Traffic and Crime*

Louis-Philippe Beland
Louisiana State University

Daniel Brent
Louisiana State University

December, 2017

Abstract

We study the link between crime and extreme traffic congestion to estimate the psychological costs of traffic. Our empirical analysis combines police incident reports with observations of local traffic data in Los Angeles from 2011 to 2015. This rich dataset allows us to link traffic with criminal activity at a fine spatial and temporal dimension. Our identification relies on deviations from normal traffic to isolate the impact of abnormally high traffic on crime. We find that traffic above the 95th percentile increases the incidence of domestic violence, a crime shown to be affected by emotional cues, but not other crimes. The result is robust to a variety of specifications and falsification tests. Since most drivers stuck in traffic do not commit domestic violence, but still bear some emotional costs, the results represent a lower bound of the psychological costs of traffic congestion.

JEL Classification: R28, D03, J12
Keywords: Traffic, Crime, Externalities, Domestic Violence

*Beland: Department of Economics, Louisiana State University, lbeland@lsu.edu. Brent: Department of Economics, Louisiana State University, dbrent@lsu.edu. We would like to thank Chelsea Ursaner, open data coordinator at the Office of Los Angeles Mayor for her precious help with the crime data.

1 Introduction

Traffic congestion is a severe problem in many cities that imposes substantial costs on the economy due to lost time, pollution, and increased gasoline expenditure. In metropolitan areas, road congestion led consumers to purchase 2.9 *additional* billion gallons of fuel and spend 5.5 billion hours sitting in traffic (Schrank et al., 2012). According to the Texas A&M Transportation Institute, an average commuter wastes 42 hours a year stuck in traffic - more than an entire week of full time work.[1] Given that most roads in the U.S. are unpriced, the externalities associated with traffic represent an enormous welfare cost to urban residents.

Sitting in traffic is an extremely unpleasant use of time for most people, and in certain circumstances traffic can be incredibly disruptive.[2] While the primary costs of traffic are due to lost time and reliability, there is research using survey data linking traffic to negative mental health outcomes, including stress and aggression (Parkinson, 2001; Hennessy and Wiesenthal, 1999; Gee and Takeuchi, 2004; Gottholmseder et al., 2009; Roberts et al., 2011; Künn-Nelen, 2016).[3] Using subjective well being data, recent research by Anderson et al. (2016) shows that the estimated costs of congestion greatly exceed typical estimates that account for lost time and reliability. This discrepancy is consistent with large psychological costs of traffic congestion, although this is not tested directly.

In this paper, we extend the literature on the costs of traffic congestion. In particular, we focus on the effect of traffic on domestic violence, which has been shown to be sensitive to emotional cues from local football teams' unexpected losses (Card and Dahl, 2011). We estimate the impact of emotional cues due to high traffic on the incidence of domestic violence in Los Angeles County. Los Angeles is a candidate for the worst traffic in the U.S.; six of the country's 10 most congested stretches of highway are in the Los Angeles metropolitan area.[4] Our primary contribution is to quantify a specific outcome of the emotional costs of traffic congestion using observational data. We also build on the literature of the economic

[1] See the Annual Urban Mobility Scorecard report from Texas A&M Transportation Institute available at: http://mobility.tamu.edu/ums/.

[2] Kahneman et al. (2004) finds that commuting is one of the least pleasant daily activities, and traffic can cause a late arrival or missing a business meeting, flight, court appearance or family responsibilities.

[3] For estimates of the value of time and reliability see among others Small et al. (2005).

[4] See the INRIX 2015 Traffic Scoreboard, available at: http://inrix.com/scorecard/.

consequences of emotional cues. Traffic will not induce most people to commit crimes but will still impose a psychological burden; therefore we consider our estimates a lower bound on the psychological cost of traffic.

Our empirical analysis combines police incident reports with observations of local traffic data in Los Angeles from 2011 to 2015. This rich dataset allows us to link traffic with criminal activity at a fine spatial and temporal dimension. Our empirical strategy relies on traffic shocks to estimate the effect of traffic on domestic violence. We find that extreme traffic (above the 95th percentile) significantly increases the incidence of domestic violence by approximately 9%. Since our primary outcome of interest, domestic violence, typically occurs in the home, we are confident that the offender faced the traffic that is typical of the commute at the location of the crime. We control for unobserved effects across space and time with fixed effects, and time-varying measures of traffic in the most recent week and month to control for changes in traffic expectations. Our results are robust to multiple specifications and falsification tests. There is no effect of traffic on lagged domestic violence incidents, no effect of evening traffic on morning domestic violence incidents, no effect of randomized traffic on domestic violence, no effect of traffic in other areas of the city, and no effect of traffic on other categories of crime such as property crime and homicides. To alleviate concerns of endogeneity we show that our results are similar when instrumenting for traffic conditions with the number and duration of severe accidents. We also investigate the differential impact of expected vs. unexpected traffic on domestic violence. We find some evidence that unexpected severe traffic leads to larger effects than expected severe traffic. However, most of the differential effects are not statistically significant. A challenge in separating the role of expectations is that it is we do not perfectly know how drivers form their traffic expectations. The results are robust to including daily and lagged internet search volume for traffic in the region as an alternative control for traffic expectations.

The effects are also economically important. Using published estimates of the costs of different crimes indicates that extreme traffic is responsible for approximately $5-22 million in annual damages due to increased incidence of domestic violence.[5] While these additional costs are small relative to the cost of lost time and pollution, we consider them to be extreme

[5]Direct and indirect cost to assaults are valued at $107,020 in 2008 dollars according to McCollister et al. (2010).

lower bounds. Most drivers who experience acute congestion will not commit crimes but still suffer some welfare loss due to stress, thereby greatly increasing the psychological costs of traffic.

The rest of the paper is organized as follows: Section 2 discusses the related literature; Section 3 provides a description of the data and presents descriptive statistics; Section 4 presents the empirical strategy; Section 5 is devoted to the main results, heterogeneity of the impacts, a series of robustness checks and investigation of the role of expectation and habituation; and Section 6 concludes with policy implications.

2 Literature review

Our paper is related to the literature on externalities associated with traffic congestion, emotional cues and the determinants of crime. Several papers find a negative impact of traffic on psychological health, anger and stress. Kahneman et al. (2004) show that commuting is one of the least enjoyable daily activities. Understanding this negative fundamental part of urban life has been the focus of several papers studying the relationship between traffic and psychosocial health. The general finding is that traffic is associated with worse psychological health and higher levels of anger and stress. Gee and Takeuchi (2004) is one of the first papers to establish a link between self-reported traffic stress and perceived physical and psychological health conditions. Gottholmseder et al. (2009) improve on the statistical methodology and find a relationship between commuting features, including travel predictability, and self-reported stress. More recent work by Künn-Nelen (2016) shows that while self-reported commuting times have an impact on self-reported health outcomes and doctor visits, there is little effect of commuting time on objective health outcomes. Both Roberts et al. (2011) and Künn-Nelen (2016) find that the effect of commuting on health predominantly manifests itself in women as opposed to men. Stutzer and Frey (2008) shows that panel respondents in Germany who have longer daily commutes report lower levels of subjective well-being. Incorporating observed traffic data with subjective well-being data in China, Anderson et al. (2016) show that the estimated costs of congestion greatly exceed typical estimates that account for lost time and reliability. In the psychology literature, traffic is shown to be associated with increased anger and aggression (Parkinson, 2001; Hennessy and Wiesenthal, 1999).

We build on this literature in several ways. First, we link observed traffic data with an observed stress-related outcome. Most of the traffic data in the existing literature relies on survey data that only captures a self-reported snapshot of traffic conditions. This mutes most of the time series variation in actual traffic conditions. Our traffic data are built on a rich panel of hourly data from different roads and directions that enables us to provide a representative depiction of actual traffic conditions. Similarly, most of the physical and psychological health effects are also based on self reported data. Conversely, our measure of the psychological costs of traffic data relies on observed crimes from police incident reports. Therefore, we significantly advance the literature on the psychological costs of traffic congestion.

This article also fits into a broad literature investigating negative externalities to traffic. The largest traffic externality is likely the value of time and fuel expenditures associated with congestion. Schrank et al. (2012) estimates these two categories cost U.S. commuters $121 billion in 2011. The economics literature has also quantified several other externalities of traffic. Ossokina and Verweij (2015) exploits a quasi-experiment that reduced traffic congestion on certain streets in the Netherlands and find that the decrease in traffic led to an increase in housing prices. Currie and Walker (2011) show that traffic reductions due to the introduction of electronic toll collection (E-ZPass) reduce vehicle emissions near highway toll plazas, which subsequently reduces prematurity and low birth weight among mothers near a toll plaza. In addition to negatively affecting infant health, Anderson (2015) uses quasi-random variation in wind direction to show that traffic has a long run effect of increasing mortality within the elderly population. Quantifying the total economic cost of traffic congestion is important when deciding how to optimally manage congestion. For example, Gibson and Carnovale (2015) show that tolling not only reduces traffic but also leads to lower levels of air pollution. Another strand of the literature focuses on policies to reduce externalities to traffic such as congestion pricing through dynamic tolling (e.g. De Borger and Proost (2013); Brent and Gross (2017); Bento et al. (2017)).

Our paper is also related to the literature on emotional cues and their impact on economic outcomes. Card and Dahl (2011) study the link between family violence and the emotional cues associated with wins and losses by professional football teams. They use police reports of violent incidents on Sundays during the professional football season in the United States. They find that upset losses (defeats when the home team was predicted to win by four or more points) lead to a 10% increase in the rate of at-home violence and the impact is larger for important games. While Card and Dahl (2011) establish an important finding, there are potentially fewer policy levers to address unexpected football losses compared to managing traffic congestion. Additionally, there are a limited number of football games whereas traffic is a daily concern for many urban residents. There are several related studies on emotional cues and economic outcomes. Eren and Mocan (2017) find that criminal sentences set by Louisiana judges for juvenile crimes are harsher following an unexpected loss by the local university's football team. Duncan et al. (2016) shows that emotional cues due to Super Bowl exposure is associated with a small, but precisely estimated, increase in the probability of low birth weight.[6] This is relevant for consumers' experience with traffic because research on dynamicly priced toll lanes shows that drivers have a larger valuation for increased reliability (Brent and Gross, 2017) and on-time arrival (Bento et al., 2017).

Our research also fits into the literature that studies the determinants of crime. Research shows that crime is affected by many different factors. For example, Schneider et al. (2016) find that domestic violence is affected by negative labor market conditions. Cui and Walsh (2015) show that following a vacant home foreclosure there is an increase in violent crime and a smaller increase in property crime. Ranson (2014) finds that weather and climate change affect crime; temperature has a strong positive effect on criminal behavior, with little evidence of lagged impacts. Herrnstadt et al. (2016) estimate the causal effect of pollution on criminal activity in Chicago and Los Angeles and find that air pollution increases violent crime in both cities. A related literature in other social sciences documents that stress increases domestic violence (e.g. Romero-Martínez et al. (2013), Riggs et al. (2000), and Umberson et al. (2003)). We document that traffic is an additional mechanism whereby

[6] There is a related literature documenting the changes in stress and behavior following a dramatic event. For example, the emotions associated with tragic events have been shown to affect birth outcomes and student performance. For birth outcomes see Eskenazi et al. (2007) following the September 11th terrorist attacks, and Currie and Rossin-Slater (2013) following Hurricane Katrina. For student performance, see Beland and Kim (2016) after a shooting in a high school and Imberman et al. (2012) after a hurricane.

stress can lead to crime.

3 Data, descriptive statistics and traffic conditions

3.1 Data sources

Data on crime in Los Angeles come from police incident reports from two sources: the Los Angeles Police Department (LAPD) and the Los Angeles Sheriff Department (LASD). The LAPD police reports represent all crimes that take place in the City of Los Angeles and were accessed via the Los Angeles Open Data website.[7] The LAPD data are available from 2011 to 2015 and contain information on the date, time, location and type of crime. The LASD police report data are obtained through Los Angeles County GIS Data Portal, and contain data for all crimes in the LASD jurisdiction.[8] The LASD serves 40 incorporated cities and all unincorporated areas of Los Angeles County. These two datasets represent the vast majority of crime in Los Angeles County.[9] We consider the following crimes: assault, domestic violence, property crime, homicides and all crimes. We control for weather that could affect both crime and traffic by collecting daily data on rain, maximum temperature and wind speed in Los Angeles from the National Oceanic and Atmospheric Administration's (NOAA) National Centers for Environmental Information.[10]

The traffic data for Los Angeles are obtained from the California Department of Transportation through the Caltrans Performance Measurement System (PeMS).[11] We access annual Station Hour datasets from 2011 to 2015 from the PeMS data clearinghouse for California District 7, restricting the stations to Los Angeles County. We focus on two major roads, I-10 and I-5, that represent primary north-south and east-west routes to downtown Los Angeles. These datasets contain over 22 million observations of hourly speeds from 543 unique stations in Los Angeles County for the two major interstates in our analysis. In order

[7] Data are available at https://data.lacity.org/ by searching for "LAPD Crime and Collision Raw Data".

