Universidad de San Andrés

Departamento de Economía

Maestría en Economía

Unravelling the Links between Public Transportation and Crime: The Case of Barcelona

Nicolás Agustín ROIG

36329759

Mentor: Daniel MONTOLIO

Buenos Aires

29 de Agosto, 2019

Tesis de Maestría en Economía de

Nicolás Agustín ROIG

"Crimen y Transporte Público: El Caso de Barcelona"

Resumen

El transporte público es central para cualquier metrópolis ya que se ofrece universalmente, facilita el desarrollo urbano y disminuye el tráfico automovilístico mejorando la calidad del aire. A pesar de sus evidentes beneficios, las consecuencias inintencionadas que la inversión pública en transporte tiene en el crimen, no son generalmente tenidas en cuenta por los hacedores de política, probablemente debido a la escasa evidencia acerca de su magnitud. A través de un "event study" analizo variabilidad temporal y espacial de una extensa base de datos de crimen para Barcelona, para estimar el impacto de la expansión de la red de Metro de Barcelona en el crimen local. En promedio, en las secciones censales que tienen al menos la mitad de su área dentro de un radio de 300 metros de la apertura de una estación, se observa un incremento del 8% en el número total de delitos ($p<0.01$), o un 14% de incremento en la tasa de criminalidad ($p<0.001$). El incremento en el crimen es explicado mayoritariamente por un efecto de largo plazo, y específicamente, asociado a un incremento en los crímenes contra la propiedad en lugar de crímenes contra las personas.

Palabras clave: transporte público, crimen, análisis espacial, event study

"Unravelling the Links between Public Transportation and Crime: The Case of Barcelona"

Abstract

Public transportation is central for any metropolis since it improves air quality by reducing automobile congestion, and facilitates compact development, conserving land and decreasing travel demand. Despite the striking benefits, the unintended effect of this type of investment on crime is hardly taken into account by policymakers, perhaps because of the little evidence of its magnitude. Exploiting spatial and time variability of a novel high-frequency crime dataset for Barcelona, I use an event study approach to examine the impact of Barcelona Metro expansion on local crime. On average, census tracts that have at least half of its area within 300 meters of a station opening expect to see an 8% increase in the total number of crimes ($p<0.01$) or a 14% increase in crime rates ($p<0.001$). The increase in crime is explained mostly by a long term effect and specifically, due to an increase in property crimes rather than crimes against persons.

Keywords: public transportation, metro, crime, spatial analysis, event study

Códigos JEL: R42, K42, Z18

I. INTRODUCTION

The expansion of the subway and rail system has transforming effects on the physical areas affected. It is natural to think that commerce and residences situated in the proximities of a new metro station are likely to change. That means that new economic opportunities would arise, not only in the legal sector but also in the illegal sector. While the positive impacts of such urban investments have been widely analyzed and accounted for, the negative impats are scarcely considered. In the present work I investigate the externalities related with a potential change in criminal activity in an area as a result of the opening of a metro station.

This investigation establishes a link between the expansions of the metro system in Barcelona over the period 2007-14 with the crime occurrences in the surrounding areas. Exploiting a rich high-frequency dataset of crime events with two-stage least-squares (2SLS) techniques, this article estimates an increase of 8% in the total amount crimes in census tracts that are within 300 meters of a metro station opening (*p-val<0.01*). Furthermore, I found that the effects are concentrated in the long run and mostly led by crimes against property. On average, the vicinity of a station opening is not affected in the short run (less than 1 year), while after 2 years estimates show an increase in crime of 12% (*p-val*<0.01) and an increase in crime rates of 20% (*p-val<0.001*).

II. A BRIEF LITERATURE OVERVIEW

The study of the determinants of crime in a theoretical manner can be traced back to Becker's (1968) seminal model. He proposed that criminals are rational individuals that find attractive to *work* in illegal contexts. An expansion in public transportation may have effects in either direction. On the one hand, new stations may increase monitoring or police presence, and thus committing a crime in that area becomes less attractive (Di Tella and Schargrodsky, 2004; Klick and Tabarrok, 2005; Draca et al., 2011). Nonetheless, more police may increase crime reporting by lowering the costs to do so or increase arrests by reducing police response times (Blanes I vidal and Kirchmaier, 2017). Crime can also be reduced because criminals that used to live in the area where a station opened might go to other places of the city (pre-existing stations) where returns are higher. On the other hand, a station opening may increase the returns to crime by producing crowds, bringing potential victims and offenders closer together (Felson et al., 1990; Brantingham and Brantingham, 1995). Myhre and Rosso (1996) argued that stations congregate easy targets, people that tend not to be alert, are tired, or are commuting and carrying things. Also, more connectivity may allow criminals from poorer areas to access richer areas to commit crimes.

Although public transportation expansions may offer criminals access to new markets or decrease their transportation costs, it is often considered an investment with positive net benefits. Even though neighbors often oppose the construction of public transport stations near their homes because they fear that crime will increase[1], the evidence from rigorous analyses of this link is inconclusive. The unintended effect of this type of public investment on crime is rarely taken into account by policymakers, perhaps because there is no evidence of its magnitude.

The results of several studies examining the effects of public transit on crime are mixed. Inlanfeldt (2003) finds some rare evidence of the link between transit and crime. His empirical analysis of the opening of new stations in the city of Atlanta shows a redistribution

[1] A study on resident's perceptions prior to the construction of a train station in Atlanta found crime as the second most major concern (Ross and Stein, 1985). In the last few years, some Santa Monica neighbors blamed the new Expo Line as responsible for the rising crime rates.

of crime from wealthy to poor areas. The Green Line light rail system in Los Angeles was found to be irrelevant to explain crime in the station neighborhoods (Ligget et al., 2003)[2]. Along with these results, studying the city of Charlotte, Billings et al. (2011) did not provide any evidence that light rail increase nor decrease crime around stations. Neiss (2015) finds that the addition of a bus line in Cleveland increased the mean property crime in the neighboring census tracts. Most of these articles face some sort of either methodological or data limitations: failing to place the analysis of the transport system in the larger metropolitan context, relying on aggregate data, or they lack a rigorous analysis of their identifying assumption. The inconclusive results show a need to clarify the relationship between public transportation and crime. With high-frequency micro data for Barcelona and by using an event study framework, I will be able to overcome the shortcomings exposed,

More compelling evidence come from Phillips and Sandler (2015) who use temporary maintenance-related closures of stations in Washington, DC's rail transit system to estimate how the availability of public transportation affects crime. Their main finding is that closures reduce crime in the vicinity of stations on the same train line. They find suggestive evidence that crime falls more at closures that happen on stations that tend to import crime. While their identification strategy was clear and convincing, it does not allow one to separate the direct effects of lowering transportation costs or investing in public infrastructure, nor to estimate medium and long term effects since it only exploit sharp micro-time series variation.

The main contribution of this study comes from exploiting a rich high-frequency dataset with geocoded crime rates. I can distinguish between types of crime (pickpocketing or violent crime, i.a.), where specifically the crime occurred (inside a station or in the street,

[2] The authors only analyzed the crime levels in the neighborhoods without considering crimes at the stations or the stations parking lots.

i.a.) and when (date and time). Though I cannot directly control for the link between potential police presence in a station and the low cost of filing a crime report, for a crime to be in my sample the victim had to file a police report at a police station, it does not suffice to meet a police official. The extension of the period under analysis allows me to study short-run and long-run impacts, while the spatial dataset allows to consider the transport infrastructure expansion in the context of the city as a whole. Policy implications are related to security concerns in urban planning. As my study precisely identify distance and time period in which a station opening is expected to influence each specific type of crime.

The remainder of the paper proceeds as follows. Section III describes the background of Barcelona Metro system and its expansion. Section IV presents the data sources, and Section V explains the empirical framework. Section VI discusses the results. Finally, Section VII provides a summary of the findings and concludes.

