Appendix

Criminal Deterrence when there are Offsetting Risks: Traffic Cameras, Vehicular Accidents, and Public Safety

Justin Gallagher and Paul J. Fisher[*]

February 11, 2019

1

Contents

1   Empirical Bayes Model

The "empirical bayes" (EB) model is the standard model used to evaluate how red light cameras (hereafter "cameras") affect traffic accidents (e.g. Abbess et al. [1981]; Hauer [2006]; Hauer [1992]; Hauer et al. [2002]). Section 1.1 summarizes how the traffic literature uses the empirical bayes model to estimate the causal effect of cameras on accidents. In section 1.4 we provide a summary of the recent literature.

It is widely claimed that the empirical bayes model corrects for mean reversion, while also improving statistical precision.[1] Mean reversion occurs as many cities (including Houston) install cameras at intersections with high accident levels in the previous year (Stein et al. [2006]). These intersections are likely to experience a decline in accidents regardless of the effectiveness of the camera program. Sections 1.2.1 and 1.2.2 discuss why the EB model does not fully correct for mean reversion.

A well known challenge in studying the cause of traffic accidents is the high variability in the number and location of accidents from year to year (e.g. Hauer [1992]). This variability increases the accident variance at city intersections and can make estimating statistically precise camera program causal treatment effects difficult. A selling point of the EB model are the extremely precise model estimates. This is true even when an EB study includes just a handful of camera intersections (e.g. Carriquiry and Pawlovich [2004]; Persaud and Lyon [2014]; Lord and Greedipally [2014]). Section 1.2.3 outlines why the statistical precision of the EB model is overstated.

Section 1.3 provides empirical evidence that the EB Model fails to correct for mean reversion. First, we simulate fake accident data and estimate the causal effect of a placebo camera program using the EB model. Estimates from

---

[1]The EB model has become largely associated with Ezra Hauer. Hauer et al. [2002] states: "The empirical bayes (EB) method for the estimation of safety increases the precision of estimation and corrects for the regression-to-mean bias" (p126). Recent applications to red light camera programs also support this view (e.g. Erke [2009], Høye [2013], Ko et al. [2013], and Ko et al. [2017]). For example, Høye [2013] writes: "Since no obvious weaknesses in the application of the EB method were found, all of these studies can be assumed to be largely unaffected by [mean reversion]" (p81).

an EB model consistently suggest a positive, statistically significant causal effect of the camera program on the number of accidents when, in fact, no such relationship exists (by construction). Second, we estimate the causal effect of the introduction of cameras in Houston using an EB model. The EB model estimates for the introduction of cameras in Houston imply a statistically significant reduction in total accidents of 30% (untrimmed panel) and 5% (trimmed panel). These results are in contrast to our referendum-based camera removal estimates that avoid the endogeneity concerns related to the camera placement and start times (e.g. mean reversion).

In summary, the EB Model's failure to correct for mean reversion and the model's overstated statistical precision together can account for the large number of studies that conclude camera programs reduce intersection accidents. In contrast to the current literature, our empirical strategy follows a "design-based" econometric approach (Angrist and Pischke [2017]) that does not rely on correctly modeling the entire accident generating process.

1.1 Estimating the Empirical Bayes Model

The first step is to specify a structural model of accidents. The structural model of accidents is referred to as the Safety Performance Function (SPF). The SPF is supposed to accurately characterize the deterministic relationship between the number of accidents at an intersection and the variables that cause these accidents. In practice, the SPF often includes a small number of intersection characteristics such as average daily traffic, number of lanes, speed limit, lane width, and right-turn-on-red prohibition (e.g. Lord and Greedipally [2014]; Mahmassani et al. [2017]). Several recent papers use only average daily traffic to model the SPF (e.g. Ko et al. [2013]; Pulugurtha and Otturu [2014]; Ko et al. [2017]).

The SPF is estimated using a sample of non-camera intersections to predict accident levels at camera intersections. The traffic literature usually estimates the SPF using a negative binomial regression. $E(k_i)$ is the expected number of accidents for camera intersection i prior to the start of the camera program

3

(pre-period).[2] $E(\lambda_i)$ is the expected number of accidents for a camera intersection during the camera program (post-period). $E(k_i)$ and $E(\lambda_i)$ are estimated by running the SPF model on a sample of non-camera intersections during the pre- and post-periods, respectively.

The next step is to calculate a weighted average of the actual and predicted number of pre-period camera intersection accidents at each camera intersection.

$$E(k_i|K_i) = w_i E(k_i) + (1 - w_i)K_i \tag{1}$$

K_i is the actual number of accidents at the camera intersection during the pre-period. The weight w_i, when estimating the model using a negative binomial regression, is

$$w_i = \frac{1}{1 + \frac{E(k_i)}{\phi}} \tag{2}$$

, where ϕ is the estimated inverse negative binomial dispersion parameter.

The goal is for the weighted average to estimate what the level of accidents would be absent any stochastic accident shocks. The concern for a researcher interested in estimating the causal effect of the camera program is that the camera intersections are likely to have higher than predicted pre-program accident levels due to positive stochastic shocks. If the model does not account for these shocks, then it is likely to over-estimate any beneficial impact of the camera program. The camera intersections are likely to experience mean reversion and lower accident levels regardless of the camera program.

The estimated number of accidents, $E(k_i|K_i)$, is then adjusted to account for the passage of time. Equation 3 adjusts the weighted average from Equation 1 using $q_i = \frac{E(\lambda_i)}{E(k_i)}$. q_i is calculated as the SPF prediction for the treated sites in the post-period, divided by the SPF prediction for the treated sites in the pre-period.

$$\pi_i = q_i E(k_i|K_i) \tag{3}$$

---

[2]The majority of studies focus on what happens when the cameras are installed. Our discussion of the EB model follows this literature and considers the time period prior to the start of the camera program as the pre-period.

The assumption is that q_i captures what would have occurred in the absence of treatment. This assumption is similar to the common trends assumption in the difference-in-differences model (e.g. Cameron and Trivedi [2005]), except with two key distinctions. First, the difference-in-differences model uses trends in the underlying data for the control group, whereas the EB model uses the ratio of the before and after predictions from a traffic model (the SPF). Second, an appealing feature of the difference-in-differences model is that the common trends assumption is transparent and can be easily evaluated.[3] Less clear is how to evaluate and test the assumption underlying q_i. We are not aware of a single camera study using the EB model that tests this assumption.