[8] The LASD data are available from 2005, but we only use 2011-2015 to match with the LAPD data. The LASD crime data are accessed at: http://egis3.lacounty.gov/dataportal/2012/03/05/crime-data-la-county-sheriff/.

[9] The maps for the LAPD and LASD jurisdiction are available from the Los Angeles Times: http://maps.latimes.com/lapd/ and http://maps.latimes.com/sheriff/.

[10] Weather data are available at: https://www.ncdc.noaa.gov/cdo-web/datasets.

[11] The data can be accessed via http://pems.dot.ca.gov/. A free account needs to be established.

to capture typical commuting patterns, we utilize the Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics for 2014 from the United States Census.[12] We also gather data on the location of metro stations from the Los Angeles Open Data website and average zip code-level income data from the U.S. Census American Community Survey.[13]

3.2 Dataset creations and descriptive statistics

To measure the impact of traffic on crime, we assign each zip code a daily time series of traffic from the closest major highway that connects to the downtown area. We focus on two major roads, I-10 and I-5, that represent primary north-south and east-west routes to downtown Los Angeles. While these are not the only means of transportation in the Los Angeles Metro area, they are likely to be correlated with traffic on other nearby routes in the same direction. The process of creating the traffic facing each zip code requires three steps. First, we assign each zip code to the closest route (I-5 or I-10) based on driving distances from the zip code centroid to the nearest on-ramp. Second, we use the LEHD Origin-Destination data to construct the typical destination zip codes for each origin zip code. Lastly, we construct the daily average travel times for the morning and evening commutes based on all the stations between the origin and destination zip codes. We elaborate on each step below.

In order to assign each zip code to the nearest road, we calculate the driving distance from the zip code centroid to the nearest on-ramp on both I-10 and I-5 using the ArcGIS Network Analyst tool and the road network for Los Angeles County obtained from the Los Angeles County GIS Data Portal. After determining the closest route to the zip code centroid, we find the nearest traffic station to the zip code centroid, which we refer to as our origin station. We exclude all zip codes where the closest route is more than 4 miles from the zip code centroid because workers in these zip codes are unlikely to commute via either I-10 or I-5. Figure A.1 presents the mapping of zip codes to roads.

We use the LEHD Origin-Destination data to calculate the destination station for each zip code. Since the LEHD data is organized at the census block level, we first use the

[12] The LEHD data is available at https://lehd.ces.census.gov/data/#lodes. We used the LODES 7.3 version.

[13] Data on metro stations are available at https://data.lacity.org/A-Livable-and-Sustainable-City/Los-Angeles-County-Metro-Rail-Station-Portal-Locat/s2k2-nqiy.

geographic crosswalk to convert to zip codes. Next, we aggregate all unique trips in each origin zip code to calculate the proportion of trips made to each destination zip code. Then we order the set of trips from the most common to the least common destination zip codes. The distribution of destination zip codes has a fat right tail - there are many destination zip codes with only a few commuters from any given origin zip code - so we consider the top 75% of destinations as the set of typical commuting patterns.[14] For every destination zip code we find the closest traffic station on the same route as the origin zip code. We exclude all destination zip codes that are more than four miles from the route assigned by the origin zip code. The furthest traffic station matched to a destination zip code serves as the end point of the route, which we refer to as the destination station.

Traffic conditions from all stations between the origin and destination zip codes comprise the daily time series of traffic. Congestion is primarily driven by the morning and evening commuting period, henceforth referred to as AM and PM traffic, so we focus on these peak traffic conditions.[15] Traffic stations are defined by a given location along a given route in a given direction. For AM traffic we use stations on lanes directed towards downtown and for PM traffic we use stations on lanes directed away from downtown.[16] Our final traffic variables are the daily travel times for each commuting period (AM or PM) for each zip code. This is constructed by dividing the length of each station in miles by the average speed and summing over all the stations between each zip code's origin and destination stations. This setup uses both the time series and spatial variation of traffic conditions, such that each zip code has a unique time series of traffic data.

Additionally, since not all workers will face the traffic that we assign them, we use the LEHD data to calculate the proportion of commuters that travel along their assigned route in each zip code. We assume all workers with the same origin and destination zip code are not commuting along the assigned route because they are either self-employed and/or do not face a significant commute. For the sample as a whole, roughly 41% of the commuters use the route that we assign them. The final traffic dataset is a panel of daily traffic observations

[14] As a robustness check we also use several other thresholds (top 25% top 50%, top 90%, top 95%, and top 99%) to determine the set of typical destination zip codes. The results are robust to alternative choices.

[15] The morning period is defined as 5:00-9:00 AM and the evening period is defined as 3:00-7:00 PM.

[16] For PM traffic the origin and destination stations are reversed to capture the commute from work to home.

for each zip code.

Our main measure of high traffic is when the evening traffic is above the 95th percentile for a given zip code in a given day of the week. We also test for several other thresholds for high traffic. Since our traffic variable is based on typical commuting patterns, and domestic violence often occurs in the home, our primary specification focuses on the effect of evening traffic on evening crime. We explore several specifications of the timing and spatial assignment of traffic including placebo tests of traffic that occurs after a crime is committed.

Figure 1 visualizes the zip codes, regions, and routes that comprise our traffic sample. Regions are defined by the direction relative to downtown. In each region every zip code has a unique starting and ending point that could spill over into other regions. For example, commuters in a zip code north of downtown may in fact work in zip codes south of downtown.

Figure 2 presents the traffic congestion by route, direction and time of day. The goal of the graph is to visualize average travel times and variation over time, so we select one representative zip code for each region of the city. This captures the time series variation in travel times by hour of day. Using the whole sample pools both time series variation and cross sectional variation over zip codes. The solid line represents average travel times and the dashed lines are plus and minus one standard deviation. The direction corresponds to the route traveling towards downtown, and each zip code is assigned traffic from two routes. For example 10:E are zip codes that are west of downtown and therefore travel to work (AM traffic) on I-10 East and return home (PM traffic) on I-10 West. Figure 2 shows both morning and evening peaks; regions of the city differ with respect to the severity of morning and evening peaks. Additionally, the northern and western regions of LA County extend further out along our routes of interest, and therefore have longer average commutes. The average travel time on I-5 or I-10 varies across zip codes and we explore specifications that exclude zip codes with short and long average commutes. Table A.1 presents key descriptive statistics (mean, standard deviation, minimum and maximum) for daily average traffic in AM and PM for zip codes in our data set. It shows a mean travel time on I-5 or I-10 of 24 minutes in the morning and 26 minutes in the evening.[17]

The crime data provide a fine spatial and temporal resolution, and in order to match the

[17]Door-to-door commuting time for a zip code is longer, we are only measuring time spent on I-5 or I-10.

crime data to the traffic data, we aggregate all crimes within a zip code over the course of the day in each of the categories to obtain our dependent variables.[18] All crimes that occur before 5 AM are assigned to the previous day and are coded as nighttime crimes. This allows crimes committed prior to 5 AM to be affected by traffic in the previous day. For example, we assume that a crime committed at 1:00 AM can be influenced by getting stuck in traffic on the way home from work. We also differentiate between evening and morning (AM and PM) crimes such that we can ensure that a crime takes place *after* commuters experience their assigned traffic. Figure 3 presents a map of average daily domestic violence incident by zip code in Los Angeles. The maximum daily average domestic violence incidents is 1.8 and many regions have an average of 0.4 and below. Figures A.2, A.3 and A.4, in appendix, present similar figures for all crimes, property crimes and assault in Los Angeles. The figures show that some zip codes have large average daily crime incidents, reaching above 20 daily crimes for some areas. Table A.2 presents key descriptive statistics for daily average crimes by zip code for total crime and evening crime. It shows, for example, that the daily average number of crimes committed in a zip code is 3.57 while the average incidence of domestic violence is 0.14. There are an average of 2.5 total crimes and 0.1 incidents of domestic violence when focusing on crimes committed in the evening.

3.3 Traffic in Los Angeles

Traffic in Los Angeles is a severe problem. According to a Texas A&M transportation Institute report, drivers in Los Angeles spend on average 80 hours or 3.5 days a year in gridlock.[19] Los Angeles has the biggest difference between normal travel times and rush hour travel times in the United States. Rush hour can be 43 percent slower than non-peak hours. According to Sorensen (2009), congestion is due to the high population density of Los Angeles metropolitan region, and the fact that parking is cheap and abundant. Most drivers do not pay the full economic and social costs of driving. A recent Los Angeles Times poll shows that traffic is the top concern for Los Angeles residents, topping personal safety,

[18] The crime data are available at finer spatial resolutions than zip codes, but when aggregating up subcategories, such as domestic violence, there is little daily variation in the crime data due to a mass of zeros.

[19] See the Annual Urban Mobility Scorecard report from Texas A&M Transportation Institute available at: http://mobility.tamu.edu/ums/

personal finances and housing costs.[20]

4 Methodology

To estimate the impact of traffic on domestic violence, we rely on deviations from normal traffic to isolate the impact of abnormally high traffic on crime. Following Card and Dahl (2011), we estimate the following Poisson count model:

$$\begin{aligned} Crime_{it} = \beta_0 + \beta_1 HighTraffic_{it} + \beta_2 E[Traveltime] + \\ \beta_3[Weather] + \beta_{MY} + \beta_D + \beta_Z + \delta_1 t + \delta_2 t^2 + \epsilon_{it} \end{aligned} \tag{1}$$

where $Crime_{it}$ is the number of domestic violence incidents in zip code i on day t. Our main analysis focuses on domestic violence but we also examine other types of crime (property, assault, homicide and all crimes). $HighTraffic_{it}$ is an indicator variable equal to one when traffic exceeds the 95th percentile of travel times for a given zip code and day of the week. E[Traveltime] is a measure of expected travel time for a given zip code. It contains information on the average traffic for the last week and month for a given zip code. Our coefficient of interest, β_1, measures the impact of traffic above the 95th percentile of traffic on our outcomes. [Weather] is a vector of weather covariates: rain, maximum temperature and wind. We use zip code level fixed effects (β_Z) to control for static spatial unobserved effect. To control for time-varying unobservables we include year-by-month (β_{MY}) and day-of-week (β_D) fixed effects as well as a quadratic time trend. We cluster the standard errors at the zip code level.

To better understand the relationship between traffic and crime, we also estimate the same model using different thresholds for extreme traffic. We also consider the possibility that traffic has heterogeneous effects on crime. First, we investigate how the timing of traffic (morning vs. evening traffic) impacts crime, and the persistence of traffic shocks on domestic violence. To ensure the validity of our results, we perform several robustness checks. We run placebo regressions of traffic on lagged domestic violence incidents, estimate randomized

[20] See the poll at http://www.latimes.com/local/lanow/la-me-ln-traffic-still-tops-crime-economy-as-top-l-a-concern-poll-finds-20151007-story.html.

traffic on domestic violence, relax the spatial assumptions regarding traffic assignment, and employ alternative estimation methods such as instrumental variables and ordinary least squares (OLS).

5 Results

5.1 Main Results

Table 1 presents the impact of high traffic (above 95th percentile) on domestic violence. Our primary estimates use traffic during the evening commute and domestic violence incidents in the evening. Column (1) uses all traffic and crime observations in Los Angeles to estimate equation (1), and shows that domestic violence is significantly higher when traffic exceeds the 95th percentile. Given that we estimate a Poisson model the coefficients can be interpreted as the approximate percentage change in crime when traffic exceeds the 95th percentile.[21] Columns (2)-(4) focus on subsamples with stronger links to the assigned travel time based on typical commuting patterns. Column (2) investigates the impact of high traffic on domestic violence, excluding the downtown area. We exclude crimes that occurs downtown since these crimes are either not affected by the typical commuting pattern, or we cannot be sure which roads the offender used. Column (3) excludes weekends and holidays, since the conventional commuting patterns do not hold and traffic is inherently less predictable. Column (4) excludes downtown zip codes, weekends and holidays. The results of column (4) show that traffic above the 95th percentile leads to an increase in domestic violence of approximately 9%. All specifications show that there is significantly more domestic violence when there is high traffic. The effects are larger when restricting the sample to those areas and time periods that represent conventional commuting behavior. Our preferred specification is column (4), which we use for the remainder of the paper except when specified. All subsequent regressions also include zip code, year-by-month and day-of-week fixed effects, as well as a quadratic time trend, weather variables and recent traffic controls as described in Section 4.

Next, we examine how the estimates change for different thresholds defining high traffic.

[21] More precisely, we estimate the change in the log of the expected count when traffic exceeds the 95th percentile.