III. INSTITUTIONAL BACKGROUND

Barcelona has a well-developed public transport system which consists of a subway network (Barcelona Metro), commuter rail, trams, buses and even funiculars and cable cars. The Barcelona Metro is the most popular mean of transportation although it has complementarities with buses and commuter trains. In 2014 the Barcelona Metro, a network of mostly underground railway lines in central Barcelona and the city's suburbs, consisted of 12 lines and 172 stations adding up to a total length of around 137 kilometers. Ridership averaged more than 1 million riders per day. During the period under analysis (2007 through

2014) there were 19 station openings in 6 different lines occurring at 8 different dates. The spatial distribution of the openings is shown in FIGURE 1.[3]

The *Plan Director de Infraestructuras* (PDI) is an instrument through which the public metropolitan transportation authority establish expansion and modernization plans for the rail system. The first plan was for 2001-2010 and had a follow up stating the advances so far in 2009. The last plan which is also relevant for my study is the one for the period 2011-2020. These documents are likely to record important events such as announcements and details of the construction of the expansion. The announcement itself could potentially have an impact on crime as documented by Billings, et al. (2011), thus I will present results from an event analysis to assess whether the openings of the stations can be considered exogenous conditional on a series of controls.

IV. DATA DESCRIPTION

A. *Data Sources*

To complete this study, I will exploit a non-public geocoded dataset collected by the Catalan Police Department. This dataset contains detailed crime reports from January 2007 through December 31, 2013. Reports were filed by both citizens and *Mossos d'Esquadra* (the autonomous police agency in Catalonia, responsible for crime prevention and investigation in the Catalan region) and correspond to all crimes that occurred in the municipalities within the range of Barcelona Metro. There are a total of 12 such municipalities within this region.[4]

[3] A second timeline of the openings specifying the station's names and lines can be found in the appendix (FIGURE A1).

[4] The criteria is that the lowest distance between any station and any point of a municipality should be at the most 2 kilometers. Setting a criteria of 1 kilometer distance give same results.

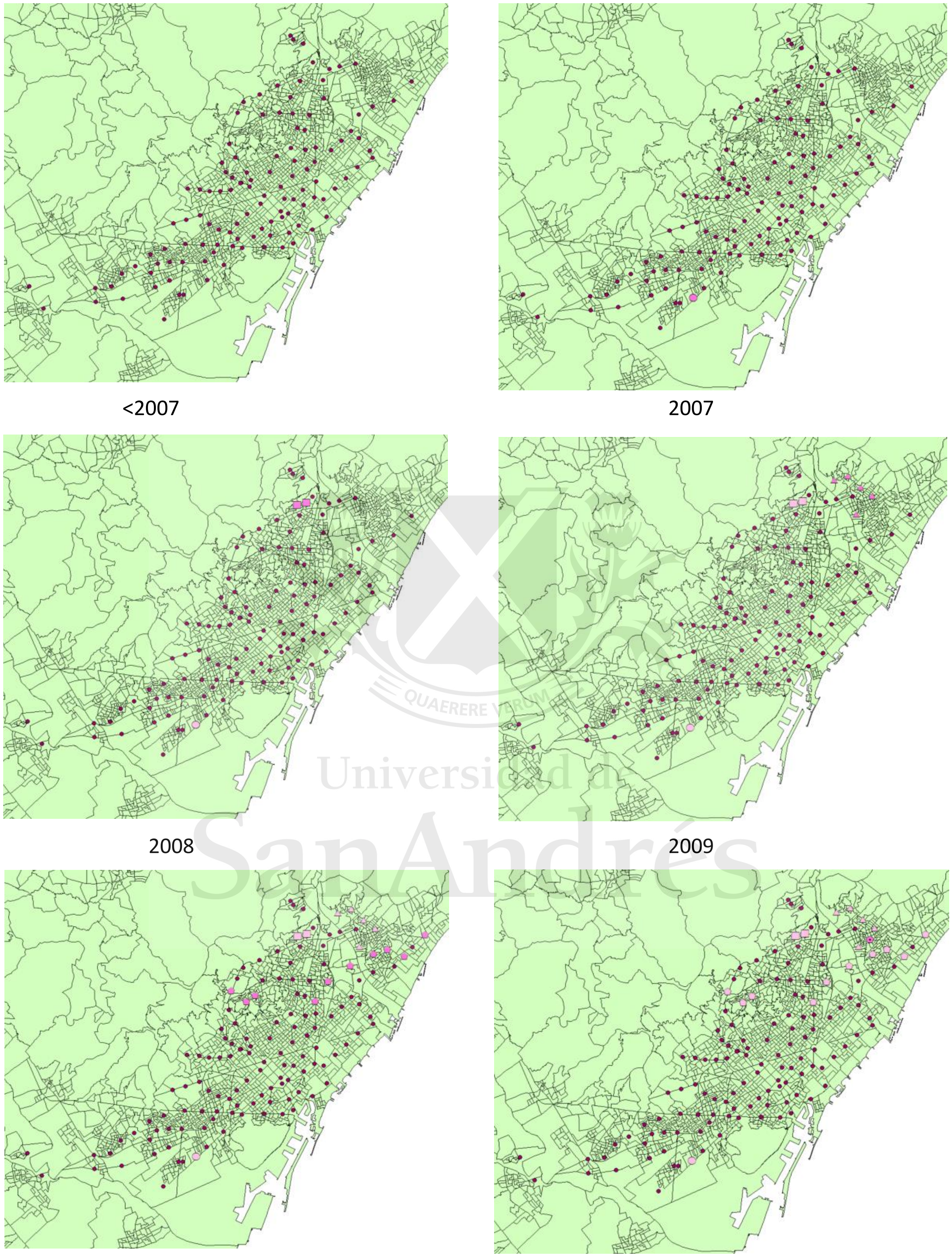

FIGURE 1. SPATIAL DISTRIBUTION OF METRO STATION OPENINGS BY YEAR

This dataset records the location and time where the crime took place, and the type of crime committed. There are 1,884,296 crimes reported varying only slightly by year.[5] There are 190 different types of crime, according to which article of the Spanish penal code was violated. Following Montolio (2018), I combined those articles and ended up with three main categories: crimes against persons, crimes against property, and other types of crimes. Then, further divided them in serious or minor and specific type of crime, making at last 28 categories as shown in TABLE A1 in the APPENDIX. Property crimes include thefts, robberies, car thefts and damages, while crimes against persons include murders, injuries, gender violence, sexual crimes and threats. Other crimes is mostly explained by traffic and consumption of drugs, and road safety. The advantage of his categorization is that crimes associated with a clear economic return are easily identifiable (property crimes).

Besides the crime dataset, I use information on Barcelona Metro stations (location, lines connectivity, and open dates) and geo-localized data on census tracts. The advantage of using census tracts is that it makes possible to generate crime rates since they provide information of the number of inhabitants living in each the relevant areas. However, the drawback is that since census tracts were not stable across time (they had their geographic areas modified), I rely on a linear interpolation[6]. First, I divided the metro area into 76,574 squares of side approximately 56 meters (See FIGURE 2). The median census tract has 11 squares. Once constructed this grid I intersected it with the crime dataset, which allowed me

[5] There is a little peak in 2009 and 2010 and then slightly decrease (from 13.45 at its maximum to 11.76 at its minimum).

[6] The reason of the update of the census tract map between 2009 and 2010 was to maintain them relatively homogenous in terms of population after the important demographic changes, in terms of migration patterns, experienced in Spain in the 2000s.

to count the number of crimes in each square for any unit of time. Throughout this paper, I will work with monthly data. There are 96 months in the period under study.

