuente: Reporte Mensual de Abusos Sexuales SCT

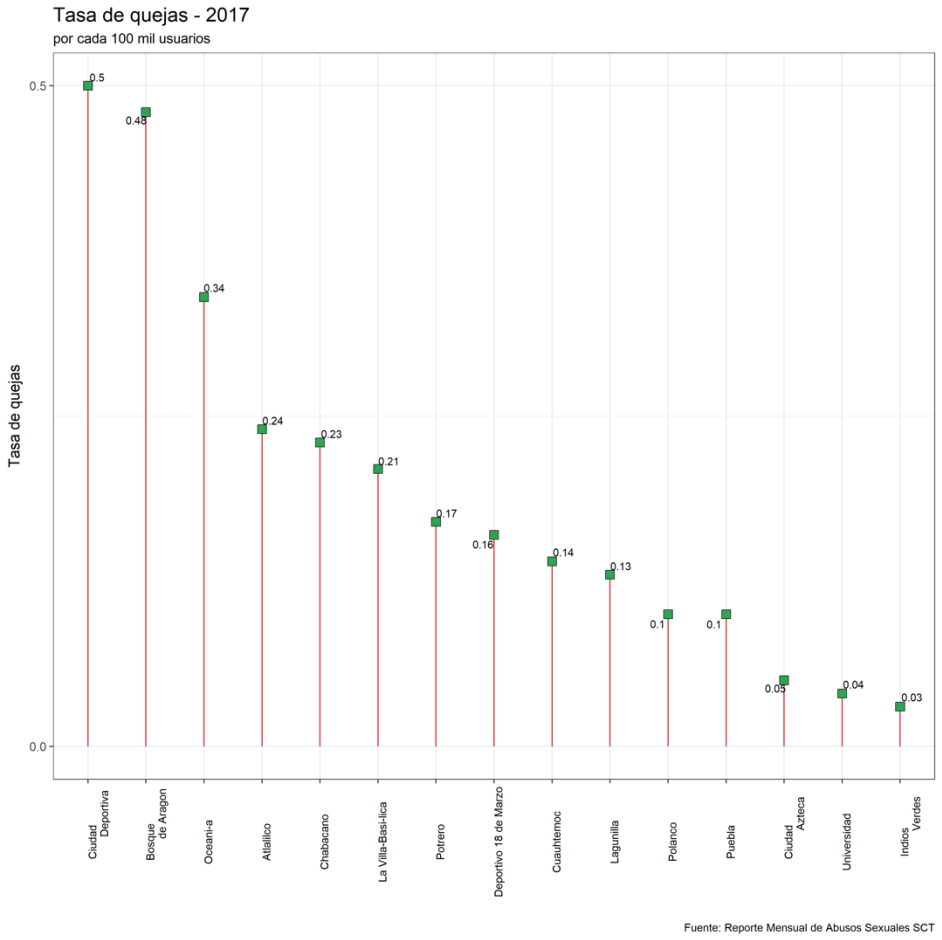
Si ahora vemos la tasa de desistimiento por 100 mil pasajeros, vemos que otra vez las estaciones de Guerrero y Chabacano están entre las estaciones con tasas más altas de, seguidas de la estación Balderas.



Por su parte las estaciones de Guerrero, Chabacano, Santa Anita y Tlatelolco son las que tienen mayores tasas de carpetas de investigación por abusos abiertas por cada 100mil pasajeros.