The estimated causal effect of the cameras on the number of accidents is calculated as $\delta = \pi - \lambda$. π is the sum of expected accidents at camera intersections in the post-period (from Equation 3). λ is the sum of the actual number of accidents at camera intersections in the post-period. δ provides a valid causal estimate of the camera program on the number of accidents at camera intersections if π_i is an unbiased estimate for the predicted accidents at each camera intersection (absent the camera program).

$$\theta = \frac{\frac{\lambda}{\pi}}{1 + \frac{Var(\pi)}{\pi^2}} \tag{4}$$

The literature often presents the EB model results as θ, which is sometimes referred to as the safety index. θ is a weighted ratio of actual accidents to estimated accidents, and is interpreted as the estimated percent decrease in accidents when $\pi < \lambda$.[4]

1.2 Internal Validity Concerns

We have three main concerns with the empirical bayes model. Section 1.2.1 outlines our concern with the EB model's reliance on estimating the "structural

---

[3]Common trends in the pre-period can be tested. There is no way to know for sure how the trend for the camera group would have evolved in the post-period without cameras.

[4]$Var[\theta] = \frac{\theta^2 [\frac{var(\lambda)}{\lambda^2} + \frac{var(\pi)}{\pi^2}]}{1 + \frac{Var(\pi)}{\pi^2}}$

relation" between accidents and the road characteristics that cause accidents (Angrist and Pischke [2017]). Section 1.2.2 discusses why the EB model is very unlikely to correct for mean reversion. Section 1.2.3 reviews why the statistical precision of the EB model is overstated. The failure to correct for mean reversion and the EB model's overstated statistical precision together can explain the large number of studies that conclude that camera programs reduce the number of intersection accidents.

1.2.1   Modeling the Safety Performance Function

The EB model relies on the correct specification of the Safety Performance Function. The starting point of the EB model is specifying the structural relation between traffic accidents and the intersection characteristics that determine the number of accidents. Estimates from the SPF are used to separately adjust the observed accident data for calendar time trends, and as a way to eliminate bias due to regression to the mean. Successful use of the SPF estimates for either purpose requires a complete and accurate model of what causes traffic accidents.

As we highlight in the previous section, most EB camera studies specify and estimate a very parsimonious safety performance function. Further, very few studies report SPF model goodness of fit statistics to indicate the amount of variation in traffic accidents that can be explained by the intersection characteristics in the SPF.[5] One exception is Powers and Carson [2004], who report a R-squared of 0.24 from a linear regression for an SPF with two accident characteristics (length of roadway segment and average daily traffic). Thus, the SPF in Powers and Carson [2004] can only capture about 25% of the variation in the number of accidents at an intersection. In other words, the model whose purpose is to adjust for time trends and to eliminate bias due to regression to

---

[5]Those studies that do report model statistics (e.g. Ko et al. [2013]; Pulugurtha and Otturu [2014]) focus on statistics such as Akaike information criterion (AIC) and quasi likelihood under independence model criterion (QIC). These statistics are used for assessing the strength of explanatory variables and comparing the relative fit of a model among a finite set of models, given the sample data. These statistics are uninformative about the overall amount of variation captured by the SPF.

the mean, can not explain approximately three-quarters of the variation.

In our view, the EB model epitomizes what Angrist and Pischke [2017] describe as the outdated "model-driven approach to regression" aimed at producing a "statistically precise account of the processes generating economic outcomes" (p1). We are skeptical that a model-driven approach can reliably estimate the causal effect of altering a single intersection characteristic (e.g. the installation/removal of cameras). For example, Table 2 panel A columns (1) and (2) in the manuscript shows that observable characteristics of Houston camera intersections differ greatly from the typical Houston non-camera intersection. It is very likely that important unobservable (or poorly measured) characteristics also differ. The EB model applies SPF coefficient estimates that are estimated from a sample of non-camera intersections to predict accident levels at camera intersections. The estimates for camera intersections are likely to be biased due to model misspecification. By contrast, our empirical strategy follows a "design-based" econometric approach that does not rely on correctly modeling the entire accident generating process.

1.2.2 Mean Reversion

Mean reversion poses a serious empirical challenge when estimating the causal effect of cameras on the number of intersection accidents. For example, manuscript Figure 4 panel B shows that accident levels in San Antonio dramatically decline in the years before and after a placebo camera program. These placebo camera intersections are selected based on having very high accident totals in 2003 using a similar method as the actual method to select Houston intersections (Stein et al. [2006]). Both the placebo San Antonio camera intersections and the Houston camera intersections display remarkably similar yearly accident trends. In particular, accident levels immediately decrease in 2004 before the Houston camera program begins.

Advocates for the EB model assert that the model corrects for mean reversion (e.g. Hauer et al. [2002]). Equation 1 is the key modeling step that adjusts for mean reversion. Equation 1 calculates a weighted average of the actual and predicted number of pre-period accidents at a camera intersection.

The assumption is that the structural model (SPF) used to estimate the predicted number of accidents will provide an estimate that is non-mean reverting and unbiased. Even if this assumption is true (Section 1.2.1 outlines why this is unlikely to be the case), Equation 1 will only eliminate the mean reversion if all of the weight is on the predicted number of accidents (i.e. $w_i = 1$). Since the pre-period accident levels are mean reverting, then the weighted average, $E(k_i|K_i)$, will also be mean reverting if Equation 1 considers the actual number of accidents (i.e. when $w_i \in [0, 1)$). However, the role of mean reversion will likely be small (and statistically insignificant) provided the weight is close to one.

Equation 2 shows how the weight, w_i, is calculated. This equation is used to calculate the weight because it minimizes the variance of $E(k_i|K_i)$ (Hauer [1997]). The size of the weight is determined by the relative magnitudes of the estimated pre-period number of accidents for intersection i and the estimated dispersion parameter. The larger the ratio between the expected accidents and the dispersion parameter the smaller is the weight. A small w_i implies more weight on the (mean-reverting) pre-period number of accidents and less emphasis on the (non-mean reverting) SPF estimate. There is an explicit trade-off between minimizing the variance and the introduction of bias due to mean reversion. In essence, the weight represents the percentage of mean reversion removed relative to a naive pre-/post- difference model.