We run separate regressions where we define high traffic as an indicator based on traffic exceeding each percentile from the 70th to the 99th. Figure 4 shows the regression results, where each bar represents the coefficient estimate of that percentile indicator, and the error bars are 95% confidence intervals generated from the robust standard errors clustered at the zip code level. Each coefficient is estimated using our preferred specification shown in column (4) of Table 1. The effect of high traffic is small and insignificant when defining the threshold from the 70th percentile to the 84th percentile. For thresholds above the 85th percentile the coefficients become statistically significant and generally increase in magnitude as the percentile threshold increases. The effect at the 90th percentile is 8.6%, the effect at 95th percentile is 9.4% and the effect at the 98th percentile is 11.0%.[22] This is consistent with threshold effects, where only traffic in the right tail causes the necessary stress to induce domestic violence. The damages from traffic can also experience thresholds effects. For example, drivers may account for certain levels of traffic when commuting, but extreme traffic will cause them to be late or miss important appointments.

Table 2 estimates the effect of traffic on different types of crimes. Column (1) replicates our preferred specification using domestic violence as the outcome variable. In column (2), we regress all crimes on traffic, and find no significant effects. Column (3)-(5) present results using assaults, property crimes, and homicides as the outcome variables. Since domestic violence is categorized as an assault, we remove domestic violence incidents from the assault category. There are no significant effects for assaults, property crime, or homicides.[23] The results show that traffic predominantly impacts crimes where stress is a contributing factor (domestic violence) as opposed to other crimes, which is consistent with a model of emotional cues where traffic shocks increase psychological stress. It is unlikely that the psychological stress from traffic would cause an increase in robberies or homicides.

[22]The 99th percentile is of similar magnitude as the 98th, but has very large standard errors because there are relatively few data points to estimate the parameter.

[23]For all crimes except homicides the results are relatively tight zeros as indicated by standard errors that are roughly half the size of the standard errors on the coefficient for domestic violence. The coefficient of the effect on homicides is reasonably large, but is very noisy, because homicides are very rare events.

5.2 Heterogeneity

Next, we examine heterogeneous impacts by the timing of traffic. In particular, we examine the effect of both morning (AM) and evening (PM) traffic shocks on domestic violence. Column (1) of Table 3 presents our baseline specification of the effect of a PM traffic shock on domestic violence in the evening. Column (2) shows that AM traffic shocks have no significant effect on domestic violence incidents. Column (3) presents a model that includes separate AM and PM indicators; only PM traffic has a significant impact on domestic violence. As a placebo test, in column (4) we estimate the effect of traffic in the evening on domestic violence that occurs in the morning. Reassuringly, we find no effect of traffic later in the day on crime in the morning.

Some commuters in each zip code do not travel by the route that we assign them; therefore, only a portion of the population in each zip code is "treated" by their assigned traffic shock. To account for this feature of the data we incorporate the fraction of commuters into our regression model. The fraction of commuters is calculated at the zip code level using the LEHD Origin-Destination data and represents the fraction of workers that are likely to use the route assigned to them as described in Section 3. Column (1) of Table 4 multiplies the 95th percentile indicator by a continuous and static variable for the fraction of commuters in the zip code, effectively scaling the traffic shock by traffic exposure. The scaled traffic shock has a much larger effect on domestic violence compared to our baseline specification. One way to interpret this coefficient is the effect of high traffic on domestic violence if everyone in the zip code commuted by the route that we assign to the zip code. Multiplying the average fraction of commuters by this coefficient ($0.42 \times 0.234 = 0.098$) approximates our baseline estimate. Columns (2) and (3) estimate our preferred specification using sample splits, based on zip codes that are above and below the median percentage of commuters, respectively. We find a significant impact of traffic on domestic violence in zip codes with a high fraction of commuters on their assigned routes (column (2)), but not in zip codes below the median (column (3)). Therefore, the results in Table 4 suggest that the effect is concentrated in zip codes where more commuters travel by their assigned routes.[24]

[24] Given that zip codes with relatively small fractions of correctly assigned commuters may still be affected by the traffic on their actual commute our estimates may be considered a lower bound.

We investigate the persistence of the effect by examining if high traffic leads to an increase in domestic violence in the days following the traffic event. We estimate traffic at time T on domestic violence in the following four days (on time $T + h$, where $h = 1, 2, 3, 4$). Column (1) of Table A.3 presents results at time T and replicates our preferred specification of Table 1. Columns (2), (3), (4) and (5) present results at time $T + 1$, $T + 2$, $T + 3$, and $T + 4$. The impact of high traffic carries forward to the following day (T+1) but not after that (T+2, T+3, or T+4). Column (6) uses domestic violence incidents at either time T or $T + 1$, as an outcome variable, showing the cumulative impact of high traffic on domestic violence. Including domestic violence incidents that occur the day after an extreme traffic event leads the impact of traffic on crime to increase slightly to 10.3%.

In order to assess the policy levers available to mitigate the effect of traffic on domestic violence, we examine how access to public transportation affects our results. Table A.4 studies how the proximity to a metro station moderates the effect of traffic on domestic violence by limiting the sample to zip codes within 1-3 miles from a metro station. The results once again indicate that high traffic leads to an increase in domestic violence, and that the effect is not reduced by access to public transportation.[25] We also examine heterogeneity in the effect of traffic on crime by dividing zip codes along three dimensions: average crime, average income and distance to downtown, using interaction terms. The results, reported in Table A.5 in the Appendix, provide suggestive evidence that traffic has a stronger link to domestic violence in zip codes that have lower income, higher average crime rates, and are closer to downtown. However, none of the interaction effects are statistically significant from zero.

5.3 Robustness

We perform several robustness checks to test the validity of the results. Our first exercise is a temporal placebo test where we regress crime in previous days ($T - 1$, $T - 2$, $T - 3$, and $T - 4$) on traffic information at time T. High traffic at time T should not affect crime in previous days. Table 5 shows that high traffic at time T has no significant impact on

[25]The effects are actually larger in zip codes close to a metro stations. This result is likely driven by the fact that these zip codes are closer to downtown, and Table A.5 suggests that impact of crime on domestic violence is concentrated in zip codes closer to downtown.

domestic violence incidents in previous days. This is reassuring as high traffic at time T leads to higher incidence of domestic violence at time T, but no increase is observed in the days prior to the high traffic (above the 95th percentile).

Next, we perform a spatial placebo test (similar to a permutation test) where we randomize traffic assignment and investigate the impact of this randomized traffic on domestic violence. For each zip code, we randomly select traffic from a zip code in an alternative region of the city (as defined in Figure 1) that has a different commuting pattern and then estimate the impact of "randomized false traffic" on domestic violence. We repeat this process 500 times and each time we set a new seed that generates the random number used to assign a zip code's randomized false traffic. This tests for the possibility that the result is driven by temporally correlated shocks not related to the specific commuting pattern. Figure 5 shows the distribution of treatment effects for randomized traffic as well as our baseline estimate. The randomized estimates are centered around zero, while our baseline estimate is in the right tail. The mean of the placebo estimates is 0.0019, and 14 out of the 500 estimates have a p-value less than 0.05 - slightly less than would be predicted by random chance. Figure 5 gives confidence that our results on domestic violence are due to exposure to high traffic.

Table 6 presents several additional robustness checks. Column (1) of Table 6 excludes zip codes with the shortest average commute time in our sample (shortest 10%), and column (2) excludes zip codes with the longest average commute time (longest 10%). Columns (1) and (2) show that our results results hold after excluding those zip codes. Column (3) removes the controls for weather covariates and column (4) removes the days before Thanksgiving, Christmas and New Years (Lag Holidays), which are known to have different traffic patterns. In columns (3) and (4), the results are qualitatively similar to our main results in Table 1. Column (5) restricts the sample to zip codes where the assigned route is at least four miles closer to the zip code centroid than the alternate route. These zip codes are more likely to be assigned the appropriate traffic conditions. This will exclude zip codes that are roughly equidistant to I-5 and I-10. Once again, the results are very similar to our baseline specification. Column (6) present results that control for date fixed effects instead of month-by-year fixed effects, and the results are qualitatively the same. Across all the specifications the results are quite similar. Column (7) tests for contemporaneous cross-city conditions.

This specification replaces the 95th percentile dummy with an indicator variable that takes a value of one on days when a zip code did not experience extreme traffic, but there was extreme traffic in one of the other regions in the city. Column (7) shows, reassuringly, that this measure has no significant impact on domestic violence.

Since the Poisson regression is sensitive to specification error when using many fixed effects, we also estimate several OLS specifications that are presented in Table 7. Column (1) uses the same controls and restrictions as our baseline specification. Column (2) presents results without weather controls, column (3) presents results without the days before Thanksgiving, Christmas and New Years, and column (5) controls for date fixed effects. Column (5) investigates the potential for the error terms to be correlated across time within a zip code and across zip codes within a month by using two-way clustered standard errors at the zip code and month level. Columns (1)-(5) of Table 7 show qualitatively the same results. The OLS results are not based on the log of expected counts (as the poisson model), but rather on the raw number of domestic violence incidents. When considering these results in the context of the average evening domestic violence rate of 0.10, the results are similar in percentage terms.

We think that, conditional on all of the fixed effects and previous traffic conditions, extreme traffic is quasi-random. However, we also exploit accidents as an additional source of quasi-random variation in traffic conditions. Table 8 presents instrumental variable estimates using the number of accidents causing traffic delays above one hour or the duration of these accidents as our instruments.[26] While traffic and accidents are correlated we believe that the timing and location of severe accidents exploits one specific source of quasi-random variation in our empirical strategy. The results indicate that high traffic due to accidents leads to an increase in domestic violence. The IV coefficients are slightly larger (15.2% and 15.6%, respectively) than the baseline model and are significant at the 10% level.

Our last set of robustness checks relax the assumptions that allow us to match zip codes to traffic conditions. Columns (1)-(5) of Table A.7 restrict the sample to zip codes that are

[26] A regression of hourly traffic on accidents, presented in Table A.6 of the Appendix, shows that there is a contemporaneous effect of accidents on travel times and that the results are persistent for roughly two hours. The accident data are collected from Caltrans. Our instrumental approach includes controls for weather, expected traffic and all fixed effects included in our baseline specification. The F-test of the first stage is 39.39 and 10.38, respectively.

within two to four miles to the closest on-ramp. Households in these zip codes are more likely to use the roads that we assign to them. Once again we find that high traffic leads to increases in domestic violence incidents. The effects of traffic on domestic violence in these specifications are larger (9.4% to 13.1%). These zip codes represent less stringent assumptions regarding the typical commuting patterns, indicating that our preferred specification may in fact be a lower bound. Table A.8 presents different methods for assigning the end of a trip from a given zip code using the LEHD Origin-Destination data. As discussed in the data section, our main estimates use the top 75% of destinations as the set of typical commuting patterns. In Table A.8, we use several other thresholds (top 25% top 50%, top 90%, top 95%, and top 99%) and the results are qualitatively the same. Overall, the results are robust to many alternative specifications, which provide confidence that high traffic leads to an increase in domestic violence in Los Angeles.

5.4 Expected vs. Unexpected Traffic

We next explore the role of expectations and habituation. The concept that unexpected traffic shocks affects domestic violence is consistent with a model of emotional cues. Although we control for traffic expectations, we cannot be certain that results are due to *unexpectedly* high traffic as opposed to simply high traffic (both expected and unexpected).

We first incorporate data on daily Google search volume for the term "traffic" in Los Angeles County.[27] Google search volume may reflect an additional source of traffic expectations since drivers can easily find out traffic conditions through an internet search. Additionally, search volume for the whole county likely reflects traffic conditions on more roads than those considered in this paper. Therefore, current and lagged search volume may serve as both a control for traffic expectations and an alternative measure of traffic for drivers who do not commute by I-5 or I-10. The regressions incorporating standardized Google search volume (mean zero and standard deviation of one) are presented in Table 9.

Column (1) simply replicates our baseline specification and column (2) adds the daily standardized Google search volume to the baseline regression model. Days that have higher

[27] Search volume is available weekly for periods over three months and daily for periods less than three months. We download three month intervals of the daily search volume and re-normalize the search volume based on overlapping dates.

search volume also have higher incidents of domestic violence; the base effect of actual traffic conditions remains largely unchanged. Next, we include lags of search volume to account for expectations formed over recent days, which is presented in column (3). The effect on the 95th percentile of traffic slightly increases to 10.4% with lagged search volume, and the coefficient on contemporaneous search volume decreases and is no longer statistically significant. We also examine threshold effects in column (4) by including an indicator for search volume above the 95th percentile. The 95th percentile of search volume has a similar magnitude to our traffic variable, and the coefficient on actual extreme traffic is largely unchanged. Column (5) includes search volume but not actual traffic, and Column (6) extends this specification to include zip codes that are further than four miles from our defined roads (I-5 and I-10). In both regressions the impact of search volume on domestic violence is positive and statistically significant.