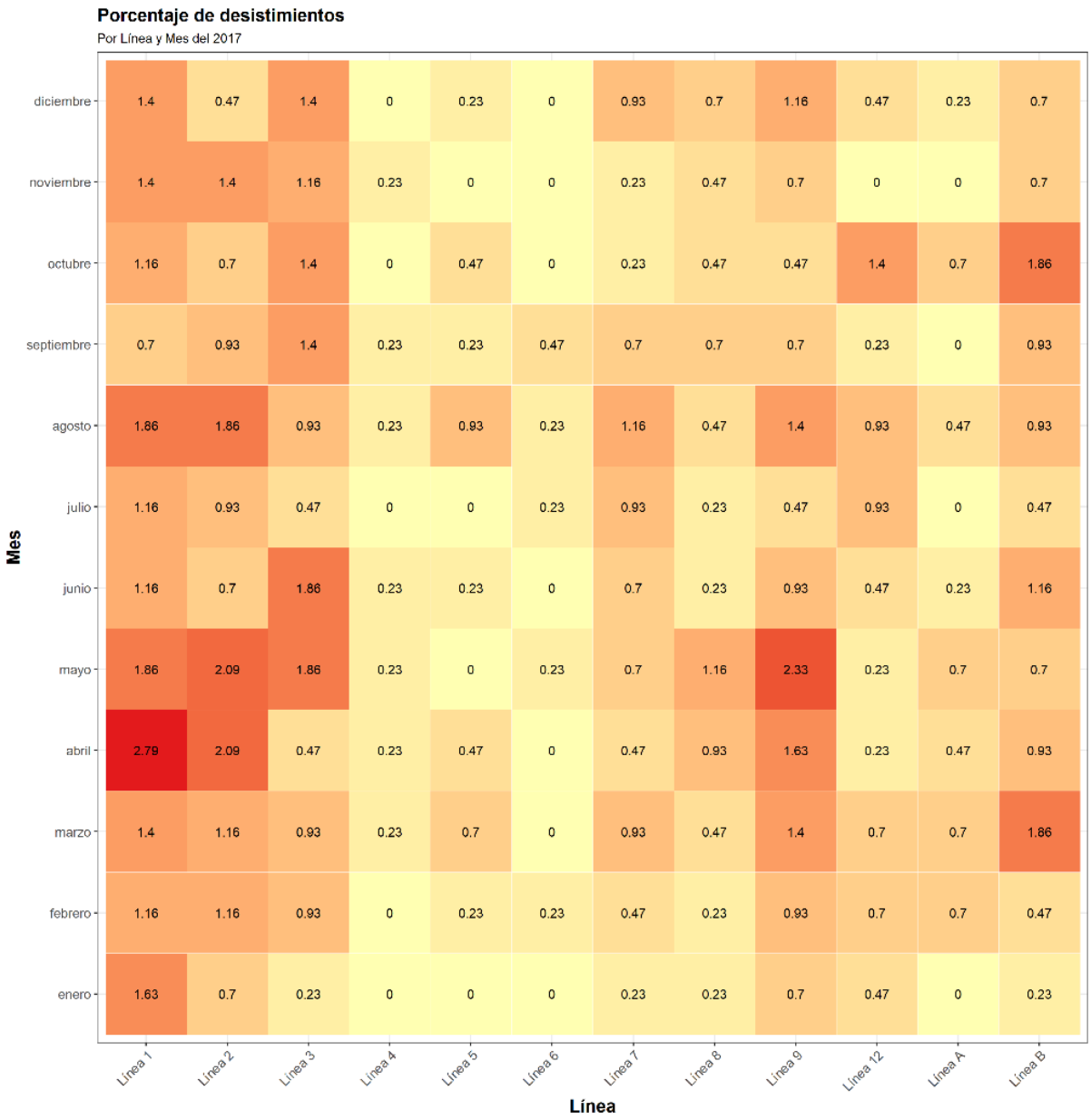


Por último, las estaciones Ciudad Deportiva y Bosque de Aragón son las que presentan tasas más altas de quejas.

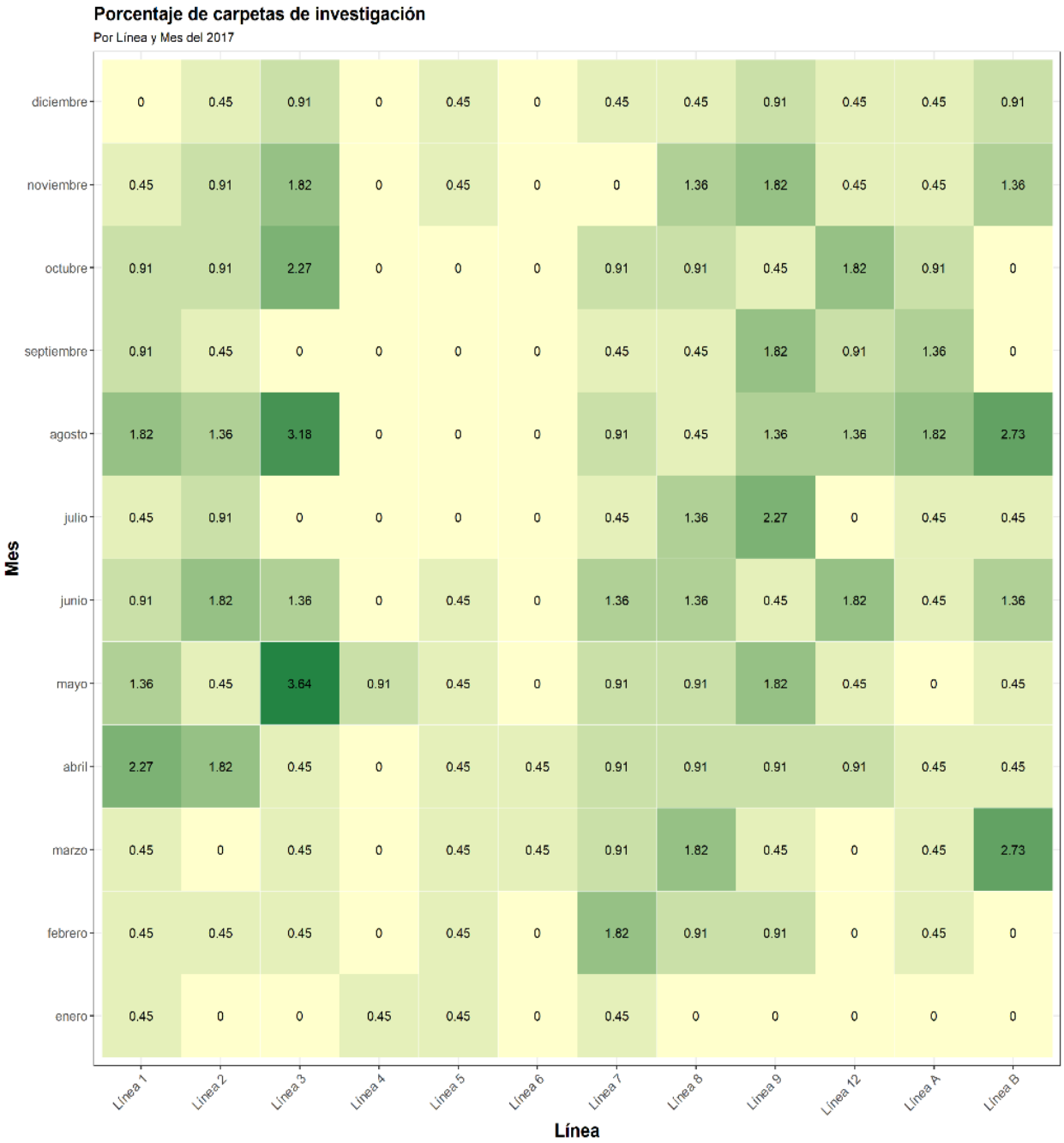


2.4 ¿Cuándo y dónde sucedieron los abusos?

Para analizar los meses y las estaciones en las que se distribuyen los mayores porcentajes de denuncias por abuso, tomamos en cuenta los disentimientos y las carpetas de investigación, ya que son los dos destinos de denuncias por abusos más recurrentes.



Para la ubicación, los porcentajes de desistimientos se acumulan tanto en las primeras tres líneas del metro, como en la Línea 9 y la Línea B. Para los meses de ocurrencia de estas denuncias, parece que no hay un patrón mensual, parece que hay denuncias que terminan en desistimientos la mayor parte del año. La observación temporal es importante ya que es una primera aproximación de que no hay una variación en la ocurrencia de denuncias después del final de la campaña.



En cuanto a las carpetas de investigación vemos que aparecen con frecuencia en prácticamente todas las líneas excepto en las líneas 5 y 6. La distribución mensual de ocurrencia, al igual que en el caso de desistimiento, no sigue ningún patrón y una vez más parece estar presente la mayor parte del año.