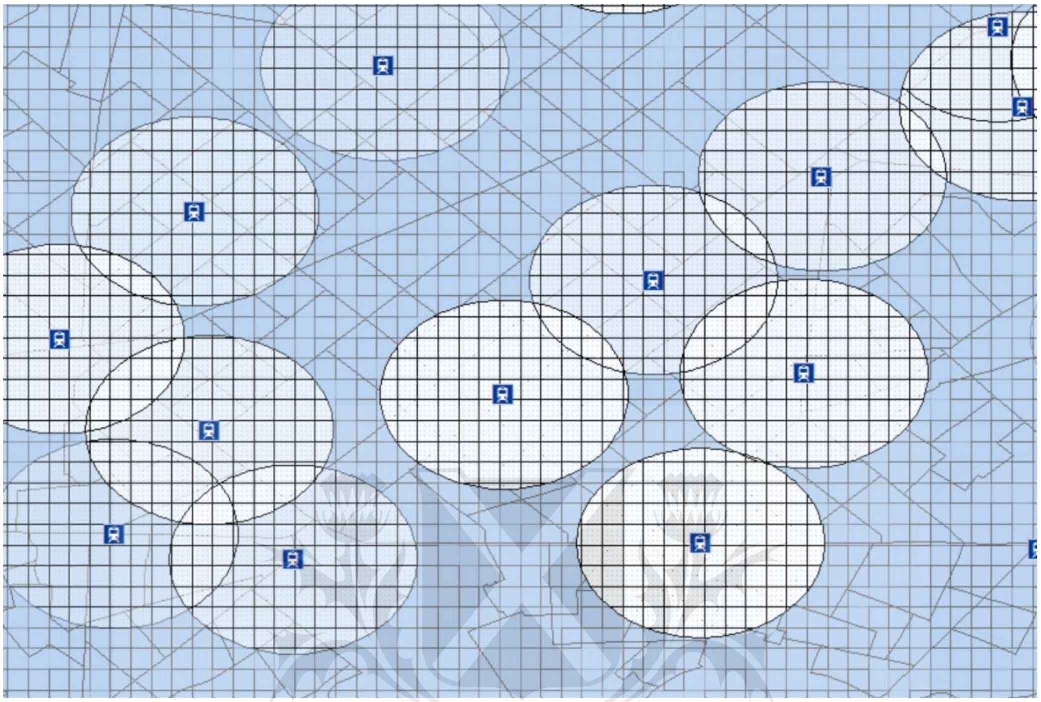

FIGURE 2. GRID USED TO CREATE TREATMENT UNITS AROUND METRO OPENINGS

Second, I created circles of 100, 300, 500 and 1,000 meters around each of the stations in the Barcelona metro system. After intersecting those buffers with the grid, I was able to identify the treated cells. Since the cross-sectional unit of analysis is census tract I had to define treatment criteria. A census tract is considered treated at a specific distance -after a station opening- if at least half of its area is within the treatment area. This decision is arbitrary, therefore I created 4 other different treatment criteria and all the results remain highly consistent.[7] To compute treatment status, I divide the number of squares in a treatment area belonging to a census tract divided by the total amount of squares in that same census

[7] Other treatment criteria consider a census tract as treated if 1%, 10%, 30% or 75% of its area is within a treatment area after the treatment period.

tract. Notice that a census tract can be treated in a 500-meter radius during some period and at a 100-meter radius during some other period. This is only possible when two stations are close enough to each other. Although this happens, it is highly unusual. To avoid problems of collinearity I defined the treatment at the closest criteria. That is, if half of the area of a census tract is both at a 100-meter radius of a new station opening but also at the area defined by the 300-meter to 500-meter ring, the census tract is considered as treated at a 100-meter radius only.

There are two main outcomes in my analysis: number of crimes and crime rate. Because of the aforementioned link between potential police presence on a metro station and ease to file a crime report, I excluded crimes that were located in the metro system in one of the robustness checks and in FIGURE A2 in the appendix.[8] Magnitudes change slightly and the estimation becomes more imprecise, nevertheless the main findings remain unaltered. I have a preference to test whether a new station makes a place more or less risky. The total amount of crimes is not a good proxy of how dangerous a zone is. A more crowded place might have a higher number of crimes even though is not a dangerous place in terms of the probability of being the victim of a crime. In this sense, crime rates may perform better. Since I do not have information on daytime population to compute that probability I will rely on the number of residents of a census tract.[9] Still, using information on the evolution of population by census tracts is complicated, as previously mentioned, a relatively large proportion of census tracts changed their geographical limits in 2010. To overcome this issue, I rely on a linear interpolation. Since I am working with the 2010's definition of census tracts,

---

[8] There are a total of 147,934 crimes reports located in one of the following: "metro infrastructure", "metro convoy", "metro stations", and "metro". Results remain highly consistent when including them.
[9] An analysis using ridership data will be possible in the future.

I have information on population for that year. Furthermore, I obtained information on the yearly evolution of inhabitants by neighborhood. There are 78 neighborhoods in the sample that on average have 23 census tracts each, though the median district has only 9 census tracts. For each census tract I computed its relative weight in terms of total district population. Then, I interpolated the evolution of inhabitants from the districts to the census tracts keeping the weights constant. After this process, I computed crime rates per 10,000 inhabitants for each census tract. Although this measure is not the most precise, it may perform better than the sum of crimes to capture how risky a census tract is.

TABLE I depicts summary statistics for census tracts in my sample by different radius and treatment criteria. There are large disparities in most of the variables selected reflecting structural differences among treatment and control areas. Most of the difference is likely to be explained by the differences in municipalities, Barcelona which is almost all in the control group differs in high degree with the other municipalities. However, my identifying assumption do not imply them to be alike. Conditional on time and census tract fixed effect I assume that treated areas would evolve as control units in the absence of the treatment. In the next section I will describe the event study framework to test this assumption.

I. EMPIRICAL STRATEGY

My purpose is to identify the causal effect of subway expansion on crime. In order to do so, I will exploit the total number of crimes per census tract during each month from January 2007 to December 2013. As I explained in the previous section I built a panel which has 1,810 census tracts and 96 months. To estimate the impact of station openings on crime, I estimate a difference-in-difference fixed effects model:

TABLE 1: SUMMARY STATISTICS

Treatment Criteria	Radius	N			Population		Density (km2)		Unemployment Rate		Crime Rate (per 10000 inhabitants)	
		Treatment	Control		Treatment	Control	Treatment	Control	Treatment	Control	Treatment	Control
1%	100-meter	78	1688	mean	1401	1463	36757	41487	0.295	0.241	44.011	70.473
				sd	*368*	*403*	*23883*	*24188*	*0.113*	*0.099*	*50.739*	*193.837*
	300-meter	212	1554	mean	1378	1471	42984	41046	0.291	0.237	40.772	73.280
				sd	*389*	*402*	*26214*	*23897*	*0.116*	*0.097*	*53.848*	*201.355*
	500-meter	314	1452	mean	1370	1479	41247	41285	0.282	0.236	41.922	75.357
				sd	*406*	*398*	*25943*	*23800*	*0.111*	*0.096*	*50.900*	*207.972*
	1000-meter	561	1205	mean	1400	1488	40299	41735	0.267	0.233	42.662	81.984
				sd	*389*	*404*	*24466*	*24053*	*0.106*	*0.096*	*48.361*	*227.188*
10%	100-meter	58	1708	mean	1402	1462	43716	41196	0.292	0.242	34.156	70.498
				sd	*381*	*402*	*22703*	*24238*	*0.118*	*0.100*	*32.463*	*192.848*
	300-meter	193	1573	mean	1361	1472	45360	40778	0.293	0.238	36.211	73.451
				sd	*390*	*401*	*25967*	*23921*	*0.117*	*0.097*	*41.413*	*200.492*
	500-meter	297	1469	mean	1360	1480	42905	40950	0.284	0.236	40.679	75.241
				sd	*397*	*399*	*25466*	*23916*	*0.111*	*0.097*	*50.784*	*206.863*
	1000-meter	554	1212	mean	1400	1487	40516	41627	0.268	0.233	42.370	81.919
				sd	*389*	*404*	*24448*	*24069*	*0.106*	*0.096*	*47.934*	*226.667*
30%	100-meter	28	1738	mean	1342	1462	49778	41142	0.298	0.243	28.822	69.942
				sd	*374*	*402*	*19672*	*24234*	*0.128*	*0.100*	*24.267*	*191.241*
	300-meter	160	1606	mean	1371	1469	48286	40580	0.300	0.238	33.064	72.986
				sd	*378*	*403*	*25308*	*23969*	*0.120*	*0.097*	*33.530*	*198.612*
	500-meter	276	1490	mean	1365	1477	44358	40708	0.288	0.236	38.565	75.142
				sd	*392*	*401*	*25343*	*23933*	*0.112*	*0.096*	*48.801*	*205.507*
	1000-meter	521	1245	mean	1398	1486	41589	41149	0.269	0.233	40.499	81.673
				sd	*393*	*402*	*24386*	*24112*	*0.106*	*0.097*	*45.346*	*223.943*
50%	100-meter	15	1751	mean	1376	1461	53574	41173	0.309	0.243	23.227	69.678
				sd	*326*	*402*	*14806*	*24228*	*0.116*	*0.100*	*18.638*	*190.549*
	300-meter	136	1630	mean	1364	1468	49804	40567	0.305	0.239	30.473	72.604
				sd	*382*	*402*	*24692*	*24016*	*0.114*	*0.098*	*30.389*	*197.202*
	500-meter	256	1510	mean	1364	1476	46078	40465	0.291	0.236	34.375	75.375
				sd	*393*	*401*	*24925*	*23974*	*0.112*	*0.096*	*35.017*	*204.576*
	1000-meter	503	1263	mean	1397	1485	41705	41109	0.270	0.233	40.607	81.021
				sd	*390*	*403*	*24436*	*24095*	*0.106*	*0.096*	*45.902*	*222.369*
75%	100-meter	5	1761	mean	1180	1461	59771	41226	0.235	0.244	28.828	69.391
				sd	*214*	*402*	*16800*	*24188*	*0.059*	*0.101*	*14.317*	*190.035*
	300-meter	105	1661	mean	1313	1469	52687	40557	0.311	0.240	30.379	71.801
				sd	*354*	*402*	*24428*	*23998*	*0.122*	*0.098*	*31.608*	*195.416*
	500-meter	219	1547	mean	1352	1475	48662	40233	0.291	0.237	32.683	74.611
				sd	*387*	*401*	*24624*	*23950*	*0.115*	*0.097*	*34.781*	*202.168*
	1000-meter	467	1299	mean	1398	1482	42825	40722	0.273	0.233	39.654	80.279
				sd	*387*	*404*	*24202*	*24167*	*0.107*	*0.096*	*45.425*	*219.532*

Notes: Summary statistics by type of treatment for different covariates including mean, standard deviation and size of each subsample. Population and density are at year 2010. Treatment is considered as if at least in one period of the whole sumple it was treated with the intensity and distance pre-specified. Unemployment rate is based on 2011 census, there are 46 missing values of a total of 1,766. Crime rates are computed for January 2007.

$$Crime_{ct} = \alpha\, 100Radius_{ct} + \beta\, 300Ring_{ct} + \gamma\, 500Ring_{ct} + \psi_c + \Omega_t + \varepsilon_{ct} \qquad (1)$$

where, $Crime_{ct}$ is either the number of crimes or crime rate in census tract c at time t. $100Radius_{ct}$ is a dummy variable that takes value 1 for all census tracts c that at least half of their area is within 100 meters of an opening station and all months after the opening and 0 otherwise. $300Ring_{ct}$ is a dummy variable that takes value 1 for all census tracts c that has at least half of their area is further than 100 meters and within 300 meters of an opening station and all months after the opening, and *100Radius* has value 0.[10] $500Ring_{ct}$ is a dummy variable that takes value 1 for all census tracts c that has at least half of their area is further than 300 meters and within 500 meters of an opening station and all months after the opening, and both *100Radius* and *300Ring* have value 0. ψ_c is a census tract fixed effect and Ω_t is a month fixed effect. The parameters of interest are α, β and γ which capture the effect of opening a station in crime for nearby census tracts.

A. *Exogeneity*

The regression strategy exposed rely on the assumption that station openings generate exogenous variation conditional on the controls included (time and spatial fixed effects). In other words, in the absence of the station opening crime would evolve similarly for treated and control census tracts (once controlled for levels). One way to test this assumption is by studying pre-treatment trends in crime. Plotting the evolution of the dependent variable by treatment group is one alternative but it lacks a statistic test. Moreover, treatment takes place at different points in time, meaning that I should either track 9 different series or plot months

[10] The 300-meter ring is defined as the 300-meter radius circle around a new station intersected with the complement of the 100-meter radius circle.

relative to treatment. The latter alternative bias the interpretation of the evolution since the time shocks would be different for different treatment units.

The most convenient valid alternative is to test for pre-treatment trends through an event study. Event studies have become very popular in the economic literature (mostly driven by the three journals focusing on applied microeconomic work) and provide several advantages.[11] Coefficient estimates can be graphed and are very intuitive to analyze potential pre-event trends. Also, it is a useful tool to study post-event effects. The estimating equation for *Crime* is:

$$Crime_{ct} = (\textstyle\sum_{k=-24}^{-1} \alpha_k \, 1\{t_c^* + k = t\} \; + \sum_{l=1}^{24} \alpha_l \, 1\{t_c^* + l = t\} \; + \\ \beta_l \, 1\{{t_c}^* - 24 > t\} + \beta_h 1\{t_c^* + 24 < t\}) 1\{t_c^* > 0\} + \psi_c + \Omega_t + \varepsilon_{ct} \qquad (2)$$

where t_c^* is the month at which the station near census tract *c* opened in one of the areas defined (100-meter radius, 300-meter radius, or 500-meter radius), and 0 otherwise. The coefficients of interest α_k , α_l, β_h and β_l trace out changes in the relationship between treatment units ($1\{t_c^* > 0\}$) and crime. Pre-treands are then study for 2 years before the intervention (*k=24*) while the 24 leads included in the regression analyze short and long run effects. β_h and β_l capture the differences between treatment and control units more than 48 months after the treatment and less than 24 months before the treatment, respectively. The omitted dummy is the month of the treatment period, estimates become relative to that period.

FIGURE 3 depicts the estimated coefficients of interest from Eq. (2) with their respective 95% confidence intervals. The Huber-White standard errors are clustered at the

[11] Refer to Sechmidheinv and Siegloch (2019) for more details.

census tract level. The left panel has number of crimes as the dependent variable, while the right panels have crime rates. Eq. (2) was estimated for each of the possible treatment areas: 100-meter radius, 300-meter radius, 500-meter radius and 1000-meter radius. The pre-treatment period coefficients are not statistically different from 0 in most of the coefficient estimates for every specification. There are only a handful of estimates statistically different