In practice, the EB model places a small weight (w_i) on the non-mean reverting SPF estimate and a large weight ($1-w_i$) on the mean-reverting actual number of accidents when evaluating camera intersections. The reason is that high accident intersections will have a larger accident to dispersion parameter ratio. For example, Ko et al. [2017] evaluate the Houston program and use a w_i of around 0.03 to 0.06 for their main EB model (installation) estimations.[6] w_i is also below 0.1 in our application of the EB model to the Houston panel dataset (see Section 1.3). Thus, roughly speaking, the EB model of Ko et al. [2017] (and our own application) correct for less than one tenth of the mean

---

[6]Ko et al. [2017] estimate that the introduction of the cameras in Houston led to a statistically significant 37% reduction in total accidents.

reversion.

1.2.3 Statistical Precision

A notable feature of the EB model is the high level of precision of its estimates. Hauer [1997] describes the EB model as incorporating two pieces of information from the pre-period: accident levels at the intersections of interest, and accident levels at similar intersections. The two pieces of information are combined using a weighted average (Equation 1). The key question is what weight to choose. Hauer [1997] advocates and derives the general form of w_i (Equation 2) under the criteria of maximizing the precision of $E(k_i|K_i)$. This approach prioritizes reducing variance at the expense of producing estimates robust to mean reversion.

We find the choice of the weight concerning for several reasons. First, the EB model is wrongly reported as being robust to mean reversion. Second, the San Antonio data (manuscript Section 3.2.2) and our EB model simulation (Section 1.3.1) show that the effect of mean reversion when evaluating a camera program can be large. Third, the model's precision can lead to an increase in the false positive rate. Only a small bias from mean reversion or via a poorly specified SPF is necessary to reject the null hypothesis that the effect of the camera program is zero. Taken together, the use of the EB model to study camera programs is likely to lead to the publication of studies that report incorrect, statistically significant results that suffer from a large amount of bias.

Finally, the EB model overstates the precision of the estimated accident treatment effect regardless of the choice of the weight. The reported variance of the estimated treatment effect does not account for the fact that key model inputs ($E(k_i)$ and $E(\lambda_i)$) are themselves estimates.

1.3 EB Model Simulation and Houston Estimates

In this section we show that the empirical bayes model fails to correct for mean reversion through model simulation, and over-estimates the safety effect

of the camera program in Houston. In section 1.3.1 we simulate fake accident data and estimate the causal effect of a placebo camera program using the EB model. Estimates from an EB model consistently suggest a positive, statistically significant causal effect of the camera program on the number of accidents when, in fact, no such relationship exists (by construction). In section 1.3.2 we use the EB model to estimate the causal effect of the introduction of the camera program in Houston on the number of accidents. The EB model implies a statistically significant reduction in total accidents of 30% (untrimmed panel) and 5% (trimmed panel).

1.3.1 Mean Reversion Simulation

We estimate camera program treatment effects from a placebo camera program using an empirical bayes model on simulated intersection accident data. We generate accident data for one thousand intersections over six years under different assumptions regarding how well the "safety performance function" (SPF) captures the underlying variation in accidents. The accident data are generated in two steps.

First, we randomly assign each intersection an average daily traffic (ADT) value using a gamma distribution. We select parameters for the gamma distribution so that the ADT mean and standard deviation (across intersections) approximately matches the observed values in our Houston sample.[7] Next, we consider the randomly assigned ADT for each intersection as the "mean ADT" for the intersection and generate random yearly ADT variation around the mean at each intersection.[8]

Second, we generate the level of accidents at each intersection using the randomly assigned intersection ADT values. We assume that the level of acci-

---

[7]The traffic engineering literature generally assumes that ADT follows a gamma distribution (Hauer [1997]). The mean ADT for the 1,003 Houston intersections is 31,488, with a standard deviation of 18,118. The ADT mean and standard deviation from our simulation are 31,488 and 18,118 (using a shape parameter of 3.017 and a scale parameter of 10,425).

[8]We generate six yearly ADT values for each intersection by taking random draws from a truncated normal distribution. The distribution is centered at each intersection's mean ADT, has a variance of 1,000, and is truncated from below at 200 (to rule out negative and implausibly low ADT values).

dents at each intersection is determined by a single intersection characteristic, ADT, along with random yearly accident variation. The rationale is to specify a simple accident generating process that can be captured by the EB model's structural model of accidents (i.e. the SPF). In practice, many EB model camera studies specify the SPF using only ADT (e.g. Powers and Carson [2004]; Ko et al. [2013]; Ko et al. [2017]). In our simulations, we vary how much of the observed accident levels are due to ADT (the deterministic factor) and how much are due to random variation.[9]

We vary how well ADT explains the level of accidents in our simulations so that we can observe the overall performance of the EB model based on the quality of the underlying accident model. That is, we simulate results when the SPF does a good job at capturing the true accident variation (ADT dependence high, unexplained variation low) and when the SPF does a poor job (ADT dependence low, unexplained variation high). When the ADT dependence is high, the SPF could eliminate most mean reversion, provided the weight in Equation 1 is near to one. Recall that, in practice, the weight (Equation 2) is selected to minimize variance of the EB model estimator.

There is no actual camera program (i.e. treatment). We assign a placebo camera to the fifty intersections in the sample that have the highest level of accidents in the pre-camera period for each ADT dependence simulation. The fifty intersections with the next highest level of accidents are selected as the control intersections. We define the first two years of the generated panel data as the pre-camera period, the next two years as having an active camera program, and the final two years as post-program.[10]

---

[9]We specify the accident generation process as: $Accidents_{it} = (\frac{ADT_{it}*y}{k}) + e_{it}$. The level of accidents at intersection i in year t depends on the ADT and e_{it}, a randomly distributed negative binomial accident component (conditional on y) with variance approximately equal to our main sample's variance. The constant $y \in (0,1)$ varies in our simulations and represents the proportion of the accident level determined by ADT. k is a normalizing factor that is set so that the mean number of accidents (across intersections) is 6.5 and approximately equal to what we observe in our Houston panel.