3. Estimación del efecto de la campaña

3.1 Método de Diferencias en Diferencias

Para la estimación del efecto de la campaña #NoEsDeHombres utilizaremos primero el método de Diferencias en Diferencias (DD). Este método es usado comúnmente para la estimación del efecto de una intervención específica o tratamiento (en este caso la campaña #NoEsDeHombres) comparando los cambios en los resultados a través del tiempo entre la población que tuvo un impacto del programa —grupo de tratamiento— y la población que no lo tuvo —grupo de control—(Imai, 2017). El método reconoce que en ausencia de una asignación aleatoria del tratamiento, los grupos de control y tratamiento pueden ser diferentes por muchísimas razones (Angrist y Pischke, 2015). Sin embargo, en algunas ocasiones, antes del tratamiento, los grupos de control y tratamiento pueden tener tendencias paralelas. Cuando esto sucede, la divergencia entre la tendencia de ambos grupos después de que el tratamiento ocurre puede vislumbrarse como el efecto del tratamiento.

El método nos permite, a diferencia de otros métodos para verificar causalidad, no tener que conocer la asignación del tratamiento, es decir, nuestras conclusiones pueden ser valiosas aunque la asignación del tratamiento no haya sido aleatoria o nos sea desconocida. Lo anterior dado que es un método con un diseño cuasi-experimental en el que el uso de datos longitudinales de los grupos de control y tratamiento nos arroja un contra factual apropiado para identificar el efecto del tratamiento.

La estimación del efecto del tratamiento en el método de DD deriva de una regresión de Mínimos Cuadrados Ordinarios con una interacción. La interacción que nosotros utilizamos es la siguiente:

$$Y_{ij} = \beta_0 + \beta_1 \, TREAT_i + \beta_2 \, POST_j + \beta_3 \, TREAT_i.POST_j + \beta_4 \, AFLUENCIA + \beta_5 CARTELES + \beta_6 \, CORRESPONDENCIA$$

En donde Y toma los cuatro destinos de denuncias diferentes (queja, desistimiento, carpetas de investigación o juzgado) según sea el caso. TREAT es una variable binomial que controla por

las diferencias fijas entre las estaciones con carteles (grupo de tratamiento) y sin carteles (grupo de control); toma el valor de 1 en las estaciones con carteles y 0 si no tienen carteles. POST es una también una variable binomial pero que controla por las variaciones a lo largo del tiempo; toma el valor de 1 después del tratamiento y 0 en otro caso. TREAT.POST es la multiplicación de las dos variables anteriores y su coeficiente es el estimador de Diferencias en Diferencias. Los covariantes de AFLUENCIA, CARTELES y CORRESPONDENCIA son variables de control por el número de personas promedio en cada estación, el número de carteles y el número de correspondencias de cada estación respectivamente.

3.1.1 Problema de identificación

El caso de análisis cuenta con un problema de identificación que tiene que ver con uno de los supuestos clave del método DD: tendencias paralelas. Al desconocer los criterios bajo los cuales se eligieron las estaciones para colocar los carteles, puede ser que estaciones con y sin carteles no sean comparables entre sí y cualquier cambio en las tendencias de abusos y denuncias se deba a cualquier otra cosa ajena a la campaña. Lo anterior se acentúa una vez que revisamos las tendencias de desistimientos y carpetas de investigación, como veremos más adelante. Al no contar con tendencias paralelas entre los grupos de tratamiento y de control antes del tratamiento, no podemos sostener que un cambio en las tendencias después del tratamiento se deba al tratamiento.

3.1.2 Resultados

Los resultados de las cuatro regresiones se presentan en la Tabla 1.

Table 1: Resultados de la regresión

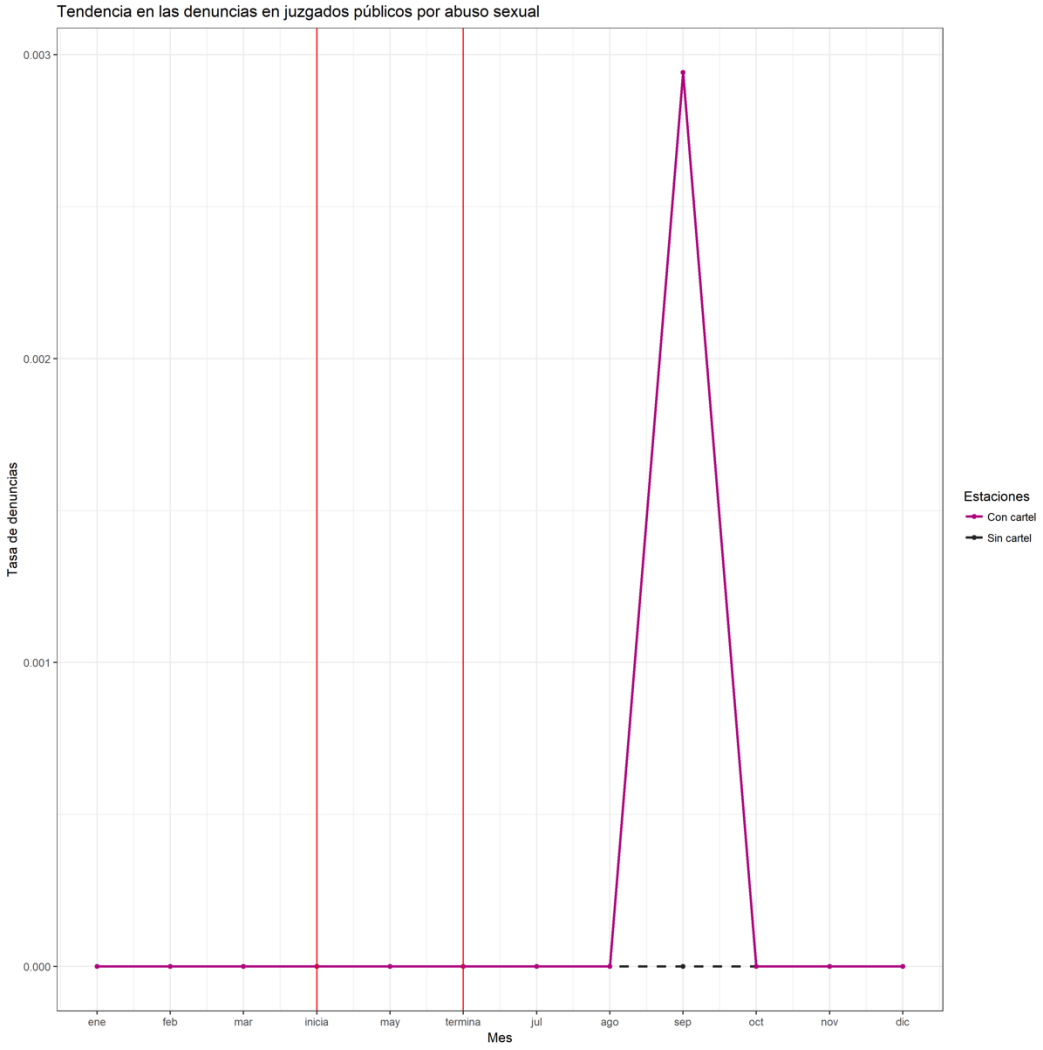
	Variables Dependientes:		
	Denuncia		
	(Carpetas)	(Juzgado)	(Desistimiento)
POST	0.005	0.000	-0.009
	(0.005)	(0.00003)	(0.006)
TREAT	-0.003	0.001	-0.012**
	(0.004)	(0.001)	(0.006)
TREATPOST	-0.007	-0.002	0.021*
	(0.009)	(0.002)	(0.012)
AFLUENCIA	-0.000***	-0.000	-0.00000***
	(0.000)	(0.000)	(0.000)
CARTELES	0.0003	0.00003	-0.00005
	(0.001)	(0.00003)	(0.001)
CORRESPONDENCIA	0.013***	0.001	0.018***
	(0.003)	(0.001)	(0.005)
CONSTANTE	0.009***	0.0004	0.027***
	(0.003)	(0.0004)	(0.006)
Note:			*p<0.1; **p<0.05; ***p<0.01