When interpreting the results that include search volume, it is important to consider the correlation between our measure of traffic and search volume. In terms of extreme traffic, only 8% of observations exceed the 95th percentile of both travel times and Google search volume.[28] Put differently, there are not many days where we define extreme traffic that also have very high search volume. The results serve as an additional robustness check and another attempt to model expectations. The impact of observed traffic is robust to a variety of specifications that incorporate traffic search volume, which strengthens the validity of the base results.

We perform several tests to attempt to disentangle the role of expectations. In column (1) of Table 10, we remove recent traffic experience in the last week and last month as controls, which represents our measure of traffic expectations. We find that the coefficient is positive and statistically significant but smaller than our baseline estimate (7.4%), but the differences are not statistically significant. Next, we divide the sample based on zip codes where traffic is more or less variable as measured by the standard deviation of travel times within the zip code. Extreme traffic in zip codes with a high standard deviation of travel times is more likely to be due to unexpected shocks. Columns (2) and (3) show the baseline specification for zip codes above and below the sample median of travel time standard deviation. Traffic

[28] The correlation between Google trends and our travel times variable is 0.14.

above the 95th percentile leads to a significant increase in domestic violence in both samples, and the coefficient is larger for zip codes with large traffic deviation (11% vs 7.8%), although this difference is not statistically significant.[29]

As an alternative way to model drivers' traffic expectations, we estimate a moving average model to predict travel times based on traffic in the last five days along with all of the fixed effects and controls that are in our baseline specification. We use the predicted travel times to generate expected extreme traffic and unexpected extreme traffic. Expected extreme traffic is defined as an indicator if predicted travel times exceeds the 95th percentile and unexpected extreme traffic is an indicator if the residual from the moving average model exceeds the 95th percentile. Column (4) of Table 10 shows that the effect of predicted traffic is positive, but roughly half the magnitude of the baseline effects and not statistically significant. Unexpected traffic, reported in column (5) of Table 10, is similar in magnitude to our preferred estimate and statistically significant at the 5% level. In Table (6), we include both the predicted and unpredicted traffic and find similar results to column (4) and (5).[30] The results on the role of expectations provide suggestive evidence that expected high traffic does impact domestic violence, and that unexpected traffic increases the magnitude of the effect. We are cautious in this interpretation for two reasons. First, some of the differential effects are not statistically significant, and second, we cannot be certain that we are accurately modeling drivers' traffic expectations.

6 Conclusion

This paper investigates the psychological costs of traffic congestion by estimating the impact of high traffic on domestic violence. We combine traffic data in Los Angeles from 2011 to 2015 to police incident reports from the Los Angeles Police Department and the Los Angeles Sheriff Department. This rich dataset allows us to link traffic with criminal activity at a fine spatial and temporal dimension. Our identification relies on extreme deviations from normal traffic to isolate the impact of abnormally high traffic on domestic violence incidents.

We find that extreme traffic (above the 95th percentile) significantly increases the inci-

[29]Similar results are found when using an interaction term for whether the zip code standard deviation in travel times is above the sample median.

[30]Results are robust to other measures of moving average, such as last 10 days or last 30 days.

dence of domestic violence by approximately 9%. We control for unobserved effects across time and space with fixed effects, and time-varying measures of traffic in the most recent week and month to control for traffic expectations. Our results are consistent with a model of emotional cues, and are robust to several specifications and falsification tests. There is no effect of traffic on lagged crime, no effect of evening traffic on morning crimes, no effect of randomized traffic or traffic in other parts of the city on domestic violence, and no effect of traffic on other categories of crime such as property crime.

We estimate the aggregate economic cost of increased domestic violence by predicting the increase in domestic incidents due to high traffic events across Los Angeles County and multiplying by the dollar value of the social cost of domestic domestic violence, which is estimated to be $121,825 in 2017 dollars (McCollister et al. (2010)). Our estimates of the economic cost of traffic-induced domestic violence range from $5-22 million dollars per year depending on the specification. Our baseline specification generates $5 million in annual damages, which rises to $11 million when scaled by the fraction of commuters in a zip code. Using the different thresholds for high traffic generates annual damages as high as $22 million annually when scaled by the fraction of commuters. Since we expect that most people who suffer some psychological costs of traffic do not actually commit crimes we consider our estimates to be an extreme lower bound; they are the tip of the iceberg.

Documenting the psychological costs of traffic provides additional support for congestion management policies that not only reduce average travel times but improve reliability by reducing the variance of travel times. Building new capacity is unlikely to reduce congestion in the long-run since the elasticity of travel demand with respect to capacity is equal to one (Duranton and Turner, 2011). Alternatively, Peirce et al. (2013) document that drivers report less stress after time-of-day pricing was implemented on a major road in Seattle. Therefore, our research documents additional benefits of congestion pricing policies, but more research is needed on how different types of tolling structures improve travel reliability and driver satisfaction.[31] There are also implications for how resources are deployed after extreme traffic events. More police and/or counseling services should be available when there

[31] A technical report by the Washington State Department of Transportation shows that dynamically priced high-occupancy toll lanes reduce peak congestion in a road in metro Seattle (WSDOT, 2012).

is high traffic.

References

Anderson, Michael L, "As the Wind Blows: The Effects of Long-Term Exposure to Air Pollution on Mortality," *NBER Working Paper Series*, 2015, (21578).

_ , Fangwen Lu, Yiran Zhang, Jun Yang, and Ping Qin, "Superstitions, Street Traffic, and Subjective Well-Being," *Journal of Public Economics*, 2016, *142*, 1–10.

Beland, Louis-Philippe and Dongwoo Kim, "The effect of high school shootings on schools and student performance," *Educational Evaluation and Policy Analysis*, 2016, *38* (1), 113–126.

Bento, Antonio, Kevin Roth, and Andrew Waxman, "Avoiding Traffic Congestion Externalities? The Value of Urgency," Technical Report, Working Paper 2017.

Borger, Bruno De and Stef Proost, "Traffic externalities in cities: the economics of speed bumps, low emission zones and city bypasses," *Journal of Urban Economics*, 2013, *76*, 53–70.

Brent, Daniel and Austin Gross, "Dynamic Road Pricing and the Value of Time and Reliability," *Journal of Regional Science*, 2017.

Card, David and Gordon B Dahl, "Family violence and football: The effect of unexpected emotional cues on violent behavior," *The Quarterly Journal of Economics*, 2011, *126* (1), 103.

Cui, Lin and Randall Walsh, "Foreclosure, vacancy and crime," *Journal of Urban Economics*, 2015, *87*, 72–84.

Currie, Janet and Maya Rossin-Slater, "Weathering the storm: Hurricanes and birth outcomes," *Journal of Health Economics*, 2013, *32* (3), 487–503.

— **and Reed Walker**, "Traffic Congestion and Infant Health: Evidence from E-ZPass," *American Economic Journal: Applied Economics*, 2011, *3* (1), 65–90.

Duncan, Brian, Hani Mansour, and Daniel I Rees, "It's Just a Game: The Super Bowl and Low Birth Weight.," *Journal of Human Ressources*, 2016.

Duranton, Gilles and Matthew A Turner, "The Fundamental Law of Road Congestion: Evidence from US Cities," *American Economic Review*, 2011, *101* (6), 2616–2652.

Eren, Ozkan and Naci Mocan, "Emotional Judges and Unlucky Juveniles," *American Economic Journal:Applied*, 2017.

Eskenazi, Brenda, Amy R Marks, Ralph Catalano, Tim Bruckner, and Paolo G Toniolo, "Low birthweight in New York City and upstate New York following the events of September 11th," *Human Reproduction*, 2007, *22* (11), 3013–3020.

Gee, Gilbert C and David T Takeuchi, "Traffic stress, vehicular burden and well-being: a multilevel analysis," *Social Science & Medicine*, 2004, *59* (2), 405–414.

Gibson, Matthew and Maria Carnovale, "The effects of road pricing on driver behavior and air pollution," *Journal of Urban Economics*, 2015, *89*, 62–73.

Gottholmseder, Georg, Klaus Nowotny, Gerald J Pruckner, and Engelbert Theurl, "Stress perception and commuting," *Health Economics*, 2009, *18* (5), 559–576.

Heaton, Paul, "Sunday liquor laws and crime," *Journal of Public Economics*, 2012, *96* (1), 42–52.

Hennessy, Dwight A and David L Wiesenthal, "Traffic congestion, driver stress, and driver aggression," *Aggressive Behavior*, 1999, *25* (6), 409–423.

Herrnstadt, Evan, Anthony Heyes, Erich Muehlegger, and Soodeh Saberian, "Air Pollution as a Cause of Violent Crime: Evidence from Los Angeles and Chicago," Technical Report, mimeo 2016.

Imberman, Scott A, Adriana D Kugler, and Bruce I Sacerdote, "Katrina's children: Evidence on the structure of peer effects from hurricane evacuees," *The American Economic Review*, 2012, pp. 2048–2082.

Kahneman, Daniel, Alan B Krueger, David A Schkade, Norbert Schwarz, and Arthur A Stone, "A survey method for characterizing daily life experience: The day reconstruction method," *Science*, 2004, *306* (5702), 1776–1780.

Künn-Nelen, Annemarie, "Does commuting affect health?," *Health Economics*, 2016, *25*, 984–1004.

McCollister, Kathryn E, Michael T French, and Hai Fang, "The cost of crime to society: New crime-specific estimates for policy and program evaluation," *Drug and alcohol dependence*, 2010, *108* (1), 98–109.

Ossokina, Ioulia V and Gerard Verweij, "Urban traffic externalities: Quasi-experimental evidence from housing prices," *Regional Science and Urban Economics*, 2015, *55*, 1–13.

Parkinson, Brian, "Anger on and off the road," *British Journal of Psychology*, 2001, *92* (3), 507–526.

Peirce, Sean, Sean Puckett, Margaret Petrella, Paul Minnice, and Jane Lappin, "Effects of Full-Facility Variable Tolling on Traveler Behavior: Evidence from a Panel Study of the SR-520 Corridor in Seattle, Washington," *Transportation Research Record: Journal of the Transportation Research Board*, 2013, (2345), 74–82.

Ranson, Matthew, "Crime, weather, and climate change," *Journal of Environmental Economics and Management*, 2014, *67* (3), 274–302.

Riggs, David S, Marie B Caulfield, and Amy E Street, "Risk for domestic violence: Factors associated with perpetration and victimization," *Journal of clinical psychology*, 2000, *56* (10), 1289–1316.

Roberts, Jennifer, Robert Hodgson, and Paul Dolan, "It's driving her mad: Gender differences in the effects of commuting on psychological health," *Journal of Health Economics*, 2011, *30* (5), 1064–1076.

Romero-Martínez, Ángel, Marisol Lila, Patricia Sariñana-González, Esperanza González-Bono, and Luis Moya-Albiol, "High testosterone levels and sensitivity to acute stress in perpetrators of domestic violence with low cognitive flexibility and impairments in their emotional decoding process: A preliminary study," *Aggressive behavior*, 2013, *39* (5), 355–369.

Schneider, Daniel, Kristen Harknett, and Sara McLanahan, "Intimate partner violence in the Great Recession," *Demography*, 2016, *53* (2), 471–505.

Schrank, David, Bill Eisele, and Tim Lomax, "TTI's 2012 urban mobility report," *Proceedings of the 2012 annual urban mobility report. Texas A&M Transportation Institute, Texas, USA*, 2012.

Small, Kenneth A., Clifford Winston, and Jia Yan, "Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability," *Econometrica*, 2005, *73* (4), 1367–1382.