[10]We chose the initial pool of 1,000 intersections and the final estimating panel size of 100 so as to approximate the sample sizes of our Houston analysis. The total accident levels for the control intersections are similar to, but lower than, those of the camera intersections in our simulation and in the actual control group (see manuscript Table 2).

11

Table 8 shows the average results of the Monte Carlo simulations. We repeat the data generation process and model estimation ten thousand times for each ADT dependence parameter. Column (1) lists the ADT dependence parameters which range from 0.1 to 0.9. For example, the first row shows simulation results when 10% of the yearly accident level at each intersection is determined by the yearly ADT at the intersection. Columns (2) and (3) in the table show the estimated percent change and standard error (in parentheses) for an empirical bayes analysis based on the camera installation and camera removal, respectively. Columns (4) and (5) show the same estimates from a simple difference-in-differences (DD) model.[11]

All of the empirical bayes installation estimates imply that turning on the placebo cameras leads to a reduction in intersection accidents at camera intersections. The estimates range from -41% when the SPF captures little of the accident variation (first row) to -6% when the SPF captures most of the accident variation (last row). All of the (mean simulation) estimates are statistically significant at conventional significance levels. Since the true effect of the placebo camera program is zero, then the measured effect is due completely to mean reversion. Evaluating the empirical bayes treatment effect at the time of removal, several years after the placebo camera selection, avoids bias from mean reversion. The estimates in column (3) are all less than 3% in magnitude and not statistically significant at conventional levels.

The difference-in-differences installation estimates in column (4) follow a similar pattern as the empirical bayes installation estimates. The estimates imply reductions in total accidents of between 39% and 58% and are statistically significant at the 1% level. As in the EB model, the largest estimates are from simulations where most of the variation in the level of accidents is random and not dependent on ADT. The mean DD removal estimates are very close to zero and not statistically significant at conventional levels.

---

[11]The installation estimates consider the first two panel years as the pre-period and the middle two years as the post-period. The removal estimates consider the middle two years as the pre-period and the final two years as the post-period.

1.3.2 Houston Sample Predictions

In this section, we use the empirical bayes model to estimate the causal effect on the number of vehicle accidents from starting (installation) and ending (removal) the Houston camera program. We estimate the EB model on the same propensity score trimmed panel as in the manuscript. We also estimate the EB model on the untrimmed panel (see manuscript Table 2). The reason for showing results for the untrimmed sample is to more closely replicate the approach of the EB model literature, which does not typically attempt to balance intersection covariates before estimating the model. The assumption is that the safety performance function (SPF) adjusts for any differences between intersections with and without cameras. We follow several recent studies and model the SPF using only information on average daily traffic (e.g. Ko et al. [2013]; Pulugurtha and Otturu [2014]; Ko et al. [2017]). Otherwise, our EB model estimation approach is the same as Lord and Greedipally [2014].[12]

Table 9 shows the results from estimating the empirical bayes model on our Houston panels. Recall that EB model estimates are usually presented as a weighted ratio of actual accidents to estimated accidents. This ratio is sometimes referred to as the safety index (θ from Equation 4). A ratio of less than one is interpreted as a reduction in accidents.

The EB model camera removal estimates for the trimmed sample (panel A, columns (4)-(6)) are very similar to our difference-in-difference estimates (manuscript Table 3, panel A). The interpretation of the all accident coefficient is that the removal of the cameras led to 6% fewer accidents. The point estimate from our preferred model in the text implies 4% fewer accidents. However, one notable difference is the dramatically higher precision of the EB model. The probability value for the EB model estimate is 0.06, while it is .70 in the difference-in-differences model. The increased precision comes from using the SPF to leverage the out-of-sample predictions, and then using this information to prioritize minimization of the estimator variance at the expense of correcting (or guarding against) mean reversion (see Section 1.2.3).

---

[12]We contacted the authors who declined to share their code or data files. We wrote our own code following the model description in Lord and Greedipally [2014].

The camera installation estimates for the trimmed sample indicate that the cameras led to fewer total accidents. This is inconsistent with the estimated removal treatment effects that take advantage of the natural experiment. The negative installation estimates are what we would expect if the model did not fully correct for mean reversion.

The EB model relies on the SPF to capture the true intersection accident risk and to adjust for risk differences between intersections. A number of studies have specified parsimonious specifications for the SPF, including using only ADT (as we do in our analysis). These estimation differences suggest that a parsimonious SPF is unlikely to capture the true intersection accident risk, and will lead to bias in the model estimates unless the out-of-sample non-treatment intersections are carefully selected.

1.4   Camera Literature Review

Table 10 lists information for 16 recent camera studies. 13 of these studies use data at the intersection-level, while three studies are done at the city-level. Eight of the 13 intersection-level studies use empirical bayes as the primary means of analysis. All of the studies estimate the causal effect of the cameras from the camera installation. Three of the studies also estimate the causal effect when cameras are removed (Hu and Cicchino (2017); Ko et al. [2017]; Pulugurtha and Otturu [2014]).

Hu and Cicchino (2017) use large city fatal accident rates to estimate the effect of having at least one red light camera. This is very underpowered as red light camera programs cover a small part of the city.[13]

Ko et al. [2017] use the empirical Bayes before-after analysis to evaluate the safety impacts of installing and deactivating cameras on different types of red-light-running accidents at 48 Houston intersections. They estimate a large reduction in angle accidents from installing the cameras, and a smaller reduction in angle accidents from removing the cameras. There are several aspects of Ko et al. [2017] that we view as questionable. First, the weights

---

[13]Fewer then 2% of Houston accidents occur at red light camera intersections.

used in their empirical bayes analysis are on average about 4% to 11%.[14] This implies that the EB model eliminates only a small fraction of the mean reversion bias that is likely to affect the installation estimates (see Section 1.2.2 for a lengthier discussion). Second, the Safety Performance Function (SPF) only considers average daily traffic (ADT). As such, the SPF is almost certainly misspecified (see Section 1.2.1 for a discussion of general concerns on modeling the SPF). Third, the Ko et al. [2017] removal estimates would need to be about 40% larger (rather than about 50% smaller) to be internally consistent with the installation estimations, after adjusting for different baseline accident levels.[15] Fourth, Ko et al. [2017] only consider rear end accidents caused by "excessive braking." The exclusion of a large number of accidents leads the paper to underestimate rear end accident levels by 90% relative to our data. Further, limiting the analysis to a subset of the total accidents makes it very difficult to estimate a reliable overall (net) causal estimate for how cameras affect accidents.