En el caso de quejas, los datos no fueron suficientemente amplios para que el modelo de regresión fuera estadísticamente significativo. Lo primero que salta a la vista es que el coeficiente de DD únicamente es estadísticamente significativo y positivo en los desistimientos, en las demás formas de denuncias no. Es decir, el efecto real de la campaña es un aumento de 2.1% en los desistimientos. Sin embargo, como veremos más adelante, este resultado no es robusto ya que para el caso de desistimientos no se cumple el supuesto de tendencias paralelas.

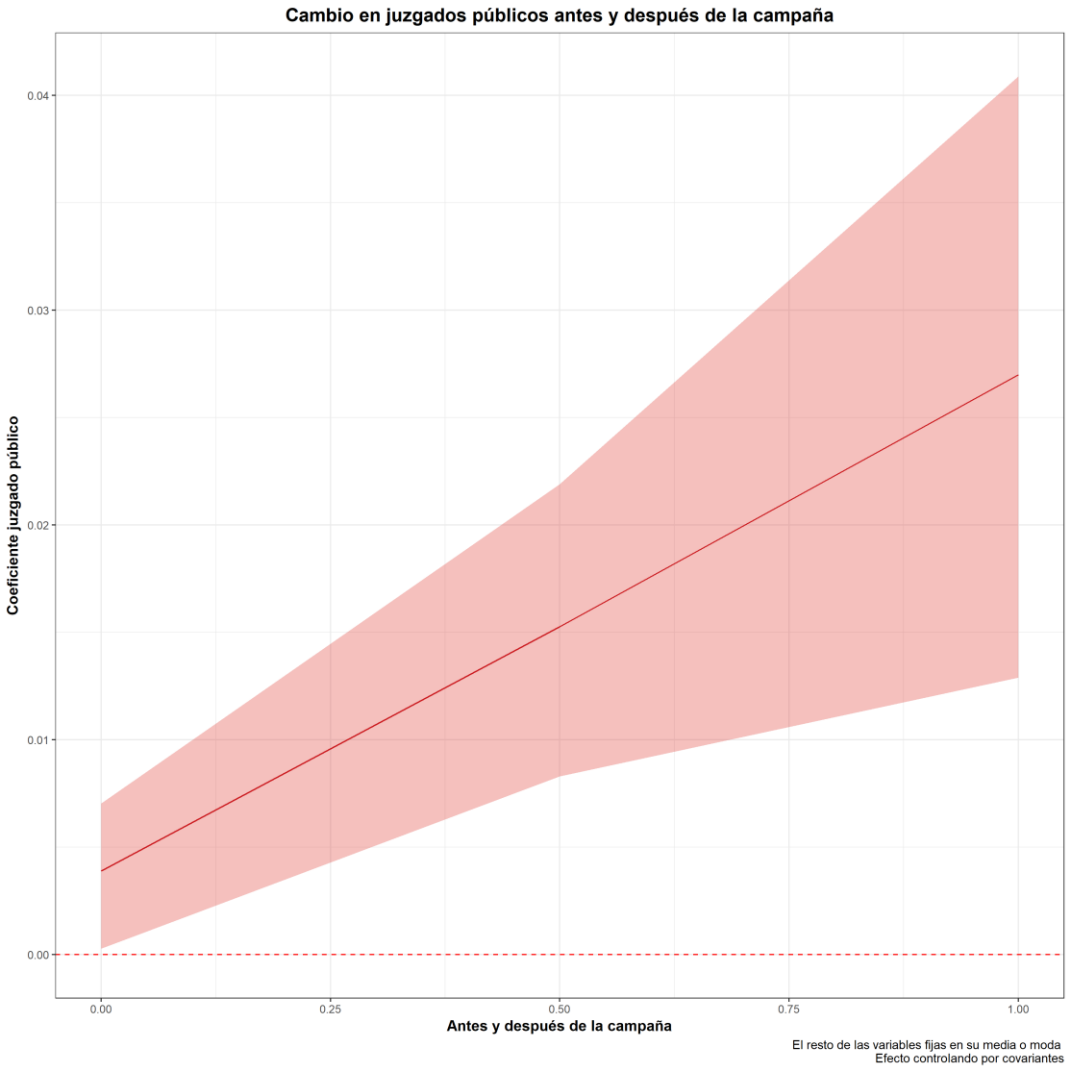
A continuación, mostraremos el comportamiento de las tendencias de cada tipo de destino de denuncia por abuso según su situación (tratados o control) y la incertidumbre de cada uno de los parámetros estimados por la regresión.