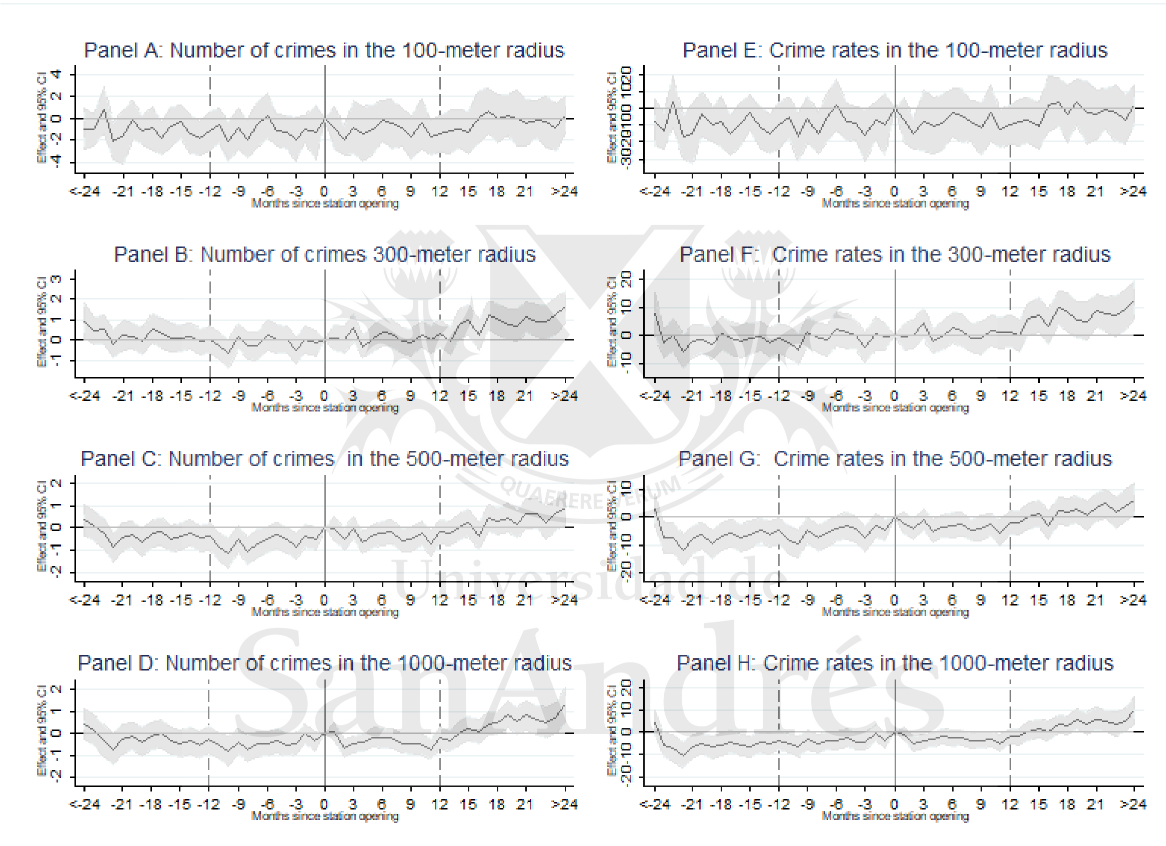

FIGURE 3. EVENT STUDY ESTIMATES OF METRO STATION OPENINGS ON CRIME

from 0 but that can be just by chance. Even if coefficients are smaller than 0, their pre-treatment trend is parallel to the x-axis meaning that there is not different evolution between control and treatment groups. Thus, conditional on census tract fixed effects station openings generate exogenous variation making it possible to estimate the difference-in-difference estimator from Eq. (1).

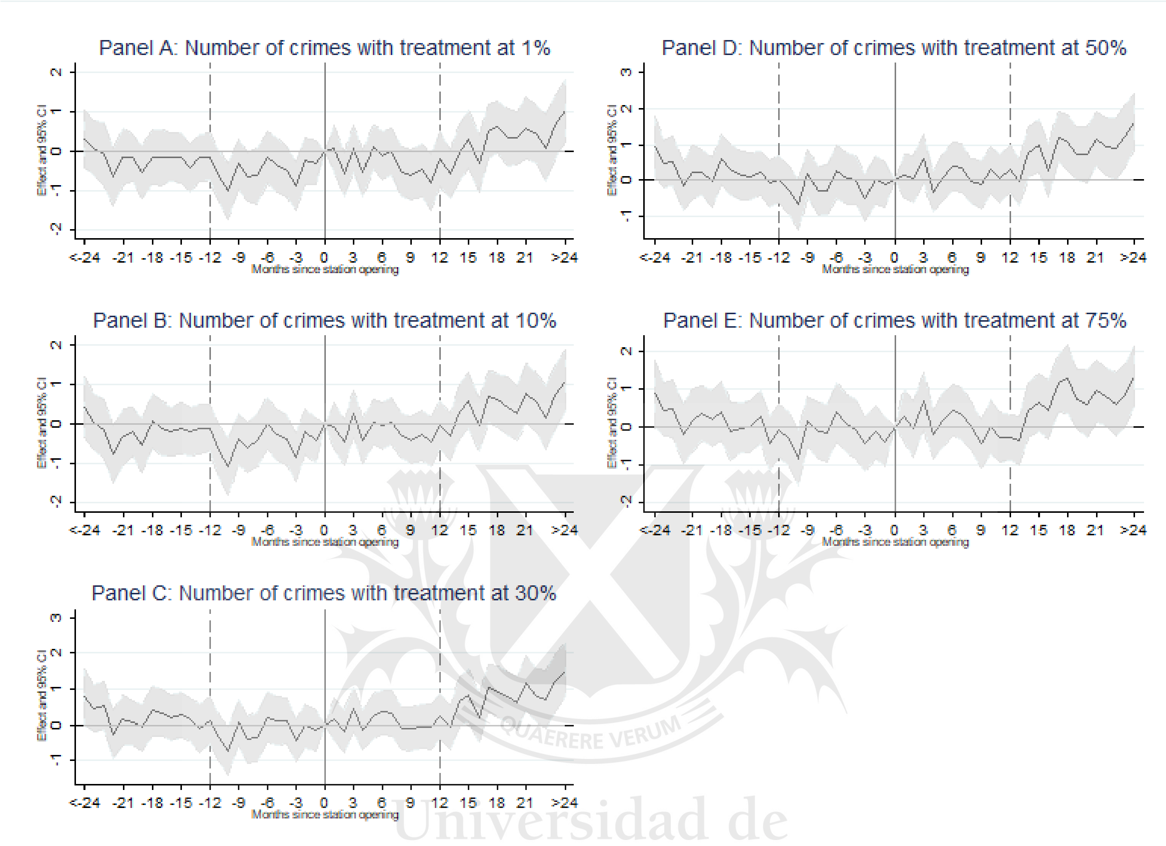

FIGURE 4. EVENT STUDY FOR 300-METER RADIUS BY TREATMENT CRITERIA

Results are largely consistent across treatment criteria and different specifications of Eq. (2). FIGURE 4 presents the outcomes from the event study for the 300-meter radius specification for each of the 5 treatment assignment criteria (1%, 10%, 30%, 50% and 75%) again with Huber-White standard errors clustered at the census tract level. The event study results also show a highly significant long-term effect which starts just 1 year after the opening of the stations. The result is stronger for the 100-meter radius though more imprecise since it has lower variability. More in depth study of the short and long run effects of the treatment on crime is undertaken in the following section. Finally, FIGURE A2 depicts the results for the same analysis but excluding crimes in the metro area. Estimations become much more volatile, in particular for the 100-meter radius.

II. THE EFFECT OF STATION OPENINGS ON CRIME

A. *Main Results*

TABLE 2 depicts the results of the main regression specified in Eq. (1). Column (A) and (D) only include the 100-meter Radius as the independent variable, results are relatively large and statistically significant ($p<0.05$). The magnitude of the increase in crime due to the station openings is on the order of a 10% increase. A census tract in the neighboring areas of a station opening expects to see almost 1 more crime per month due to the opening. The effect does not fade out as you expand the radius from 100 to 500 meters as shown in columns (B) and (C). The right panel reports results with crime rates as the dependent variable and they are still positive and highly significant (most p<0.001). The estimated effect for crime rates is even higher than for number of crimes, the ATE for the 300-meter radius is 13.90% of the mean level.