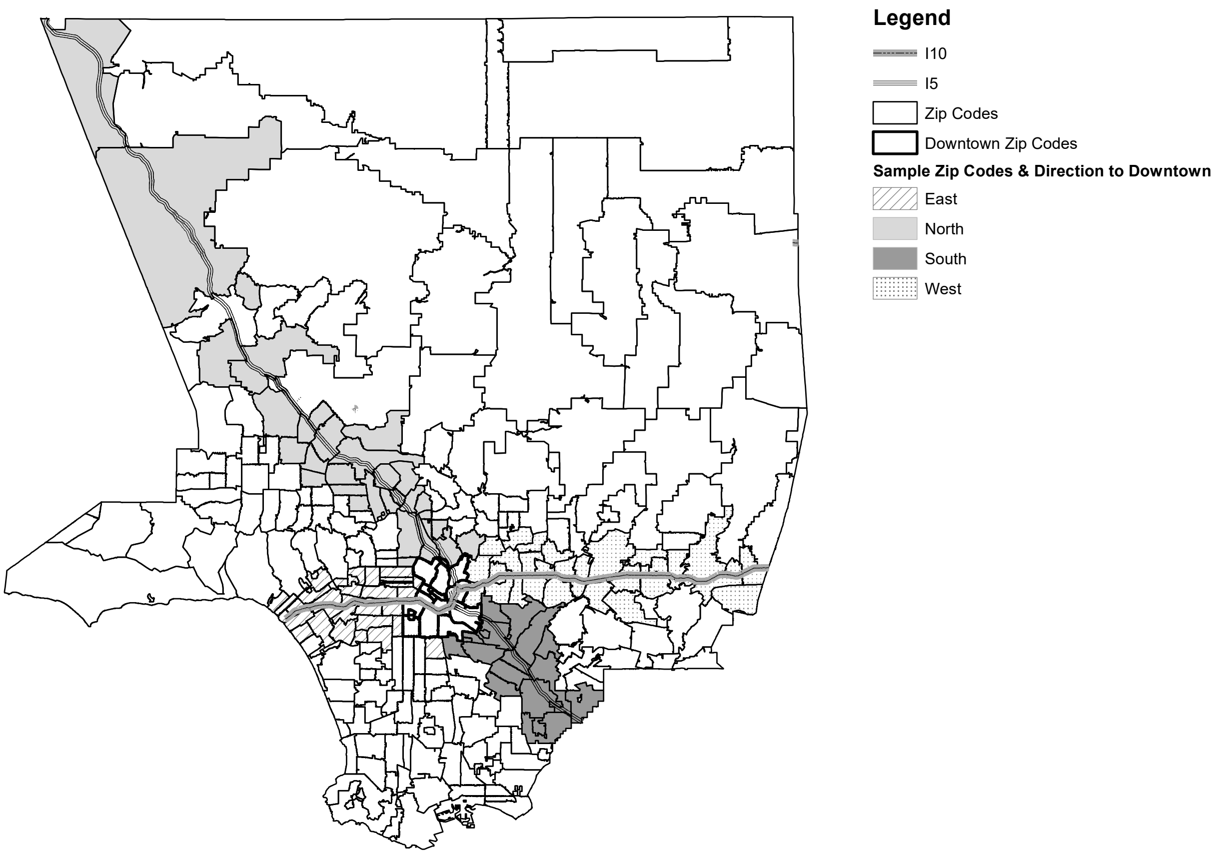
Sorensen, Paul, "Moving Los Angeles," *ACCESS Magazine*, 2009, *1* (35).

Stutzer, Alois and Bruno S Frey, "Stress that doesn't pay: The commuting paradox," *The Scandinavian Journal of Economics*, 2008, *110* (2), 339–366.

Umberson, Debra, Kristin L Anderson, Kristi Williams, and Meichu D Chen, "Relationship dynamics, emotion state, and domestic violence: A stress and masculinities perspective," *Journal of Marriage and Family*, 2003, *65* (1), 233–247.

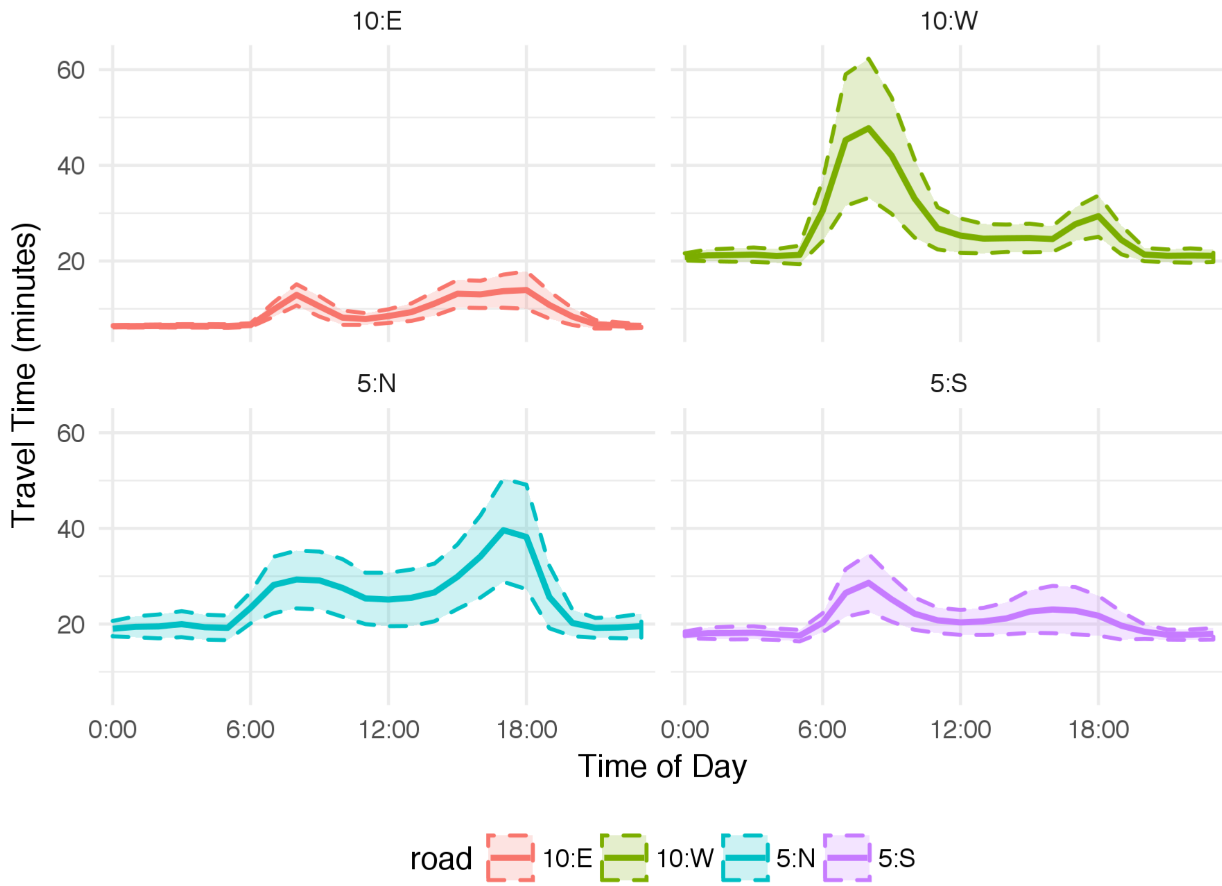
WSDOT, "SR 167 HOT Lanes Pilot Project Fourth Annual Performance Summary," Technical Report May 2008, WSDOT 2012.

Figure 1: Map of Sample



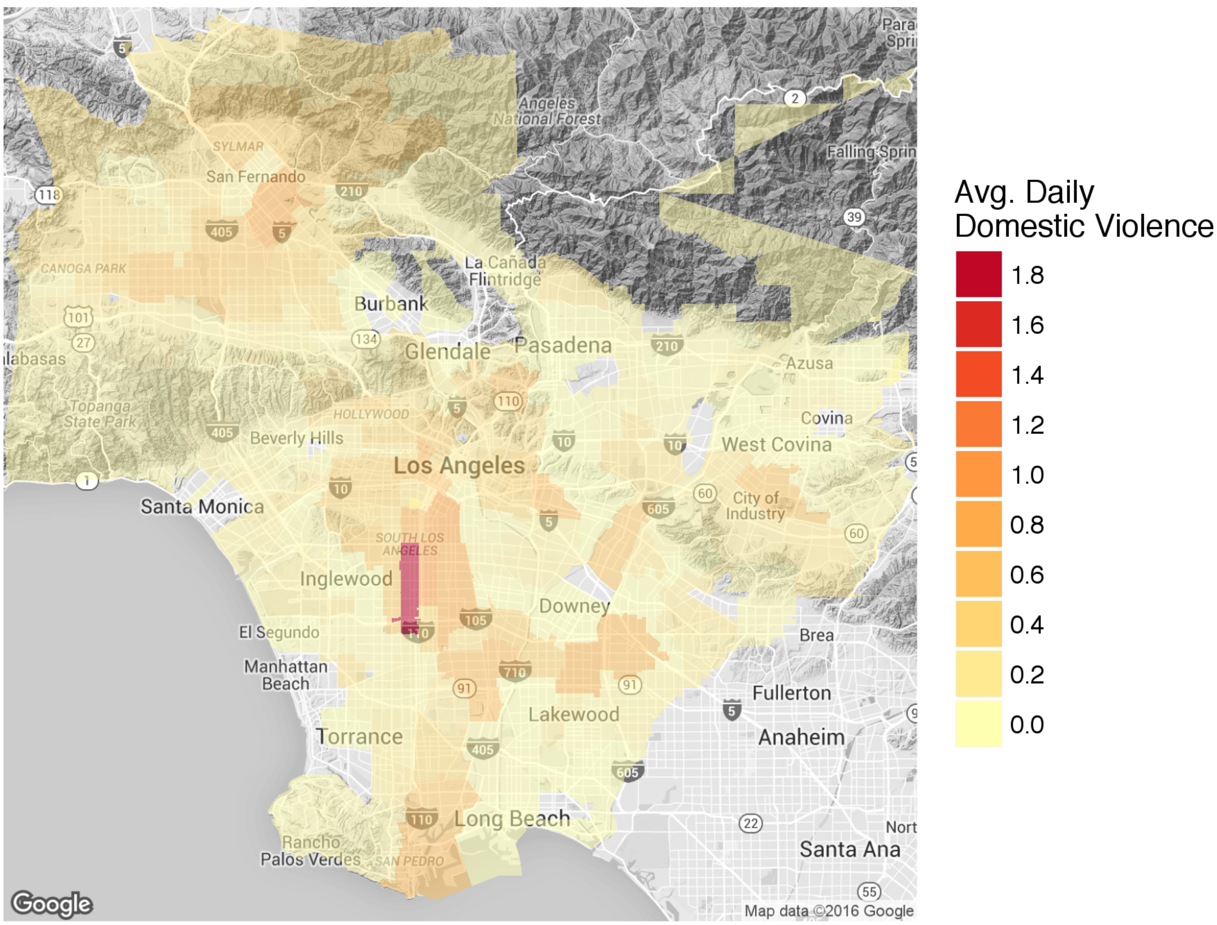
Notes: The figure shows shows the zip codes in the county along with the roads used in the analysis. The sample of zip codes with centroids within four miles of an I-5 or I-10 on-ramp are shaded or patterned. The pattern defines the commuting region as determined by the direction to downtown Los Angeles.
Sources: Los Angeles County GIS Portal.

Figure 2: Traffic by Route, Direction and Time of Day



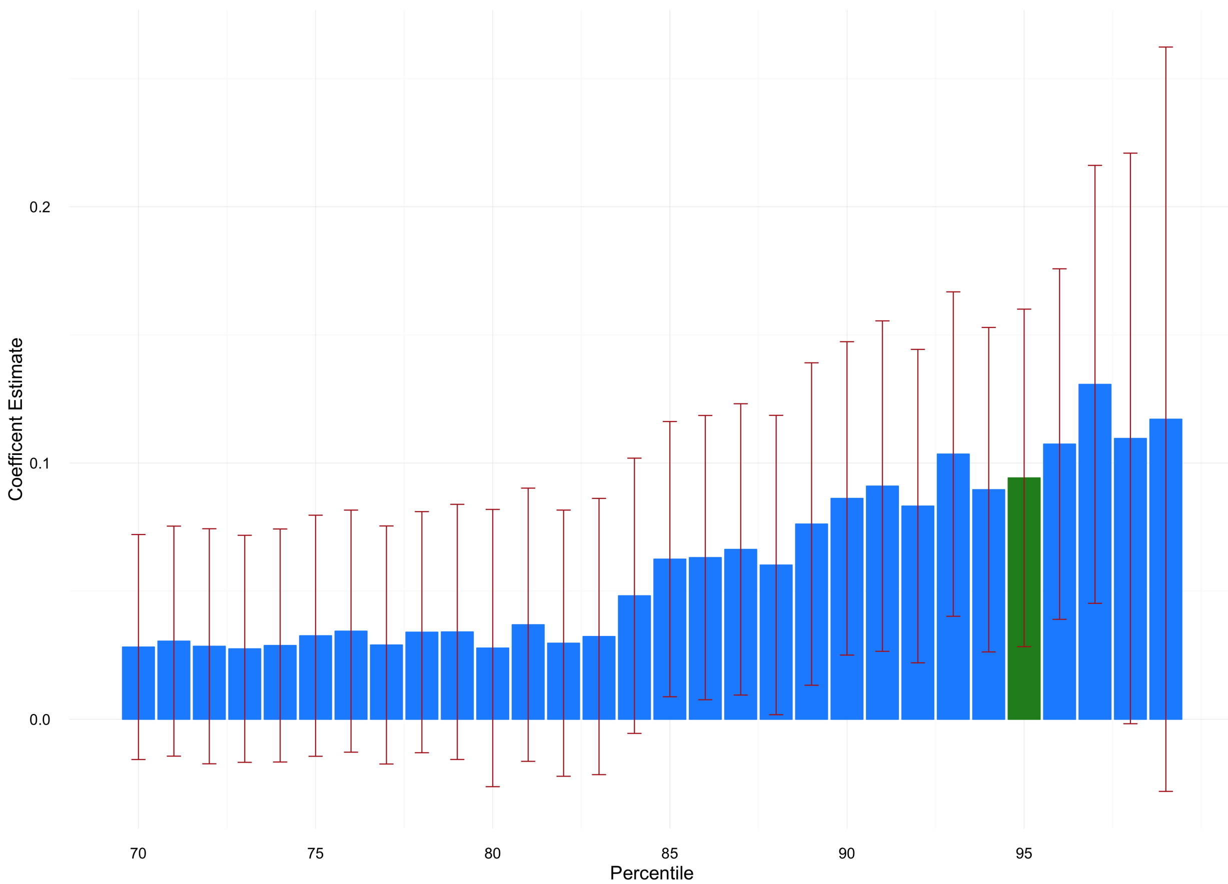
Notes: Travel times only measure the time on I-5 or I-10, not the door to door travel times.
Sources: Los Angeles County GIS Portal.