3.1.2.1 **Juzgado Público**

En el caso de juzgado cívico vemos que la tasa de denuncias es cercana a cero tanto en estaciones que fueron parte de la campaña como en aquellas donde no se implementó casi en todo el periodo. El único momento en el que la tendencia se despega de cero es en el mes de septiembre en el que la tendencia aumenta.



Revisando la incertidumbre de la estimación del parámetro de juzgado público, vemos que el impacto es positivo y significativo, sin embargo la incertidumbre es grande. Esto se debe, de entrada, a dos cosas: la primera causa es la autocorrelación serial derivada del método DD (Bertrand, Deflo, y Mullainathan, 2004); la segunda es que en el caso específico de juzgados cívicos no parece cumplirse el supuesto de tendencias paralelas entre los grupos de tratamiento

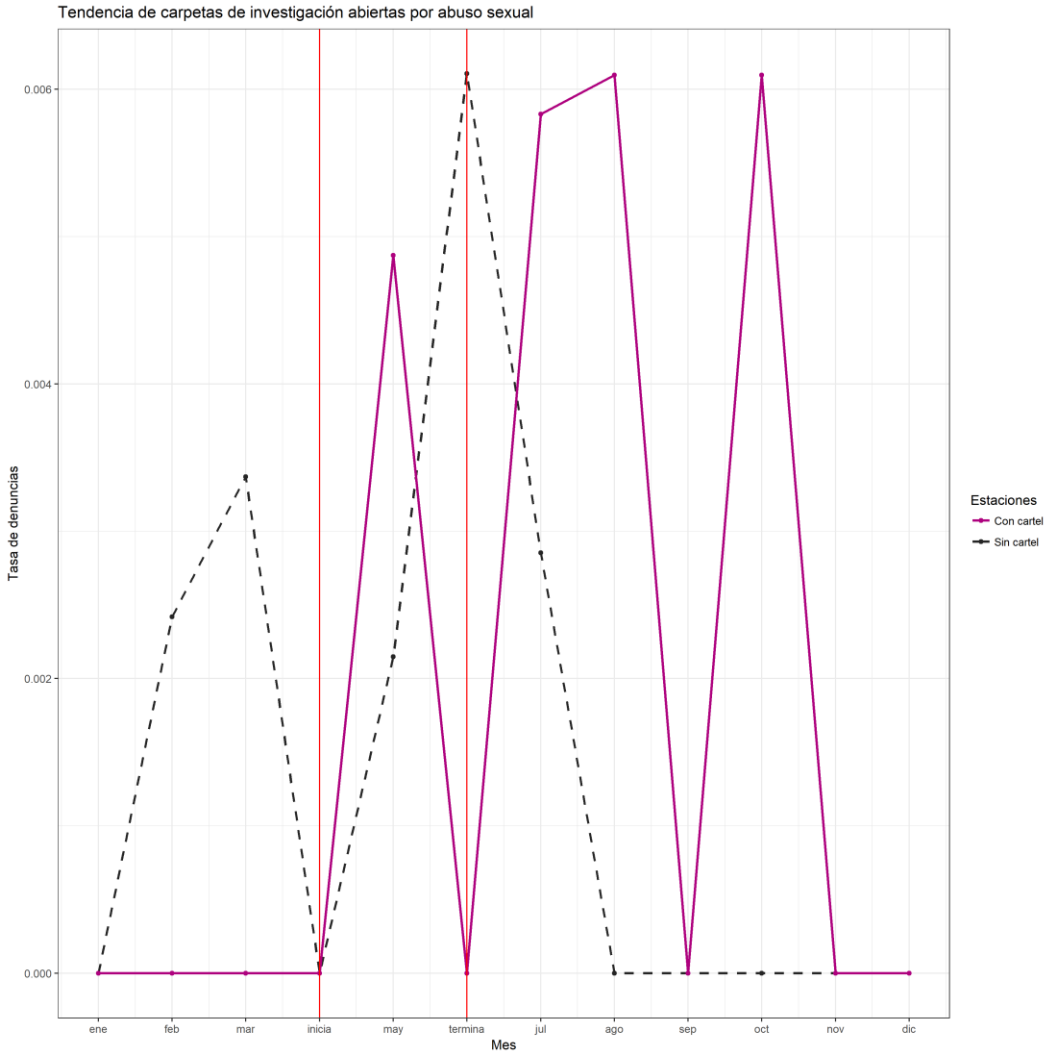


y control, por lo que la conclusión derivada de la estimación del parámetro podría no ser robusta.