B. *Robustness*

In this subsection, I further test the main results to assess its internal validity. To do so, I first focused on regressions including the *300-meter Radius* explanatory variable only. TABLE 3 presents the results for the main regression under different specifications. Column (A) has all crimes excluding crimes at the metro system as the dependent variable. The exclusion of crimes in the metro facilities only increased the variance of its estimate. This follows from the fact that a 100-meter radius around a metro station is almost full of station facilities, excluding those crimes lowers significantly the variability of my explained variable. Second, there are four out of the nineteen total openings that occurred during the period under study that made an expansion of a pre-existing metro facility. Although there is a new station (connecting a new line), the environment has not changed from not having a metro facility to having one. Therefore, I split the two and include both of them as regressors, results can be found in Column (B). Although differences in magnitude both coefficients are positive and significant (*p<0.05).*

TABLE 2: THE EFFECT OF STATION OPENINGS ON CRIME

	No. of Crimes			Crime Rate		
	(A)	(B)	(C)	(D)	(E)	(F)
100-meter Radius	0.893*	1.006*		9.202**	10.762***	
	(0.412)	(0.438)		(3.017)	(3.255)	
100 to 300-meter Ring		0.861**			10.683***	
		(0.317)			(2.579)	
300 to 500-meter Ring		0.613*			8.869***	
		(0.311)			(2.524)	
300-meter Radius			0.907**			10.165***
			(0.296)			(2.344)
Census tract fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	169,536	169,536	169,536	165,120	165,120	165,120
R-squared	0.891	0. 891	0. 891	0.879	0.879	0.879
Mean of dependent variable	11.11	11.11	11.11	73.11	73.11	73.11

Notes: Least-squares dummy variables (LSDV) regressions. Huber-White standard errors clustered at the census tract level are in parentheses; ***, ** and * denote statistical significance at the 0.1%, 1% and 5% levels, respectively.

Third, column (C) and Column (D) estimate Eq. (1) with robust standard errors clustered at the municipality-month and district-month groups, respectively. The standard error only lowered making the identification more precise. Both coefficients are significant at the 0.1% level. Fourth, column (E) treats observations with number of crimes over the 95% percentile as having the value at that percentile (32 crimes per month). The magnitude of the effect lowered to 0.59 crimes per month per census tract but the precision of the estimate increases by showing a p-value smaller than 0.001. Column (F) excludes observations with zeros throughout the sample. As expected, the estimated effect is slightly larger.

Fifth, column (G) depicts the results for the computation of the DID estimator with the sample of 76,574 squares of side 56 meters including grid fixed effects and clustering standard errors at the census tract level. The effect in this last specification raises to 25% of the mean variable (*p-val<0.001*). The increase in the magnitude might be due to better identification of the treatment units and the inclusion of a larger number of controls (grid fixed effects). Estimates remain highly significant across specifications.

TABLE 3: ROBUSTNESS

	No. of Crimes						
	(A)	(B)	(C)	(D)	(E)	(F)	(G)
300-meter Radius	0.839**		0.907***	0.907***	0.596***	0.993**	0.055***
	(0.264)		(0.151)	(0.122)	(0.174)	(0.321)	(0.011)
300-meter Radius – Expansion		1.854**					
		(0.693)					
300-meter Radius – New Station		0.667*					
		(0.275)					
Census tract fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grid fixed effects	.	.	.	.	.	.	Yes
Observations	173,760	173,760	173,760	173,760	169,536	155,980	7,351,104
R-squared	0.892	0.892	0.892	0.892	0.850	0.891	0.680
Number of clusters	1766	1766	1152	7296	1766	1744	1766
Mean of the dependent variable	11.11	11.11	11.11	11.11	7.79	12.08	0.24

Notes: Least-squares dummy variables (LSDV) regressions. Huber-White standard errors clustered at the census tract level are in parentheses unless noticed; ***, ** and * denote statistical significance at the 0.1%, 1% and 5% levels, respectively. Column (A) excludes crimes at the metro system as the dependent variable. Column (B) presents the results for those stations that only opened a new connection (*Expansion*) and those which are proper new stations (*New Station*). Column (C) and Column (D) estimate equation (1) with robust standard errors clustered at the municipality-month and district-month groups, respectively. Column (E) treats observations with number of crimes over the 95% percentile as having the value at that percentile (32 crimes per month). Column (F) excludes observations with zeros throughout the sample. Column (G) depicts the results for the grid sample, including grid fixed effects and clustering standard errors at the census tract level.

C. *Short and Long Run Effects*

In this section, I investigate the presence of heterogeneous effects depending on the time passed since the opening of a metro station (short versus long-run effects). Results depicted in FIGURE 32 suggest the study of short and long run effects. In order to do so, I defined four periods after the treatment: first semester (0-6 months), second semester (6-12 months), the second year (12-24 months) and more than two years (>24 months). TABLE 4 supports the findings of FIGURE 3: for almost every estimation the coefficients for the short run (less than a year after treatment) are not significant and much lower in size even with negative signs. Therefore, the main impact captured in TABLE 2 is driven by the longer-term effect (more than two years after treatment), this result holds for both total number of crimes occurred and crime rates.

D. *Heterogeneous Effects by Type of Crime*

In this subsection, I present the results by type of crime. Those are introduced in TABLE 5[12]. As shown in TABLE A1 the most common crimes are property crimes, 87% of the sample. Each census tract has on average almost 5 property crimes per month. The introduction of new stations has an increasing effect on minor property crimes of 20% in the longer run (>24 months), while serious property crime only increases in the order of 5% in the long-run and shows a 5% decrease in the short run (*p-val<0.001)*. For all the other specification, the effect is almost always zero or negative in the short run. Property crimes are the ones leading to the increasing effect that station openings have on crime. Crimes against persons (minor) are affected significantly by the openings but only in the 100 to 300

[12] The estimations presented in TABLE 5 do not exclude crimes committed within the metro premises. As for our previous estimates, results excluding these crimes are fully consistent with those presented in this section.

meter ring. It remains to be explained why this result does not hold on the 100-meter radius with a decreasing effect of 5% (*p-val<0.001*) in the short run for serious property crimes.

TABLE 4: SHORT AND LONG TERM EFFECTS OF STATION OPENINGS ON CRIME

		No. of Crimes				Crime Rate			
		(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)
300-meter Radius	*0-6 months*	-0.021				2.690			
		(0.218)				(1.731)			
	6-12 months	-0.107				2.123			
		(0.261)				(2.102)			
	12-24 months	0.710*				8.735***			
		(0.304)				(2.446)			
	>24 months	1.428***				14.281***			
		(0.371)				(2.906)			
100-meter Radius	*0-6 months*		0.223		0.232		4.058		4.271
			(0.430)		(0.434)		(2.961)		(2.997)
	6-12 months		-0.163		-0.142		2.201		2.543
			(0.364)		(0.371)		(2.939)		(2.989)
	12-24 months		0.752		0.800		7.896**		8.479**
			(0.409)		(0.417)		(2.916)		(2.986)
	>24 months		1.353**		1.438**		12.589***		13.499***
			(0.498)		(0.510)		(3.637)		(3.737)
100 to 300-meter Ring	*0-6 months*			-0.109	-0.107			2.201	2.229
				(0.216)	(0.217)			(1.779)	(1.784)
	6-12 months			-0.226	-0.223			1.394	1.434
				(0.261)	(0.262)			(2.178)	(2.185)
	12-24 months			0.583	0.589			8.177**	8.245**
				(0.306)	(0.307)			(2.553)	(2.563)
	>24 months			1.302***	1.314***			13.810***	13.925***
				(0.368)	(0.370)			(2.972)	(2.989)
Census tract fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations		169,536	169,536	169,536	169,536	165,120	165,120	165,120	165,120
R-squared		0.891	0.891	0.891	0.891	0.879	0.879	0.879	0.879
Mean of dependent variable		11.11	11.11	11.11	11.11	73.11	73.11	73.11	73.11

Notes: Least-squares dummy variables (LSDV) regressions. Huber-White standard errors clustered at the census tract level are in parentheses.; ***, ** and * denote statistical significance at the 0.1%, 1% and 5% lvels, respectively.