Pulugurtha and Otturu [2014] evaluate a camera program at 32 signalized intersections in Charlotte, NC, and find a near zero effect on accidents from camera installation and a large decline in accidents following removal. However, similar to Ko et al. [2017], they specify a SPF where the level of accidents only depends on ADT (and no other roadway characteristic). The use of this SPF is very likely to generate biased predictions for the level of accidents even in the absence of cameras (treatment). The paper also produces internally inconsistent results (again similar to Ko et al. [2017]). This inconsistency is best explained through mean reversion leading the cameras to appear more effective at reducing accident when they are installed.

Two other analyses conducted by private consulting companies and not

---

[14]This is an approximation since we use their mean accident levels and the weights have a non-linear relationship to accident rates.

[15]On average, the number of accidents is lower at the time of camera removal in Houston across all city intersections. This is also true in Texas cities without a camera program (see manuscript Figure 4). The base year accidents levels for calculating percentage effects are inflated with installation estimates as they have higher accident levels. This makes a 20% reduction in accidents from installation estimation different in welfare terms then a 20% reduction in accident levels from a removal estimation.

published in peer review journals deserve highlighting (Lord and Greedipally [2014];Mahmassani et al. [2017]).[16] Mahmassani et al. [2017] estimate the impact of cameras in Chicago on accidents using an empirical bayes model. The study specifies a SPF with 21 variables (rather than just one or two as in Ko et al. [2017] and Pulugurtha and Otturu [2014]). Nevertheless, we are doubtful that any model that relies on the exact specification of the accident generating process to correct for mean reversion will lead to unbiased estimates. First, this approach introduces a second source of bias (omitted variable bias from the misspecification of the SPF) as a strategy to correct for the first type of bias. Second, even if there is no misspecification of the SPF, this strategy will only correct for all of the mean reversion if the weight on the SPF-based estimate is 100%. Mahmassani et al. [2017] do not publish the size of the weight, but by our calculations, the weight is only about 17%.[17]

Further, Mahmassani et al. [2017] consider a sample of Chicago accidents that only includes accidents that result in an injury. This sampling decision has two drawbacks. First, the sample of injury accidents is not representative of all accidents. Certain types of accidents are associated with higher injury rates. For example, angle accidents are associated with a higher injury rate than non-angle accidents (see manuscript Table 1). Thus, excluding non-injury accidents is likely to exaggerate the importance of (any) reduction in angle accidents and minimize the importance of (any) increase in non-angle accidents, and thereby lead to a biased estimate for total accidents. Overall, the bias will lead to results that exaggerate the effectiveness of a camera program at reducing total accidents. Second, excluding non-injury accidents, which are far more common than injury accidents, makes overall welfare analysis much more challenging.

The Lord and Greedipally [2014] analysis is similar to Mahmassani et al. [2017]. Lord and Greedipally [2014] also use an empirical bayes model to analyze Chicago's camera program, specify a parsimonious SPF, and the SPF has a low weight in the model (about 35%). Lord and Greedipally [2014] also ignore non-injury accidents. In summary, both Lord and Greedipally [2014]

---

[16]We thank a referee for alerting us to these studies.

[17]The authors of both reports declined to share either their data or estimation code.

and Mahmassani et al. [2017] fail to account for mean reversion and fail to estimate the full welfare impacts from the cameras due to sample selection.

2 Data Appendix

2.1 Data Sources

2.1.1 Vehicle Accidents

The 2003-2014 accident data from the Texas Department of Transportation's (TxDOT) Crash Records Information System (CRIS) includes all reported motor vehicle traffic accidents in Texas. We use information on all accidents in the cities of Houston and Dallas during this time period. The 2010-2014 CRIS data were downloaded via the TxDOT online database by month and year http://www.txdot.gov/government/enforcement/data-access.html

As of 2016, TxDOT did not retain CRIS information prior to 2010. All existing information prior to 2010 was transferred to the University of Texas at Austin Center for Transportation Research (CTR) (https://ctr.utexas.edu/). Researchers at CTR, who frequently use the data to investigate traffic-related questions and publish in peer-reviewed academic journals, initially refused to grant access to the 2003-2009 CRIS data.

We submitted a public information request to the University of Texas at Austin Center for Transportation Research to release the CRIS data. The public information request was appealed to the Texas attorney general's office by the University of Texas. In a letter dated January 23, 2017, the assistant attorney general of Texas ruled that the University of Texas must release the CRIS data. The University of Texas at Austin open records coordinator electronically transferred the complete record of accidents in CRIS from 2008-2009 on February 15, 2017. Michael R. Murphy, Research Engineer at the University of Texas at Austin Center for Transportation Research, subsequently sent data from 2003-2007 on a CD on April 28, 2017. Documents related to the Public Information Request and all of the original CRIS data are posted here: https://justinpgallagher.com/research.html. The original CRIS data

files can also be downloaded from Dataverse (https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/GGLKEM.).

The CRIS data include a description of the accident type for each accident. Table 1 provides a list of accident type in Houston from 2003-2005 by frequency of the type of the accident. Overall, 21% of the accidents are angle accidents and 79% are non-angle accidents.

2.1.2 Red-light Cameras

We use information on red-light camera intersections from the annual TxDOT red-light camera reports (Texas Department of Transportation (2009-16)). The earliest available reports are from 2009. These reports are compiled and published by the state of Texas using information submitted by municipalities. The reports provide the location of each camera. In addition, municipalities with a camera program are required to submit annual information on each camera, including: the date of installation, intersection speed limits, total tickets issued, and estimates of average daily traffic.