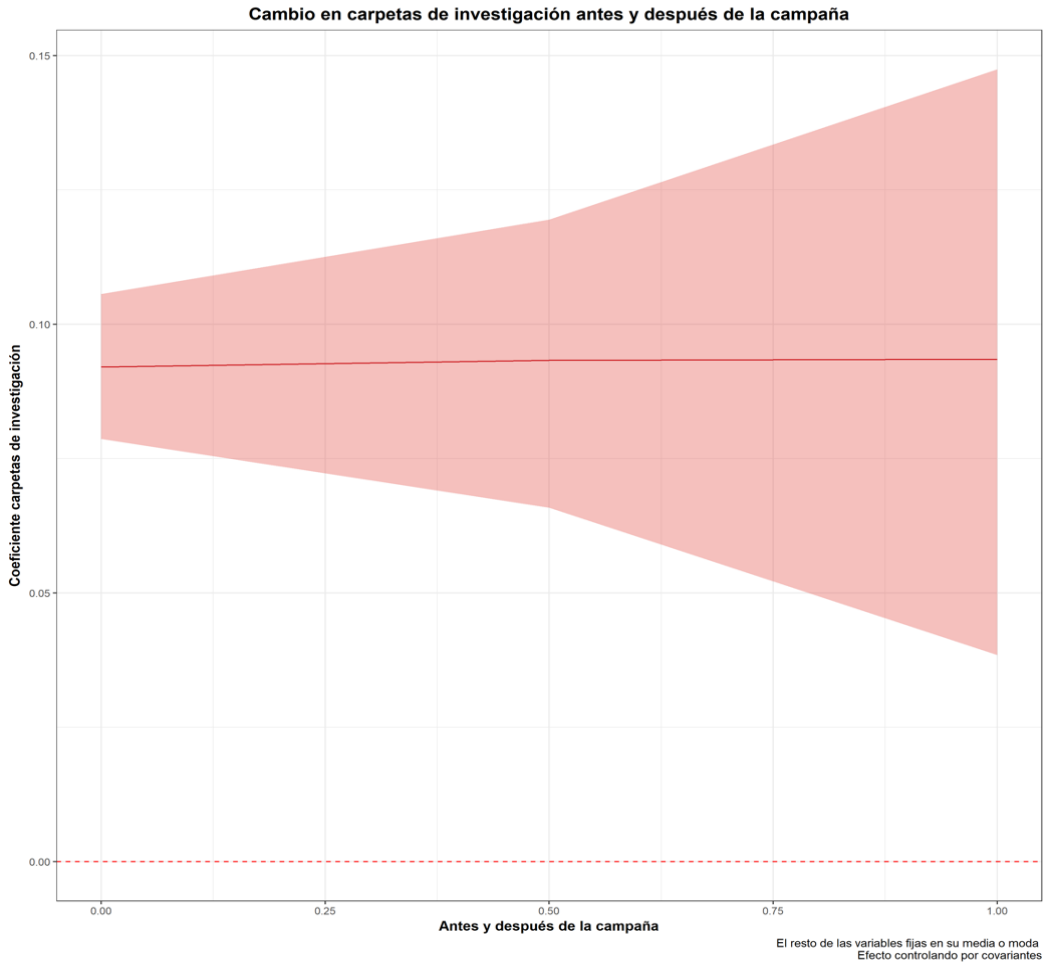
3.1.2.2 Carpetas de Investigación

Lo primero que notamos en el caso de carpetas de investigación es que las tendencias de las estaciones con y sin campañas no son paralelas. Si bien, en esta ocasión sí tenemos varias observaciones de carpetas de investigación, al no seguir las mismas tendencias antes del tratamiento, no podemos inferir que el cambio de tendencias después del tratamiento se deba al tratamiento.

.

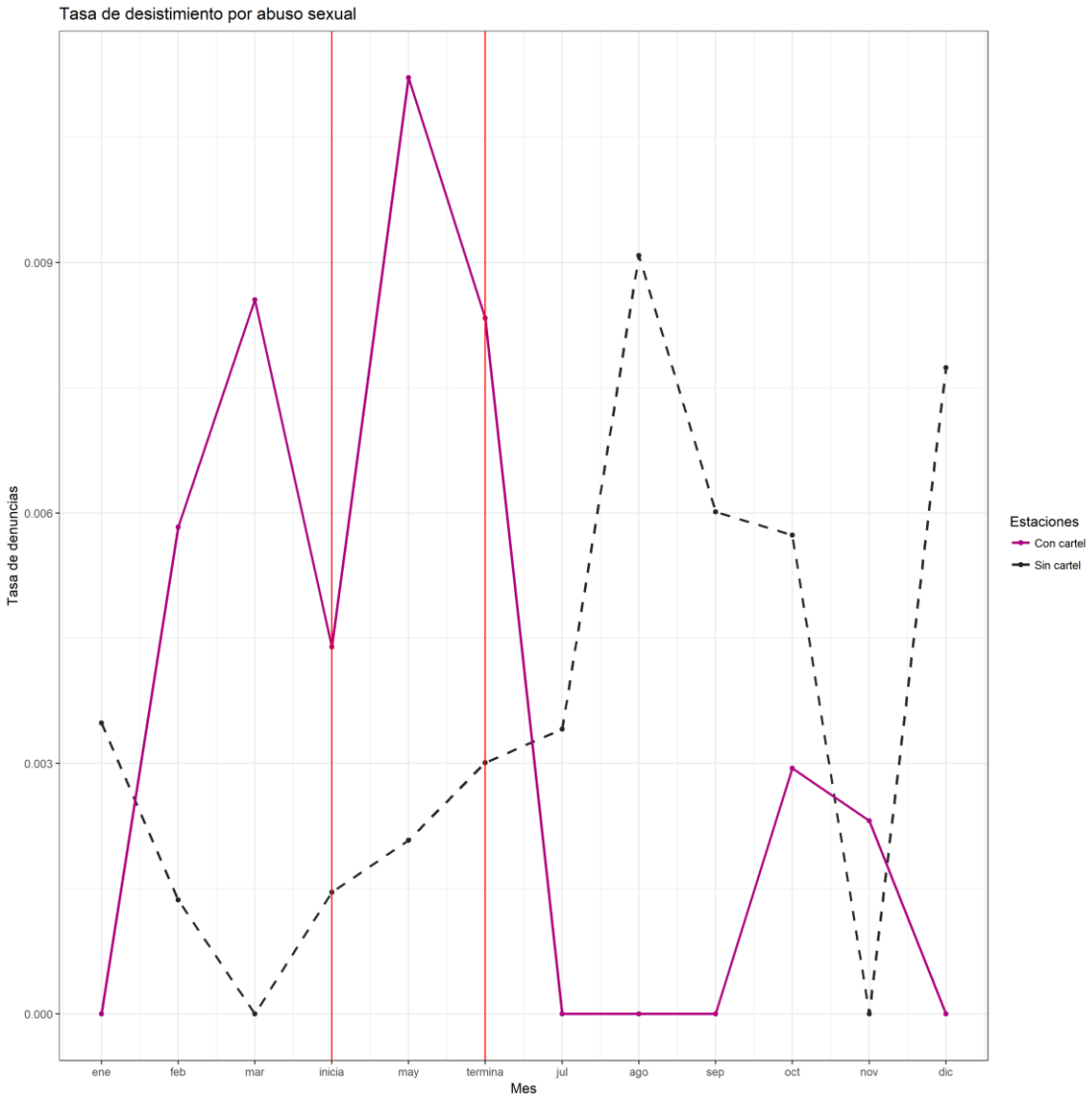


Al revisar la significancia estadística de los cambios en las tendencias después del tratamiento, nos damos cuenta que no es significativo.