TABLE 5: THE EFFECT OF STATION OPENINGS ON CRIME BY TYPE OF CRIME

		Property crimes (minor)		Property crimes (serious)		Crimes against persons (minor)		Crimes against persons (serious)		Other crimes	
		(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)
300-meter Radius	*0-6 months*	-0.030		-0.013		0.023		-0.041		0.024	
		(0.113)		(0.069)		(0.022)		(0.021)		(0.021)	
	6-12 months	0.062		-0.230***		0.050*		0.012		-0.013	
		(0.099)		(0.067)		(0.024)		(0.025)		(0.021)	
	12-24 months	0.582***		0.025		0.068***		0.007		0.009	
		(0.090)		(0.055)		(0.018)		(0.018)		(0.016)	
	>24 months	1.013***		0.267***		0.044***		0.000		0.061***	
		(0.087)		(0.045)		(0.013)		(0.014)		(0.012)	
100-meter Radius	*0-6 months*		0.416*		-0.276*		0.067		-0.128*		0.085
			(0.179)		(0.132)		(0.064)		(0.059)		(0.060)
	6-12 months		0.089		-0.256		0.021		-0.070		-0.038
			(0.159)		(0.144)		(0.059)		(0.067)		(0.046)
	12-24 months		1.003***		0.118		0.053		-0.026		0.024
			(0.170)		(0.124)		(0.053)		(0.050)		(0.035)
	>24 months		1.066***		0.262**		0.085*		-0.032		0.059*
			(0.113)		(0.094)		(0.037)		(0.039)		(0.028)
100 to 300-meter Ring	*0-6 months*		-0.110		-0.041		0.003		-0.022		0.005
			(0.066)		(0.073)		(0.023)		(0.022)		(0.021)
	6-12 months		-0.013		-0.290***		0.050*		0.033		-0.011
			(0.069)		(0.068)		(0.025)		(0.026)		(0.022)
	12-24 months		0.409***		-0.067		0.067***		0.011		-0.008
			(0.058)		(0.058)		(0.020)		(0.019)		(0.017)
	>24 months		0.885***		0.174***		0.024		0.006		0.049***
			(0.051)		(0.047)		(0.014)		(0.014)		(0.013)
Census tract fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations		169,536	169,536	169,536	169,536	169,536	169,536	169,536	169,536	169,536	169,536
R-squared		0.881	0.881	0.855	0.855	0.442	0.442	0.365	0.365	0.659	0.659
Mean of dependent variable		4.872	4.872	4.831	4.831	0.469	0.469	0.439	0.439	0.504	0.504

Notes: Least-squares dummy variables (LSDV) regressions. Huber-White standard errors are in parentheses; ***, ** and * denote statistical significance at the 0.1%, 1% and 5% levels, respectively.

III. CONCLUSIONS

By studying the expansion of the metro system in Barcelona during the period 2007-2014, I assessed the impact of having a station opening on crime in the neighboring areas. Exploiting a high-frequency dataset of crime events, I estimated that the average treatment effect is an increase in 1 crime per month *($p<0.01$)* in a census tract that has at least half of its area within 300 meters of a station opening or an increase in 10 crimes per 10,000 inhabitants per month ($p<0.001$). It reflects an increase of 8% in the total number of crimes or an increase of 14% in crime rates. These results remain robust to different specifications.

Further exploring how the effects evolve, I found that the results are largely explained by long-run impacts. Less than one year after a station opens, crime in the vicinity is not affected on average. After two years the increase in crime is estimated to be between 10% and 26% depending on the specification and results are highly significant ($p<0.001$). Finally, I studied whether all types of crimes were affected in the same way, I found that the impact was mostly driven by crimes against property. Thus, I would conclude that the increase in crime is explained mostly by a long term effect and specifically, due to an increase in property crimes rather than crimes against persons. This may be of interest for policymakers that can decide on urban design to prevent crime. Rather than embracing the idea that public transportation is overprovided, this piece of research intend to shed light on how complimentary policies can be design to ameliorate any negative externalities in crime[13].

Criminological literature pointed out that the appeal of a site as a target for a crime depends, among others, on the type of land uses, level of surveillance, accessibility, environmental factors and perceived opportunities for escape. Also, station crime is strongly related to ridership. I cannot discard that the increase in crime is due to an increase in density of population or daytime passersby or because of a flourishing economy, station openings may have transformed residential to business areas while creating hotspots for crime. Exploring these mechanisms sets an agenda for further research to better understand the effect of public transportation expansion on crime.

[13] After a systematic review on the literature of hot spots policing, Braga et al. (2014) concludes that problem-oriented policing interventions are cost-effective policies.

REFERENCES

Becker, G., 1968. Crime and punishment: an economic approach. Journal of Political Economy 76 (2).

Billings, S., Leland, S., Swindell, D., 2011. The effects of the announcement and opening of light rail transit stations on neighborhood crime. Journal of Urban Affairs 33 (5).

Blanes I Vidal, J., Kirchmaier, T., 2017. The effect of police response time on crime clearance rates. Review of Economic Studies 85 (2), 855–891.

Braga, P., Papachristos, A. V., Hureau D. M., 2014. The effects of hot spots policing on crime: An updated systematic review and meta-analysis, Justice Quarterly, 31:4, 633-663.

Brantingham, P., Brantingham, P., 1995. Criminality and Place Crime Generators and Crime Generators and Crime Attractors. European Journal of Criminal Policy and Research 3, *5*-26.

Bowes, D. R., Ihlanfeldt K. R., 2001. Identifying the impacts of rail transit stations on residential property values. Journal of Urban Economics 50 (1).

Di Tella, R., Schargrodsky, E., 2004. Do police reduce crime? Estimates using the allocation of police forces after a terrorist attack. American Economic Review 94 (1).

Draca, M., Machin, S., Witt, R., 2011. Panic on the streets of London: Police, crime, and the July 2005 terror attack. American Economic Review, 101 (5), 2157–2181.

Felson, M., Dickman, D., Glenn, D., Kelly, L., Lambard, G., Maher, L., Nelson-Green, L., Ortega, C., Pretser, T., Rajedran, A., Ross, T., Tous, L., Veil, J., 1990. Preventing Crime at Newark Subway Stations. Security Journal 1 3, 137-142.

Ihlanfeldt, K., 2003. Rail transit and neighborhood crime: the case of Atlanta, Georgia. Southern Economic Journal 70 (2).

Jackson, C.K., Owens, E.G., 2011. One for the road: public transportation, alcohol consumption, and intoxicated driving. Journal of Public Economics 95.

Klick, J., Tabarrok, A., 2005. Using terror alert levels to estimate the effect of police on crime. Journal of Law Economics 48 (1), 267–279.

Liggett, R., Loukaitou-Sideris, A., Iseki, H., 2003. Journeys to crime: Assessing the effects of a light rail line on crime in the neighborhoods. Journal of Public Transportation, 6 (3): 85–115.

Montolio, D., 2018. The Effects of Local Infrastructure Investment on Crime. Labour Economics 52.

Myhre, M., Rosso, F., 1996. Designing for Security in Meteor: A Projected New metro Line in Paris in Clarke, R. (ed) Preventing Mass Transit Crime, Monsey, NY-Criminal Justice Press.

Neiss, M., 2015. Does Public Transit Affect Crime? The Addition of a Bus Line in Cleveland. The Journal of Economics and Politics 22 (1).

Phillips, D. C., Sandler D., 2015. Does Public Transit Spread Crime? Evidence from temporary rail station closures. Regional Science and Urban Economics 52 (1): 13-26.

Ross, C., Stein, J., 1985. Business and Residential Perception of a Proposed Rail Station Implications for Transit Planning.Transportation Quarterly 39 4, 483-93.

Schmidheiny, K., Siegloch, S., 2019. On Event Study Designs and Distributed-Lag Models: Equivalence, Generalization and Practical Implications. IZA Discussion Paper No. 12079.