2.1.3 Intersection Engineering Characteristics

We collect information on a number of structural intersection characteristics, including whether the intersection has a median separating traffic, the speed limit, the number of lanes, and whether the intersection includes a frontage road. A frontage road is defined as a road running parallel to a highway that is often used as an access point to the highway. The intersection characteristics were collected using afore mentioned TxDOT reports, Google Maps and Google MyMaps from June-July 2016. The dates of the images used to collect the data roughly match the end of our panel period.

2.1.4 Average Daily Traffic

We collect average daily traffic (ADT) from three sources. The first source is the red-light camera enforcement intersection reports (TxDOT (2009-16)). Average daily traffic data from the TxDOT reports are likely from around

the year the reports were first filed in 2008 or the year of installation. ADT data retained in the annual reports do not change year-to-year at the same intersection. We infer, in turn, that the ADT information was not collected again for these reports after the initial reporting year.

We also use information from two other sources that provide traffic counts in Houston and Dallas at numerous street locations (North Central Texas Council of Governments [2016] and City of Houston [2017]). The rationale for using street-based (rather than intersection-based) ADT information is so we will have a consistent ADT measure for camera and non-camera intersections in each city. The street-based ADT measures also allow for multiple ADT measurements for a subset of our camera intersections.

Intersections are assigned ADT values using GIS software by summing the ADT values for all roads at the intersection. We take the ADT points and join them to the closest road segment using GIS software. We then identify all intersections that have an ADT measurement for two or more approaches. We sum ADT from each approach if two approaches are covered. If more than two approaches are covered, then we calculate the average ADT across all approaches and multiply it by two (e.g., if all four approaches on two roads report ADT).

2.1.5   Red-light-running Tickets

We use two sources for the number of red-light-running tickets. First, red-light-running tickets issued at each camera controlled intersection in both Houston and Dallas are available from the camera reports (Texas Department of Transportation [2009-16]). The camera ticket data are reported by fiscal year (July 1 - June 30), beginning in 2008-09. There is no published 2010-11 annual report for Houston even though the Houston cameras were in operation for four months.

We also obtained red-light-running ticket data on tickets issued by Houston police. These data exclude tickets issued via the traffic cameras. We requested the data via an open records request (Municipal Court Record Number: 341JUNE17). We received the monthly intersection count of all red-light-

19

running tickets from April 2006 to December 2016.

2.1.6 Additional Information for Welfare Calculation

Accident Costs

We use accident cost estimates from a recent National Highway Traffic Safety Administration (NHTSA) report (Blincoe et al. [2015]). The NHTSA report compiles accident costs by accident severity using the KABCO scale. The scale is: K= killed, A = incapacitating injury, B = non-incapacitating injury, C = complaint of pain, and 0 = no injury. The scale matches the police coded accident data in CRIS, with the exception that CRIS uses "possible injury" instead of "complaint of pain." We also combine the "unknown" and "no injury" categories in the CRIS accident data into a single grouping to correspond to the KABCO no injury category. Note that the NHTSA report explicitly accounts for misreporting of injuries, by either those involved in the accident or the police officers filling out the reports. For example, the NHTSA report estimates that on average there are $7,789 (2010 $) in injury-related costs for the no-injury category (Blincoe et al. [2015], p251).

We use the total accident costs for the four non-fatal KABCO categories in Table D-1 (Blincoe et al. [2015], p251). These categories include estimates for direct injury related costs (medical, lost wages, and legal costs), traffic congestion costs due to the accident, property damage, and lost quality of life from accident sustained injuries.

We use the Department of Transportation's recommended value of statistical life, $8,860,000 (2010 $), as the cost of a fatal accident (Blincoe et al. [2015]).

Houston Program Cost and Revenue

We obtained Houston red light camera program cost and revenue information via the Houston city budget report for the Digital Automated Red Light Enforcement for 2009 and 2010 (https://www.houstontx.gov/budget/09budadopt/index.html and https://www.houstontx.gov/budget/10budadopt/index.html). We use operating and maintenance es-

timates from fiscal years 2008-2009 and 2009-10. The cost data are total expenditures net of transfers to the other government agencies through cost sharing programs and transfers to the cities general fund. The average cost per camera per year (across the two reporting years) is $169,509.

The total average camera revenue from red light running tickets from 2008-9 and 2009-10 is $11,381,937. As we write in the manuscript, the welfare model considers the welfare benefit of raising revenue from fixed lump sum fines rather than distortionary fiscal taxes. We follow Barrage [2016] who calculates that a dollar raised via fiscal taxes, on average, has a marginal cost of public funds of 1.49. In our analysis, for every dollar raised by the camera program, we subtract 49 cents from the cost of running the program. The yearly per camera revenue (gain) is $85,502.

The Houston budget documents do not provide information on the fixed costs of installing the cameras. We attempted to collect this information via a public records request but were unsuccessful. Instead we use estimates on the purchase and installation of a standard digital camera system from Maccubbin et al. [2001]. These estimates are also used by the Centers for Disease Control and Prevention (https://www.cdc.gov/motorvehiclesafety/calculator/factsheet/redlight.html). The cameras cost an estimated $50,000 to $60,000 to purchase and $25,000 to install. There are 66 Houston camera intersections. Most intersections require more than one camera to capture the traffic entering the intersection from all of the cross streets. There are a total of 152 cameras at the 66 camera intersections. Total estimated fixed costs are $12.9 million.

We do not include the estimated fixed camera costs in our welfare model analysis. The main reason for this is that the fixed costs roughly equate with the higher estimated ticket revenue from the first two years of the camera program. The camera revenue estimates we use are based off of fiscal years 2008-9 and 2009-10, which on average are about two (for 2007 cameras) and three (for 2008 cameras) years after the date of camera installation. We know that the number of tickets issued at red light cameras decreases over time as drivers learn of camera location and adjust behavior (see manuscript

figure 2 panel B). The yearly cost and revenue estimates are after the initial introduction of the cameras and when the cost and revenue are at levels we may expect to continue indefinitely.

Number of Residents

The number of Houston residents age 18-64 is from a 2013 report by the Houston Planning and Development Department (http://www.houstontx.gov/planning/Demographics/docspdfs/SN/AgeTotalPopulation.pdf). The source of the underlying data is the 2007-2011 American Community Survey.