Figure 3: Map of Domestic Violence in Los Angeles



Sources: Los Angeles Police Department and Los Angeles Sheriff Department.

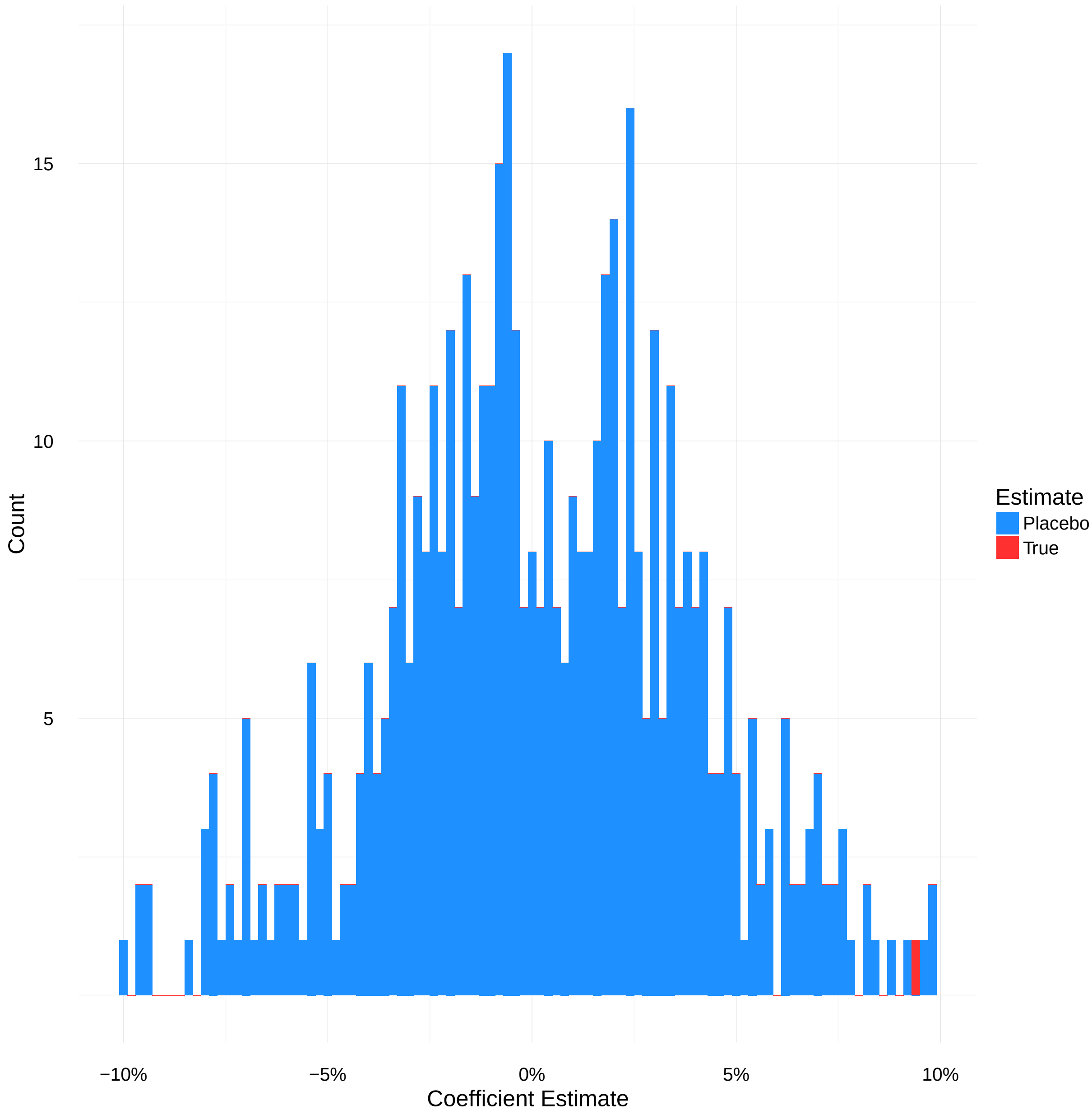
Figure 4: Effect of Traffic on Domestic Violence - Different Thresholds



Notes: The figure shows estimates of indicator variables for high traffic using percentiles starting at the 70th percentiles. Each bar represents the coefficient estimate of that percentile indicator, and the error bars are 95% confidence intervals generated from the robust standard errors clustered at the zip code level. Each coefficient is estimated using our preferred specification shown in column (4) of Table 1.

Sources: California Department of Transportation, Los Angeles Police Department, Los Angeles Sheriff Department, NOAA National Centers for Environmental Information and Longitudinal Employer-Household Dynamics.

Figure 5: Distribution of Parameter Estimates for Randomized Traffic



Notes: For each zip code we randomly select traffic from a zip code in a region of the city with a different commute pattern and then estimate the impact of "false traffic" on crime. We repeat this process 500 times and each time we set a new seed that generates the random number used to assign a zip code's false traffic. The graph shows the distribution of treatment effects for false traffic as well as our preferred estimate.
Sources: California Department of Transportation, Los Angeles Police Department, Los Angeles Sheriff Department, NOAA National Centers for Environmental Information and Longitudinal Employer-Household Dynamics.

Table 1: Base Results

	(1) All Observations	(2) No Downtown	(3) Workdays	(4) No Downtown & Workdays
95th Percentile	0.0625**	0.0715**	0.0868***	0.0942***
	(0.0254)	(0.0287)	(0.0295)	(0.0336)
Observations	193,409	168,108	132,241	114,778

Notes: The regression models include zip code, year-month, and day of week fixed effects, a quadratic time trend, rain, maximum temperature, wind speed, and average zip code level traffic in the last week and last month as controls. Robust standard errors clustered at the zip code are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Sources: California Department of Transportation, Los Angeles Police Department, Los Angeles Sheriff Department, NOAA National Centers for Environmental Information and Longitudinal Employer-Household Dynamics.

Table 2: Different Crimes

	(1) Domestic Violence	(2) All	(3) Assault	(4) Property	(5) Homicide
95th Percentile	0.0942***	0.0151	-0.0279	0.00334	0.113
	(0.0336)	(0.0109)	(0.0190)	(0.0151)	(0.227)
Observations	114,778	132,243	126,004	128,501	71,118

Notes: The regression models include zip code, year-month, and day of week fixed effects, a quadratic time trend, rain, maximum temperature, wind speed, and average zip code level traffic in the last week and last month as controls. Robust standard errors clustered at the zip code are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Sources: California Department of Transportation, Los Angeles Police Department, Los Angeles Sheriff Department, NOAA National Centers for Environmental Information and Longitudinal Employer-Household Dynamics.

Table 3: Morning and Evening Traffic and Crime

	(1) PM Traffic	(2) AM Traffic	(3) AM & PM	(4) PM Traffic AM Crime
95th Percentile (PM)	0.0942***		0.0933***	-0.000492
	(0.0336)		(0.0337)	(0.0582)
95th Percentile (AM)		-0.0451	-0.0431	
		(0.0457)	(0.0458)	
Observations	114,778	114,778	114,778	104,800

Notes: The dependent variable in columns (1)-(3) is evening domestic violence counts and in column (4) it is morning domestic violence counts. The 95th Percentile (PM) and 95th Percentile (PM) correspond to extreme traffic in the evening and morning, respectively. The regression models include zip code, year-month, and day of week fixed effects, a quadratic time trend, rain, maximum temperature, wind speed, and average zip code level traffic in the last week and last month as controls. Robust standard errors clustered at the zip code are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Sources: California Department of Transportation, Los Angeles Police Department, Los Angeles Sheriff Department, NOAA National Centers for Environmental Information and Longitudinal Employer-Household Dynamics.

Table 4: Differential Effects by Commuting Patterns

	(1) Interaction	(2) Above Median	(3) Below Median
95th Percentile × %Commuters	0.234***		
	(0.0766)		
95th Percentile		0.114***	0.0665
		(0.0435)	(0.0493)
Observations	114,778	57,390	57,388

Notes: The percentage of commuters is calculated at the zip code level using the LEHD Origin-Destination data and represents the fraction of commuters that are likely to use the route assigned to them. The median percentage of commuters across zip codes is 41%. The regression models include zip code, year-month, and day of week fixed effects, a quadratic time trend, rain, maximum temperature, wind speed, and average zip code level traffic in the last week and last month as controls. Robust standard errors clustered at the zip code are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Sources: California Department of Transportation, Los Angeles Police Department, Los Angeles Sheriff Department, NOAA National Centers for Environmental Information and Longitudinal Employer-Household Dynamics.

Table 5: Placebo Test - Lags of Domestic Violence on Traffic

	(1) T-1	(2) T-2	(3) T-3	(4) T-4
95th Percentile	0.0495	-0.00137	0.00921	0.0634
	(0.0672)	(0.0586)	(0.0398)	(0.0476)
Observations	90,300	90,374	90,292	91,141

Notes: The regression models include zip code, year-month, and day of week fixed effects, a quadratic time trend, rain, maximum temperature, wind speed, and average zip code level traffic in the last week and last month as controls. Robust standard errors clustered at the zip code are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Sources: California Department of Transportation, Los Angeles Police Department, Los Angeles Sheriff Department, NOAA National Centers for Environmental Information and Longitudinal Employer-Household Dynamics.

Table 6: Robustness

	(1) Exclude Short	(2) Exclude Long	(3) No Weather	(4) Lag Holidays	(5) Clear Choice	(6) Date FE	(7) False Traffic
95th Percentile	0.0607*	0.0659**	0.0928***	0.0954***	0.101***	0.0695*	-0.00835
	(0.0331)	(0.0325)	(0.0341)	(0.0338)	(0.0344)	(0.0281)	(0.0210)
Observations	106,044	104,801	114,778	113,727	96,066	114,778	114,778

Notes: Exclude Short and Exclude Long drop the bottom and top 10 % of zip codes by average travel time. No Weather removes weather control variables. Lag Holidays removes the day before Thanksgiving, Christmas and New Years. Clear Choice restricts the sample to zip codes where the difference in distance between I-5 and I-10 is at least 4 miles. Date FE replaces month-by-year fixed effects with date fixed effects. False Traffic replaces the 95th percentile dummy with a one for days when a zip code did not experience extreme traffic, but there was extreme traffic in one of the other three regions. Except if otherwise noted, regression models include zip code, year-month, and day of week fixed effects, a quadratic time trend, rain, maximum temperature, wind speed, and average zip code level traffic in the last week and last month as controls. Robust standard errors clustered at the zip code are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Sources: California Department of Transportation, Los Angeles Police Department, Los Angeles Sheriff Department, NOAA National Centers for Environmental Information and Longitudinal Employer-Household Dynamics.

Table 7: OLS Specifications

	(1) OLS Main	(2) OLS No Weather	(3) OLS Lag Holidays	(4) OLS Date FE	(5) OLS Two-Way
95th Percentile	0.0100***	0.0096**	0.0102***	0.0080*	0.0100**
	(0.0037)	(0.0037)	(0.037)	(0.0041)	(0.0038)
Observations	132,243	132,243	132,243	132,243	132,243

This table shows estimates from linear regression models as opposed to a Poisson model. No Weather removes weather control variables. Lag Holidays removes the day before Thanksgiving, Christmas and New Years. Date FE replaces month-by-year fixed effects with date fixed effects. Two-way clusters the standard error at both the zip code and month level. Except if otherwise noted, regression models include zip code, year-month, and day of week fixed effects, a quadratic time trend, rain, maximum temperature, wind speed, and average zip code level traffic in the last week and last month as controls. Robust standard errors clustered at the zip code or zip code and month level (column (5)) are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Sources: California Department of Transportation, Los Angeles Police Department, Los Angeles Sheriff Department, NOAA National Centers for Environmental Information and Longitudinal Employer-Household Dynamics.