APPENDIX

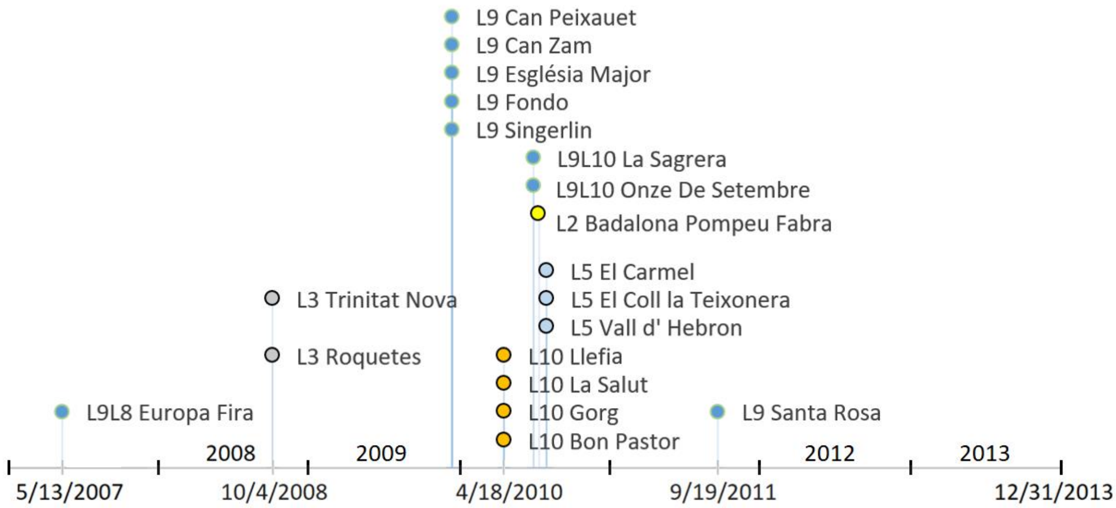

FIGURE A1. TIMELINE OF METRO STATION OPENINGS

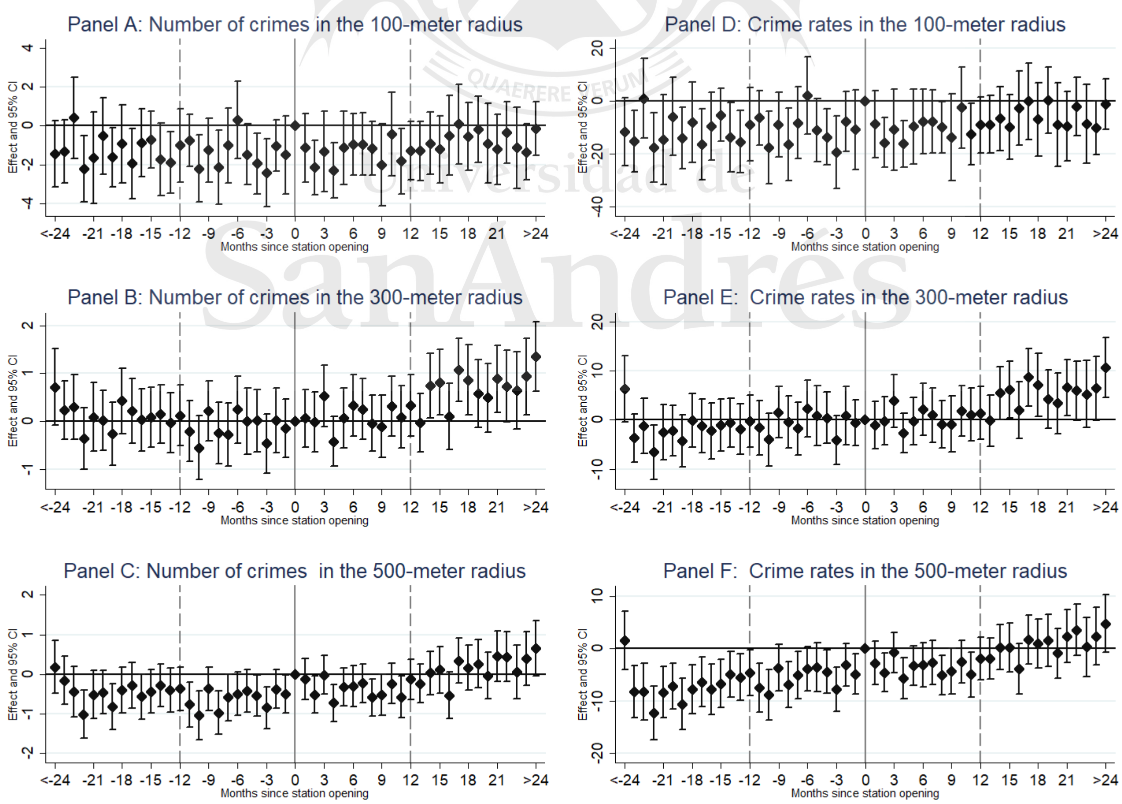

FIGURE A2. EVENT STUDY ESTIMATES OF METRO STATION OPENINGS ON CRIME EXCLUDING THOSE IN THE METRO SYSTEM

TABLE A1: TYPES OF CRIME

<table>
<tr><th>Main Type</th><th>Gravity</th><th>Specific Type</th><th colspan="3"># of crimes</th><th colspan="3">% of total</th></tr>
<tr><td rowspan="11">Crimes against persons</td><td rowspan="7">Serious</td><td>Gender violence</td><td>30,631</td><td rowspan="7">30,631</td><td rowspan="11">110,082</td><td>1.63%</td><td rowspan="7">3.95%</td><td rowspan="11">8.17%</td></tr>
<tr><td>Threats</td><td>17,400</td><td>0.92%</td></tr>
<tr><td>Injuries</td><td>11,772</td><td>0.62%</td></tr>
<tr><td>Others</td><td>5,375</td><td>0.29%</td></tr>
<tr><td>Sexual</td><td>4,923</td><td>0.26%</td></tr>
<tr><td>Family</td><td>3,499</td><td>0.19%</td></tr>
<tr><td>Murder</td><td>865</td><td>0.05%</td></tr>
<tr><td rowspan="4">Minor</td><td>Injuries</td><td>42,399</td><td rowspan="4">79,451</td><td>2.25%</td><td rowspan="4">4.22%</td></tr>
<tr><td>Threats</td><td>26,259</td><td>1.39%</td></tr>
<tr><td>Family</td><td>9,258</td><td>0.49%</td></tr>
<tr><td>Gender violence</td><td>1,535</td><td>0.08%</td></tr>
<tr><td rowspan="10">Crimes against property</td><td rowspan="6">Serious</td><td>Theft</td><td>286,349</td><td rowspan="6">819,055</td><td rowspan="10">1,645,015</td><td>15.20%</td><td rowspan="6">43.47%</td><td rowspan="10">87.30%</td></tr>
<tr><td>Robbery</td><td>249,512</td><td>13.24%</td></tr>
<tr><td>Car theft</td><td>193,237</td><td>10.26%</td></tr>
<tr><td>Fraud</td><td>56,350</td><td>2.99%</td></tr>
<tr><td>Damages</td><td>33,065</td><td>1.75%</td></tr>
<tr><td>Burglary</td><td>542</td><td>0.03%</td></tr>
<tr><td rowspan="4">Minor</td><td>Theft</td><td>664,336</td><td rowspan="4">825,960</td><td>35.26%</td><td rowspan="4">43.83%</td></tr>
<tr><td>Damages</td><td>119,301</td><td>6.33%</td></tr>
<tr><td>Fraud</td><td>42,103</td><td>2.23%</td></tr>
<tr><td>Car theft</td><td>220</td><td>0.01%</td></tr>
<tr><td rowspan="7" colspan="2">Others</td><td>Law and order serious</td><td>23,459</td><td rowspan="7" colspan="2">85,365</td><td>1.24%</td><td rowspan="7" colspan="2">4.53%</td></tr>
<tr><td>Drugs</td><td>11,204</td><td>0.59%</td></tr>
<tr><td>Environment serious</td><td>144</td><td>0.01%</td></tr>
<tr><td>Road safety</td><td>38,010</td><td>2.02%</td></tr>
<tr><td>Arson</td><td>291</td><td>0.02%</td></tr>
<tr><td>Law and order minor</td><td>10,596</td><td>0.56%</td></tr>
<tr><td>Environment minor</td><td>1,661</td><td>0.09%</td></tr>
</table>