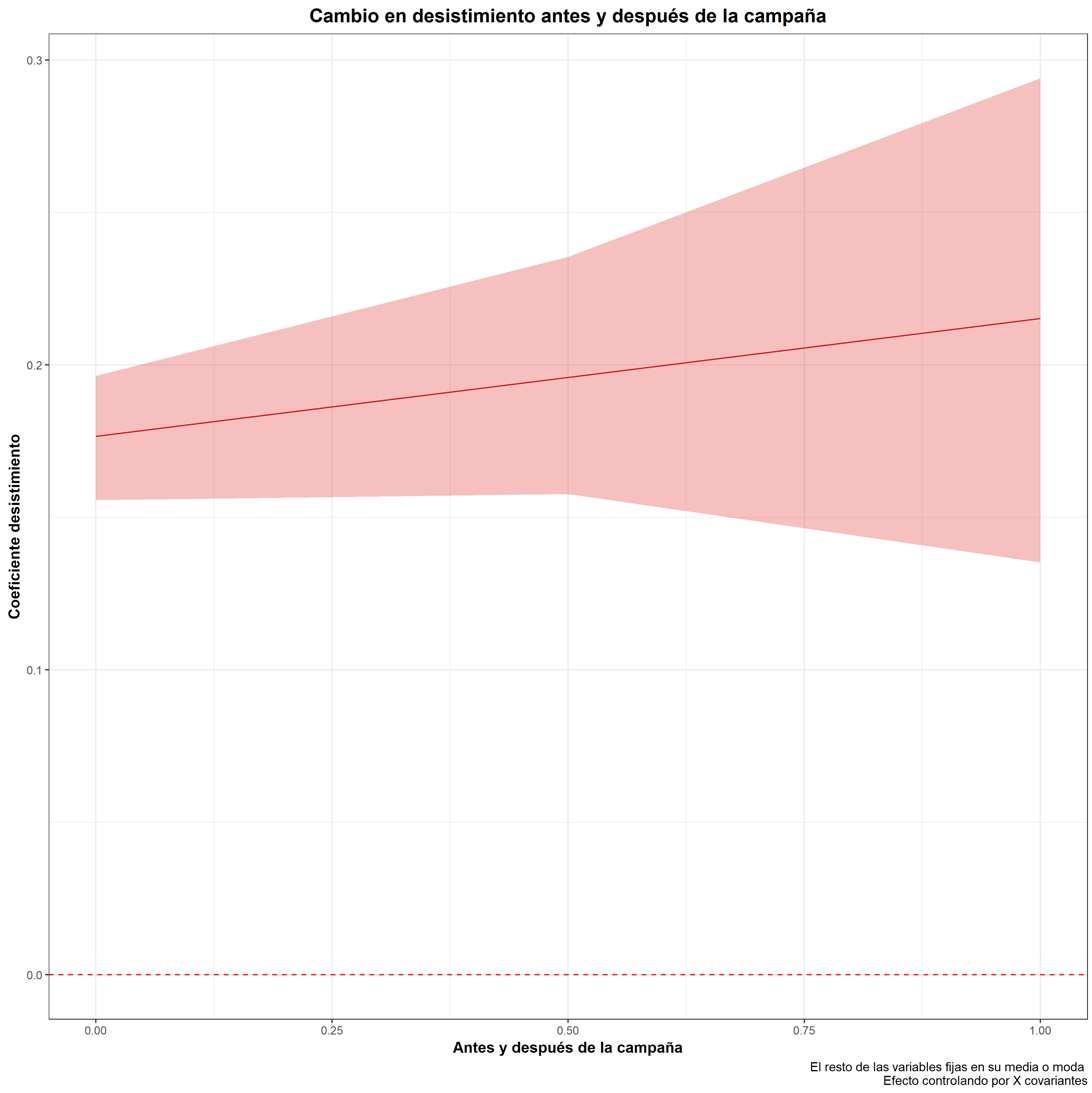


3.1.2.3 Desistimientos

Para desistimientos, vemos que una vez más las tendencias de los dos grupos, con tratamiento y sin tratamiento, no son paralelas. Por lo que en este caso, al igual que en carpetas de investigación, no podemos sostener que un cambio en la tendencia de los desistimientos después del tratamiento se deba únicamente al tratamiento.



Al revisar la magnitud e incertidumbre, vemos que el coeficiente alcanza a apenas a ser significativo.



3.1.3 Conclusiones del análisis con el método Diferencias en Diferencias

1) Los resultados del análisis nos dicen que no contamos con evidencia para sostener que la campaña #NoEsDeHombres haya tenido un efecto en el número de denuncias por abuso sexual en el Sistema de Transporte Colectivo.

2) Parece ser que el efecto del tratamiento sobre desistimiento es positivo, lo cual significaría que de heco aumentó la proporcion de mujeres que desistieron de llevar su

denuncia ante autoridades superiores. Sin embargo, tenemos que tomar este resultado en el contexto del modelo, para empezar sólo es significativo al 90%; además, la trayectoria en el tiempo de desistimiento, no cumple con el supuesto de tendencias paralelas, por lo que no podemos sostener que el cambio positivo se deba exclusivamente a la campaña. Esto es, el cambio se puede deber a variaciones periódicas en los desistimientos ajenas a la campaña; el resultado no es robusto.

3.2 Matching

Dado que en el ejercicio anterior no contamos con el supuesto principal del método DD, otra forma de estimar el efecto del tratamiento es *coarsened matching* (CME). Matching es una técnica que busca parear observaciones que sean idénticas (o suficientemente parecidas entre si) en todos los atributos que podrían modificar el efecto del tratamiento, excepto que una observación sí recibió dicho tratamiento y otra no. Asi podemos estimar el efecto del tratamiento con más certeza. Como señalamos anteriormente, al no tener tendencias parecidas las estaciones que recibieron la campaña con las que no la recibieron lo que nos queda es comparar la estaciones más semejantes en las variables que podrían sesgar la relación entre denuncias y la campaña.

De esta manera, emparejamos a las estaciones en el tiempo **por línea y número de correspondencias**. ¿Por qué por estas dos variables? El número de correspondencias puede afectar el tipo de pasajeros que transitan por cada estación, entonces es una manera de controlar por las características heterogeneas e inobservables de los pasajeros. El número de línea puede servir como un proxy imperfecto de estatus socioeconómico del area en la que se encuentra la estación.