Table 8: Instrumental Variables Approach

	(1) IV Number Accidents	(2) IV Duration Accidents
Second Stage	0.1516*	0.1556*
	(0.0829)	(0.0831)
Observations	92,488	92,413
F-Test	39.39	10.38
First Stage	0.0004***	0.0003***
	(0.0001)	(0.0001)

Notes: This reports instrumental variable regressions where extreme traffic is instrumented with the number (column (1)) and duration (column (2)) or severe accidents lasting more than one hour. The regression models include zip code, year-month, and day of week fixed effects, a quadratic time trend, rain, maximum temperature, wind speed, and average zip code level traffic in the last week and last month as controls. Robust standard errors clustered at the zip code are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Sources: California Department of Transportation, Los Angeles Police Department, Los Angeles Sheriff Department, NOAA National Centers for Environmental Information and Longitudinal Employer-Household Dynamics.

Table 9: Google Trends

	(1) Base	(2) Search Volume	(3) Lags	(4) 95th Percentile	(5) No Traffic	(6) All City
95th Percentile	0.0942***	0.0879***	0.104***	0.0907***		
	(0.0336)	(0.0337)	(0.0360)	(0.0339)		
Google		0.0427**	0.0262		0.0458***	0.0319***
		(0.0170)	(0.0190)		(0.0170)	(0.00975)
Google_{T-1}			0.0206			
			(0.0223)			
Google_{T-2}			-0.0127			
			(0.0155)			
Google_{T-3}			-0.000695			
			(0.0172)			
Google_{T-4}			0.0113			
			(0.0140)			
Google_{T-5}			-0.00231			
			(0.0142)			
Google 95th				0.0905**		
				(0.0411)		
Observations	114,778	114,778	108,416	114,778	114,778	290,621

Notes: These models utilize google search volumes for "traffic" in the Los Angeles metropolitan area. The regression models include zip code, year-month, and day of week fixed effects, a quadratic time trend, rain, maximum temperature, wind speed, and average zip code level traffic in the last week and last month as controls. Robust standard errors clustered at the zip code are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Sources: California Department of Transportation, Los Angeles Police Department, Los Angeles Sheriff Department, NOAA National Centers for Environmental Information and Longitudinal Employer-Household Dynamics, and google trends data.

Table 10: Expected vs. Unexpected Traffic

	(1) No Expectations	(2) High Variance	(3) Low Variance	(4) MA Expected	(5) MA Unexpected	(6) MA Both
95 Percentile	0.0736**	0.1100**	0.0779*			
	(0.0318)	(0.0512)	(0.0457)			
95 Percentile Expected				0.0464		0.0437
				(0.0455)		(0.0458)
95 Percentile Unexpected					0.0770**	0.0760**
					(0.0344)	(0.0348)
Observations	114,778	53,640	61,138	114,778	114,778	114,778

Notes: No Expectations removes controls for recent traffic. High (Low) Variation Traffic splits the sample based on whether the zip code is above (below) the median in terms of the standard deviation of travel time. Moving Average models predict traffic conditions using the previous five days along with all other control variables and fixed effects. The Expected and Unexpected variables are constructed using the predictions and residuals from the moving average model, respectively. The regression models include zip code, year-month, and day of week fixed effects, a quadratic time trend, rain, maximum temperature, wind speed, and average zip code level traffic in the last week and last month as controls. Robust standard errors clustered at the zip code are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Sources: California Department of Transportation, Los Angeles Police Department, Los Angeles Sheriff Department, NOAA National Centers for Environmental Information and Longitudinal Employer-Household Dynamics.

Appendix

Table A.1: Daily Average Traffic by Zip Code and Time of Day

	Mean	Std. Dev.	Min	Max	Observations
Travel Time (AM)	24.38	15.48	3.60	126.35	160272
Travel Time (PM)	26.12	14.18	3.77	149.01	160272

Notes: The table presents summary statistics for daily average travel times in morning and evening commutes for the zip codes in our sample. Travel times reflect time spent on either I-5 or I-10 and not door-to-door travel times.
Sources: California Department of Transportation

Table A.2: Daily Average Crime by Zip Code and Time of Day

All Crimes					
	Mean	Std. Dev.	Min	Max	Observations
All	3.57	4.43	0.00	70.00	160378
Assault	0.48	0.98	0.00	25.00	160378
Domestic Violence	0.14	0.42	0.00	8.00	160378
Property	1.57	2.20	0.00	34.00	160378
Homicide	0.00	0.06	0.00	2.00	160378
Evening Crimes					
	Mean	Std. Dev.	Min	Max	Observations
All	2.15	2.89	0.00	60.00	160378
Assault	0.32	0.74	0.00	18.00	160378
Domestic Violence	0.10	0.34	0.00	8.00	160378
Property	0.93	1.46	0.00	33.00	160378
Homicide	0.00	0.05	0.00	2.00	160378

Notes: The table presents summary statistics daily average crime by zip code for total crime and evening crime.
Sources: Los Angeles Police Department and Los Angeles Sheriff Department.

Table A.3: Leads of Crime

	(1) T	(2) T+1	(3) T+2	(4) T+3	(5) T+4	(6) T and T+1
95 Percentile	0.0942***	0.1108***	0.0069	0.0157	-0.0005	0.1031***
	(0.0336)	(0.0329)	(0.0447)	(0.0360)	(0.0416)	(0.0247)
Observations	114,778	115,932	114,595	112,017	113,168	115,932

Notes: The regression models include zip code, year-month, and day of week fixed effects, a quadratic time trend, rain, maximum temperature, wind speed, and average zip code level traffic in the last week and last month as controls. Robust standard errors clustered at the zip code are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Sources: California Department of Transportation, Los Angeles Police Department, Los Angeles Sheriff Department, NOAA National Centers for Environmental Information and Longitudinal Employer-Household Dynamics.

Table A.4: Heterogeneity by Access to Public Transportation

	(1) 1 mile	(2) 1.5 miles	(3) 2 miles	(4) 2.5 miles	(5) 3 miles
95th Percentile	0.130***	0.128***	0.128***	0.133***	0.113***
	(0.0490)	(0.0483)	(0.0436)	(0.0423)	(0.0416)
Observations	31,189	38,678	49,906	59,884	68,617

Notes: These models limit the sample to zip codes within a certain distance to a metro station. The regression models include zip code, year-month, and day of week fixed effects, a quadratic time trend, rain, maximum temperature, wind speed, and average zip code level traffic in the last week and last month as controls. Robust standard errors clustered at the zip code are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Sources: California Department of Transportation, Los Angeles Police Department, Los Angeles Sheriff Department, Los Angeles Open Data Website, NOAA National Centers for Environmental Information and Longitudinal Employer-Household Dynamics.

Table A.5: Interaction by Zip Code Characteristics

	(1) Base	(2) Income	(3) Crime	(4) Distance	(5) All
95 Percentile	0.0942***	0.1089*	0.0970***	0.1092***	0.1284**
	(0.0336)	(0.0558)	(0.0346)	(0.0387)	(0.0563)
95 Percentile * High Income Zip		-0.0250			-0.0279
		(0.0651)			(0.0629)
95 Percentile * Low Crime zip			-0.0441		-0.0371
			(0.1347)		(0.1391)
95 Percentile * Farther Zip				-0.0544	-0.0560
				(0.0654)	(0.0654)
Observations	114,778	114,778	114,778	114,778	114,778

Notes: The regression models include zip code, year-month, and day of week fixed effects, a quadratic time trend, rain, maximum temperature, wind speed, and average zip code level traffic in the last week and last month as controls. Robust standard errors clustered at the zip code are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Sources: California Department of Transportation, Los Angeles Police Department, Los Angeles Sheriff Department, U.S. Census American Community Survey, NOAA National Centers for Environmental Information and Longitudinal Employer-Household Dynamics.

Table A.6: Contemporaneous and Lagged Effects of Accidents on Traffic

	(1) Number	(2) Number	(3) Duration	(4) Duration
Number of Accidents	3.413***	3.316***		
	(0.121)	(0.121)		
Number of Accidents$_{T-1}$		2.250***		
		(0.121)		
Number of Accidents$_{T-2}$		1.105***		
		(0.121)		
Number of Accidents$_{T-3}$		0.175		
		(0.121)		
Duration of Accidents			0.0106***	0.0104***
			(0.000576)	(0.000576)
Duration of Accidents$_{T-1}$				0.00706***
				(0.000578)
Duration of Accidents$_{T-2}$				0.00297***
				(0.000578)
Duration of Accidents$_{T-3}$				0.0000783
				(0.000579)
Observations	120,960	120,948	120,960	120,948

Notes: The dependent variable is hourly travel times. The regression models include zip code, year-month, and day of week fixed effects, a quadratic time trend, rain, maximum temperature, and wind speed as controls. Robust standard errors clustered at the zip code are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Sources: California Department of Transportation and NOAA National Centers for Environmental Information.

Table A.7: Zip Codes Near On-ramps

	(1) 2 miles	(2) 2.5 mile	(3) 3 miles	(4) 3.5 miles	(5) 4 miles
95th Percentile	0.131^{**}	0.121^{***}	0.106^{***}	0.111^{***}	0.0942^{***}
	(0.0548)	(0.0435)	(0.0383)	(0.0356)	(0.0336)
Observations	46,160	72,360	87,327	99,805	114,778

Notes: These models limit the sample to zip codes within a certain distance to an I-5 or I-10 on-ramp. The regression models include zip code, year-month, and day of week fixed effects, a quadratic time trend, rain, maximum temperature, wind speed, and average zip code level traffic in the last week and last month as controls. Robust standard errors clustered at the zip code are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Sources: California Department of Transportation, Los Angeles Police Department, Los Angeles Sheriff Department, NOAA National Centers for Environmental Information and Longitudinal Employer-Household Dynamics.

Table A.8: Different Specifications for Traffic Assignment

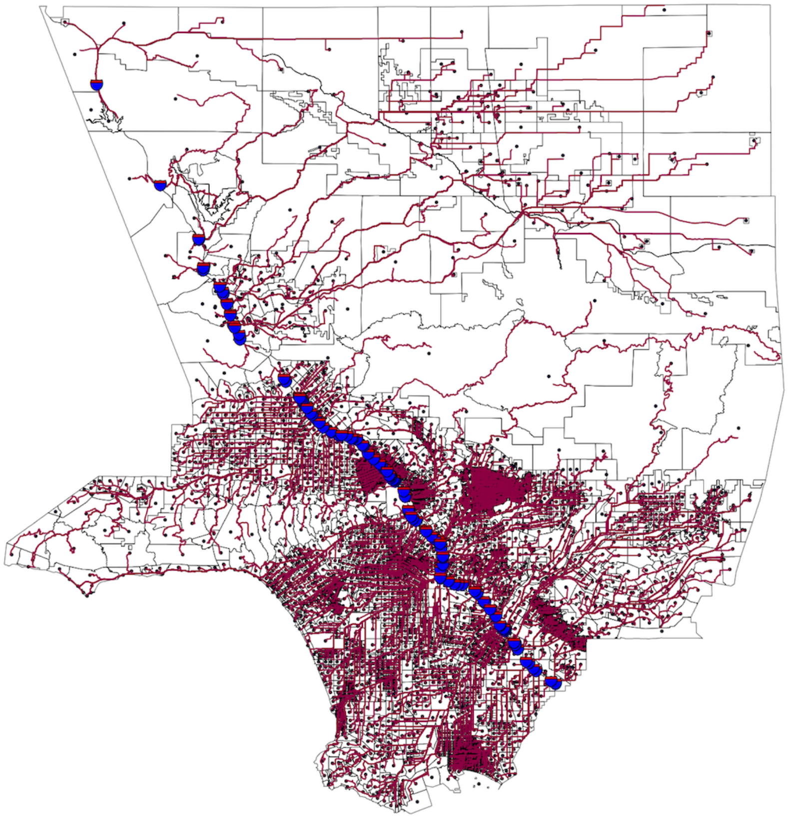
	(1) 99th	(2) 95th	(3) 90th	(4) 75th	(5) 50th	(6) 25th
95th Percentile	0.0859**	0.116***	0.132***	0.0942***	0.0683*	0.0779**
	(0.0409)	(0.0403)	(0.0403)	(0.0336)	(0.0353)	(0.0361)
Observations	114,778	114,778	114,778	114,778	114,778	114,778

Notes: These models test different specifications for the determining the destination zip code using the Longitudinal Employer-Household Dynamics Origin-Destination data. The regression models include zip code, year-month, and day of week fixed effects, a quadratic time trend, rain, maximum temperature, wind speed, and average zip code level traffic in the last week and last month as controls. Robust standard errors clustered at the zip code are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