Mean Wage

The Federal Reserve Bank of St. Louis calculates the average weekly wages for "employees in total covered establishments" in the Houston-Sugar Land-Baytown MSA using Bureau of Labor Statistics information (https://fred.stlouisfed.org/series/ENUC264240010SA). We calculate the average hourly wage in Houston by averaging across the four 2011 quarters and dividing by 40.

Number of Persons per Vehicle

The average number of persons per vehicle is from US Department of Transportation report that analyzes the 2001 National Household Travel Survey (USDOT [2003]). We use the mean across all personal vehicle trips. A link to the report can also be found here: https://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/highlights_of_the_2001_national_household_travel_survey/html/table_a14.html

Length of Red Light Signal

We calculate the average length of a red light in Houston using two sources. The duration of yellow lights and red-light lag times (i.e., when both directions are red) are provided for each camera controlled intersection in the annual red-light camera enforced intersection reports (Texas Department of Transporta-

tion [2009-16]). We estimate green-light durations using the *Traffic Signal Operations Handbook*, which is published by the Texas Transportation Institute and provides guidelines for minimum green-light times based on speed limit (Bonneson et al. [2009], p2-5). We calculate that the average wait time caused by having the cross street signal as green is 0.91 minutes.

We also use the *Traffic Signal Operations Handbook* to calculate the estimated length of the turning only phase (when through traffic on both streets have a red signal). The *Traffic Signal Operations Handbook* provides a formula for calculating the recommended left-turn only phase. The formula considers whether the cross roads are major or minor roads, the traffic volume, and the number of lanes (Bonneson et al. [2009], p31). We have all of this information for the 66 camera intersections. Using the formula we calculate that the average length of the turning only phase is 20.88 seconds.

Number Additional Vehicles Stopping at Camera Intersections

We use an estimate of the number of additional vehicles stopping at camera intersections (rather than continuing through the intersection) to estimate the travel time delay in Houston under the camera program.

The intuition behind our calculation is that we can use the observed number of annual red-light running tickets to work backwards to determine the number of vehicles that were running red lights before the cameras were installed. We know that the number of tickets decreases dramatically between the first and second years after camera installation (manuscript figure 2 panel B). Further, we can use observational (count) estimates from the literature for the reduction in red-light running from the year before the cameras (year 0) to the first year after the cameras (year 1).

We use two sources of information to estimate the reduction in vehicles passing through an intersection when the light turns red. A conservative estimate from the literature is that the number of red-light running vehicles decreases by 42% in the first year of a camera program, relative to the year before the program (McCartt and Hu [2014]). Next, we use red-light ticket data from two camera intersections in Dallas to benchmark how the number

of red-light tickets change in the second and third years of a camera program (Texas Department of Transportation (2009-16)). These are the only two camera intersections with cameras installed (in either Houston or Dallas) after the mandated annual camera reports began documenting the number of tickets in 2009, and which had an installation month in June or July. The installation month is important since the ticket data are reported by fiscal year (July 1-June 30). While we are only using two camera intersections for our benchmark, it is well documented that the number of tickets decreases over (at least) the first few years of a red-light camera program as drivers learn about the cameras. For example, Fisher [2017] examines ticketing patterns at more than 300 Chicago red light cameras including during their installation. During the first year the number of tickets issued drops by approximately 50%. The number of tickets continues to slowly decline for (at least) several more years.

We estimate that there were 573,500 vehicles that ran a red light at one of the 66 camera intersections in the year before cameras were installed. This translates to 721,463 minutes of additional waiting at a traffic intersection (calculated by multiplying by a 1.26 minute delay), 0.54 minutes of additional waiting per person (dividing by the population), and 0.0082 minutes of waiting per person per camera (dividing by 66).

Our baseline estimate is conservative. Since our calculations are derived only of off of vehicles that actually ran a red light. The baseline calculation does not account for vehicles under the camera program that stop when the light is yellow and might have passed through the intersection before the light turned red. This group of vehicles is likely to be large, potentially of a similar magnitude as the number of vehicles that stop running red lights.

2.1.7 Houston Referendum Media and Online Search Information

Figure 3 in the manuscript provides evidence from three media sources on how well aware Houston residents were of the red light camera voter referendum.

Newspaper Stories

The source of the newspaper stories is Houston's largest daily newspaper, the

Houston Chronicle (https://www.chron.com/). We use the search feature provided on the website to search for stories that include "red light camera" or "Annise Parker". We verified that each story from the search was unique, contained one of the search phrases, and had a publication date between November 2008 - November 2012.

TV News Stories

The TV news story information is from closed captioning data provided by Metro Monitor (https://metromonitor.com/). The closed captioning data cover all programing on Houston's ABC, CBS, FOX, and NBC affiliates. We use a search function available to subscribers to search the closed captioning text information for "red light camera" or "Annise Parker". We verified that each story from the search was unique, contained one of the search phrases, and had a publication date between November 2008 - November 2012.

Online Searches

The online search frequency is from Google Trends (https://trends.google.com/trends/). We limited the search geography to Houston, TX and the time period to November 2008 - November 2012. We searched simultaneously for the terms "red light camera" and "Annise Parker".

2.2 GIS Data Processing

We start with the police-recorded (accident-level) CRIS data for each accident in the cities of Houston, Dallas, and San Antonio. Next, we limit the sample to accidents that occur within 200 feet of an intersection. We use GIS software (ESRI ArcMap) to buffer the intersections by 200 feet and intersect them with the accident (latitude/longitude) point shapefile to create an accident-level output file containing all accidents within 200 feet of an intersection and 50 feet of a road.

We use the US Census TIGER/Line USA Major Roads shapefile as the street map in GIS to determine the list of intersections in each city. All

camera intersections in Houston and Dallas are included on the street map. We exclude non-camera intersections in Houston within one-half mile of a camera intersection, for camera intersections may impact driving behavior at nearby intersections (e.g., Shin and Washington [2007]; Høye [2013]). We also exclude Houston non-camera intersections if the intersection does not have associated ADT information, or if there are no recorded accidents from 2003-2014. Finally, as a reliability check, we visually check all remaining intersections in GIS to confirm that there are no duplicate intersections and that we are not double-counting an intersection.