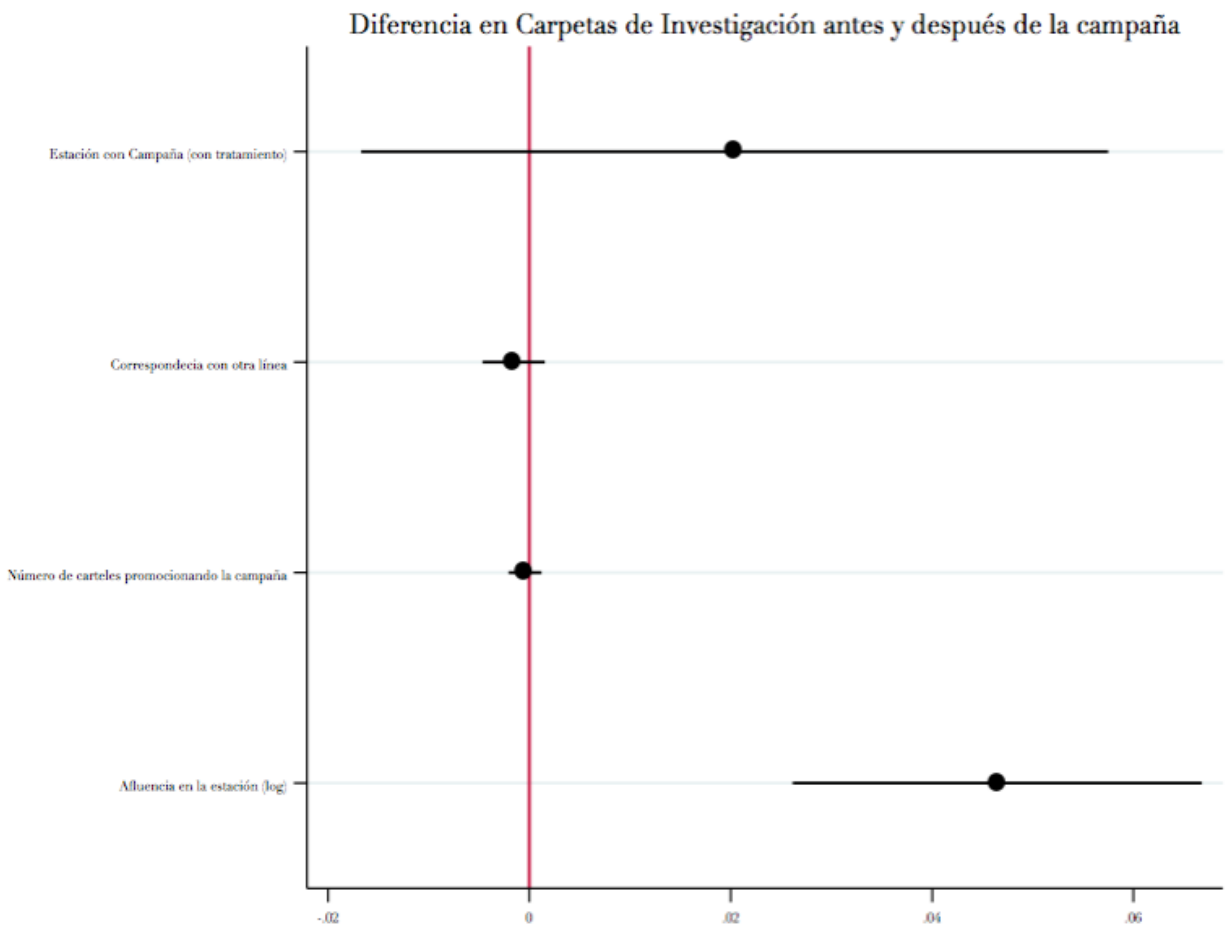
Esta técnica de emparejar a las estaciones con su contrafactual más cercano se le llama *coarsened matching*. El desbalance antes de hacer el pareo, es decir qué tanto difieren las observaciones del grupo de tratamiento y del grupo que no recibió el tratamiento a lo largo de estas dos variables es de .36, es decir, el número de linea y el número de correspondencias no se distribuían

aleatoriamente entre estaciones con y sin la campaña. Después de emparejar las observaciones la diferencia se redujo a casi 0. Esto es, las estaciones tratadas y de control son casi idénticas una vez que ponderamos por línea y número de correspondencias. En el proceso de ponderación se perdieron 9 observaciones tratadas y ninguna sin tratar.

3.2.2 Resultados

Una vez hecho el emparejamiento, corremos una regresión de mínimos cuadrados en la que la variable dependiente sea el cambio en las tasas según cada denuncia antes y después de la campaña. Para las diferencias en las tasas de juzgados, desistimiento y quejas, el valor F nos indica que ninguno de los anteriores modelos es significativo, es decir nuestras variables independientes no están explicando a las dependientes. El modelo planteado sí es significativo para el caso de carpetas de investigación.

Los resultados se esbozan en la siguiente gráfica:



Vemos que no existe una diferencia significativa estadísticamente en el número de denuncias entre las estaciones con y sin carteles de la campaña #NoEsDeHombres, aun cuando lo único que diferencie a una estación de otra sea tener carteles con la campaña. Por otro lado, lo que mejor explica la diferencia en la tasa de carpetas de investigación antes y después de la campaña es la afluencia.

3.2.3 Conclusiones del análisis con el método de Coarsened Exact Matching

El método CEM nos permitió llegar mediante un proceso distinto a las mismas conclusiones. Una vez que pareamos las observaciones más parecidas entre sí y las comparamos tal que alguna diferencia significativa entre ellas podría ser atribuible a la campaña, vemos que en realidad no hay ningún cambio estadísticamente significativo que se le pueda atribuir a la campaña. En otras palabras, los cambios en las tasas de denuncias antes y después de la campaña #NoEsDeHombres no son explicados por la campaña.

Referencias

ONU Mujeres. 2017. Violencia contra las mujeres y las niñas en el transporte público de la Ciudad de México. México.

Angrsit, Joshua D., & Jörn- Steffen Pischke. 2015. *Mastering 'Metrics*. New Jersey: Princeton University.

Imai, Kosuke. 2017. *Quantitative Social Science An Introduction*. New Jersey: Princeton University.

Bertrand, M., Duflo, E., & Mullainathan, S. 2004. "How Much Should We Trust Differences-in-Differences Estimates? ". *Quarterly Journal of Economics*.