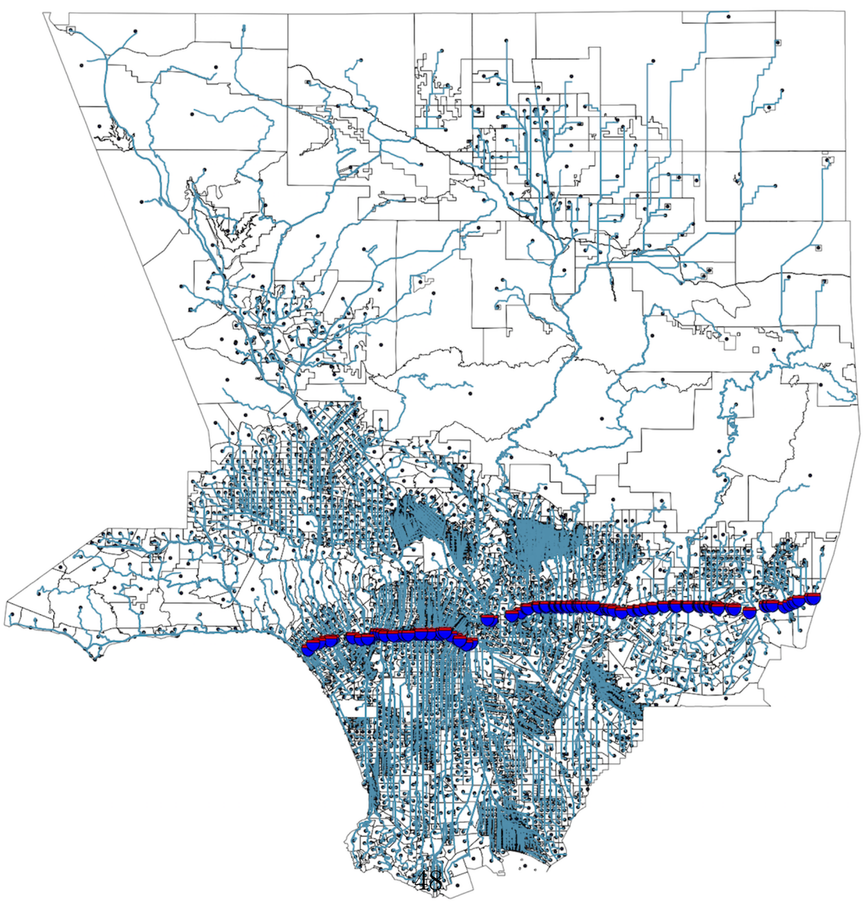
Sources: California Department of Transportation, Los Angeles Police Department, Los Angeles Sheriff Department, NOAA National Centers for Environmental Information and Longitudinal Employer-Household Dynamics.

Figure A.1: Mapping Zip Codes to Roads

(a) I-5



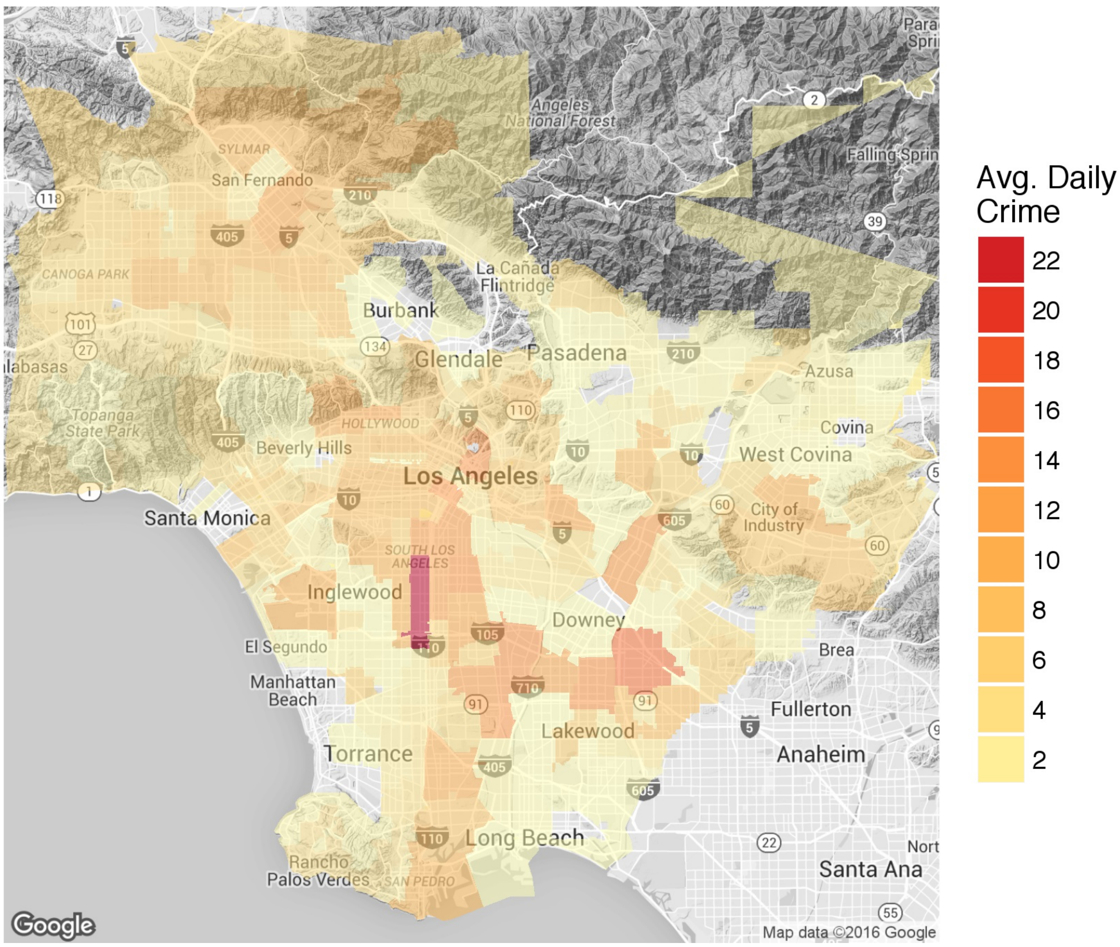
(b) I-10



Sources: California Department of Transportation

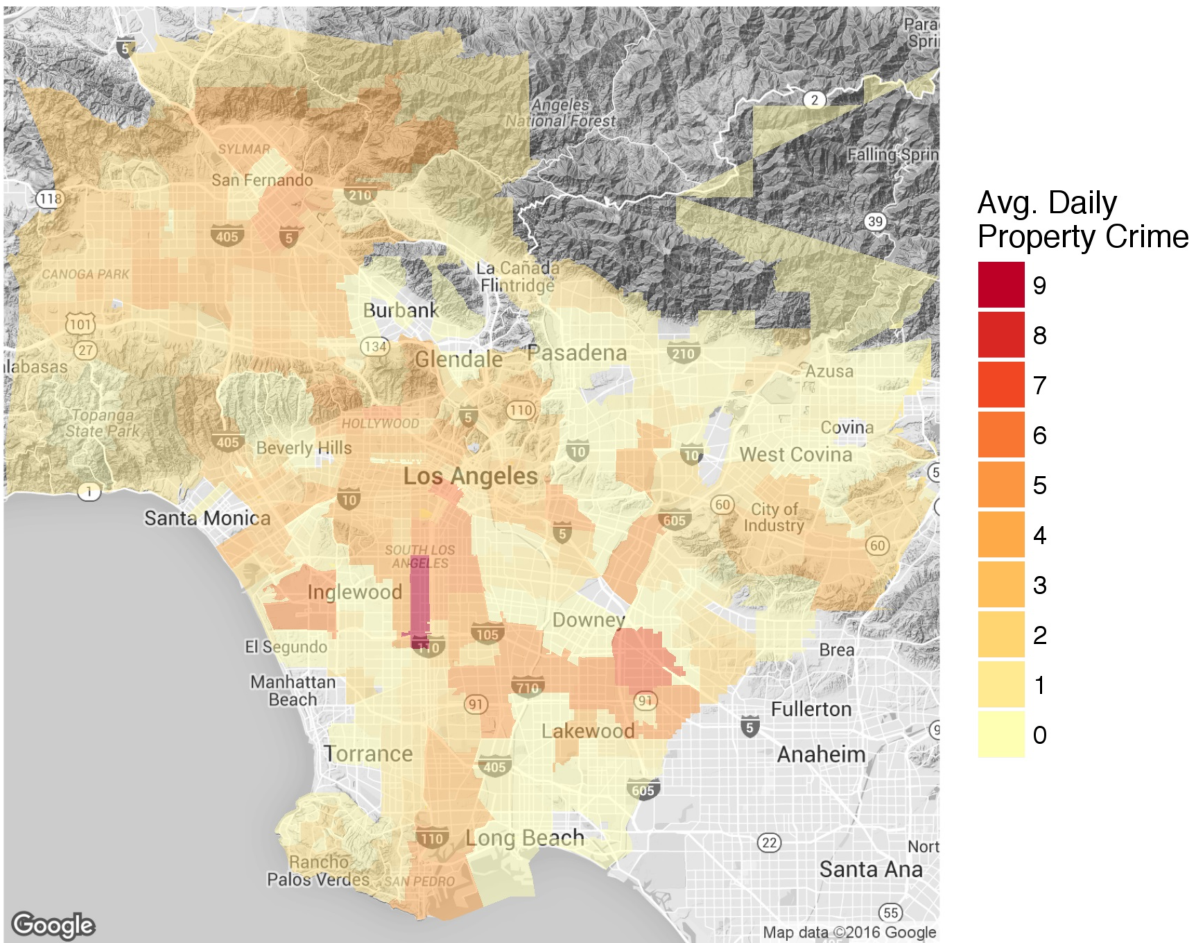
Sources: California Department of Transportation and Los Angeles County GIS Portal.

Figure A.2: Map of Crimes in Los Angeles



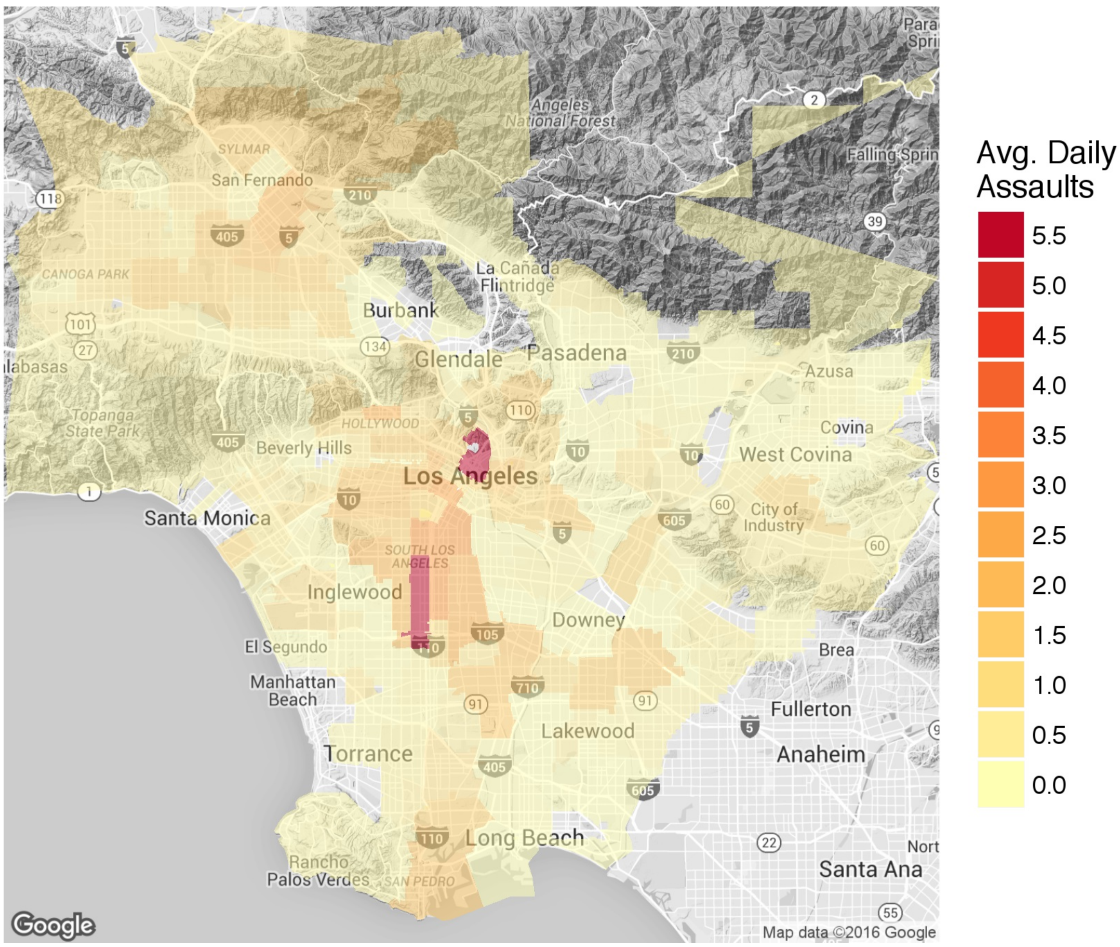
Sources: Los Angeles Police Department and Los Angeles Sheriff Department.

Figure A.3: Map of Property crime in Los Angeles



Sources: Los Angeles Police Department and Los Angeles Sheriff Department.

Figure A.4: Map of Assault in Los Angeles



Sources: Los Angeles Police Department and Los Angeles Sheriff Department.