2.3 San Antonio Intersection Risk Index

We determine the most dangerous intersections in San Antonio using a simple weighting method. We adapt the weighting formula that Stein et al. [2006] utilized to rank order Houston intersections by accident risk. Our intersection risk index is a weighted sum of San Antonio accidents for 2003 using the following weights: 4 = fatal, 3 = incapacitating, 2 = non-incapacitating or possible, 1 = unknown injury or no injury. Under the weighting scheme of Stein et al. [2006], fatal accidents receive a weight of 3. Still, the two weighting schemes produce similar intersection risk rankings. We select the 66 most dangerous intersections to use as our placebo treatment intersections, since this is the same number of 2006 and 2007 Houston camera intersections.

3 Robustness Analysis

3.1 Houston and Houston-Dallas Samples

This section provides additional analysis for the Houston and Houston-Dallas samples, including figures referenced in the manuscript. Unless otherwise noted, these samples consider accident data from 2008-2014 and estimate the propensity score model using intersection data from the pre-referendum period (2008-2010).

Figure 1 shows the distribution of the propensity scores in the Houston and Houston-Dallas samples. The overlap in the propensity scores for the treatment and control intersections is best for the Houston sample. This is one reason why our preferred estimates, and those that we emphasize in the text, are from the Houston sample. The distribution of the propensity scores in the Houston-Dallas sample is more bimodal. The Dallas camera propensity scores tend to be clustered below 0.5 and the Houston camera propensity scores above 0.5.

Figure 2 shows the location of the camera enforced (treated) and non-camera enforced (control) intersections in our Houston sample. Overall, there is a similar geographic distribution. Most of the camera and non-camera control intersections in our sample are to the west of downtown.

Figure 3 is constructed in exactly the same was as Figure 6 in the manuscript, except that this appendix figure plots accident data for Houston intersections in the Houston-Dallas sample (rather than Houston intersections in the Houston sample). Figure 3 panel A shows the pre-referendum level of angle (y-axis) and non-angle (x-axis) accidents for each Houston camera intersection. Intersections in our Houston-Dallas sample are marked with the black symbols, and those intersections dropped via propensity score trimming with the white symbols. Figure 3 panel B shows the pre-referendum to post-referendum shift in the number of angle and non-angle accidents for each camera intersection. One potential concern is that the camera intersections in our Houston-Dallas sample may respond differently to the camera program than those intersections dropped from the analysis. Panel B shows that this is mostly not the case. The shift in accidents for the camera intersections in our sample (black symbols) is similar to that for camera intersections not in our sample (gray symbols), after accounting for yearly accident trends and fixed intersection characteristics. If anything, the Houston intersections included in the sample suggest a more positive view of the camera program, as compared to those excluded from the sample. On average, ending the camera program led to both more of an increase in angle accidents and less of a decrease in non-angle accidents for the in-sample intersections (as compared to intersections not in

our sample).

Table 1 provides a list of accident causes (types) in Houston from 2003-2005 by frequency and type of accident. There are 45 different accident types, 10 of which include the word "angle" and other details (e.g., "Angle: Both Going Straight"). Five non-angle accident types ("OMV other," "other," "not reported," "undetermined," and "reported invalid") are combined into the category "other" listed in the table. Overall, 21% of the accidents are angle and 79% non-angle. The most common non-angle accident type is "Single Vehicle - Going Straight."

In Table 2, we divide the non-angle accidents into five subgroups we label as: head on, single vehicle, turning, rear end, and other. The table shows coefficient estimates from our difference-in-differences model using the Houston and Houston-Dallas samples, while limiting the dependent variable to each of the accident subgroups. The results indicate that, when the cameras are turned off, the reduction in non-angle accidents (manuscript Table 3) is mostly attributable to a reduction in rear end accidents.

Table 3 estimates the effect on average daily traffic (ADT) from ending the camera program using a difference-in-differences OLS model. The estimation samples only include intersections in our Houston and Houston-Dallas samples that have at least one ADT observation both before (2007-2010) and after (2011-2014) the referendum. The estimates imply a modest increase in traffic at Houston camera intersections after electronic monitoring ended of between 4% and 18%, although none of the estimates is statistically significant. We note, however, that if there is measurement error in the interpolation procedure used to assign the ADT data to intersections (see Section 2.1.4), then the ADT estimates are likely to be attenuated towards zero.

Table 4 shows four additional robustness results. Our strategy to estimate the effect of the cameras on the number of accidents is to use the voter referendum as a natural experiment. At the same time, since our panel of accident data begins in 2003, we are also able to estimate the introduction of the Hous-

ton cameras in 2006 and 2007.[18] Panel A estimates the effect of introducing the cameras using our main Houston sample. Panel B estimates the effect of introducing the cameras using a second Houston sample. In Panel B, accident characteristics from the three years before the first Houston camera was installed (2003-2005) are used to select our treatment and control intersections using our logit model. The pre-trimmed Houston intersections are the same as in our main Houston sample. However, the final control and treatment intersections are selected based on pre-program (rather than pre-referendum) accident characteristics.[19]

The estimates in panels A and B both suggest that turning on the cameras reduced the number of overall traffic accidents by about 30%. This finding is in contrast to our camera removal estimates that leverage the exogenous referendum. However, the entry results are not too surprising given the placebo camera program simulation results (see Table 8 column 4) discussed in Section 1.3.1. An analysis of the introduction of a placebo camera program using a difference-in-differences model suggests that there is a large (39% to 58%) and statistically significant reduction in total accidents attributable to the (non-existent) camera program.

There is a difference between the simulation results in Table 8 and the entry results in Table 4 in how we select the control group of intersections. In the simulation exercise, the "treated" intersections are the 50 intersections with the highest number of accidents in the pre-period, while the control group of intersections are the 50 intersections with the next highest number of accidents. In Table 4 panels A and B we attempt to more formally control for pre-trends by selecting the estimating sample using a logit model. If we assume that

---

[18]We estimate the model using 2003-05 as the pre-period and 2008-10 as the post-period. We drop 2006 and 2007 so as to maintain a balanced panel. We do not know when the cameras were installed during the two installation years. Even if we did, monthly accident data that would permit a partition of the accident data during the installation years into the pre- and post-periods is not available.

[19]The accident characteristics in the logit model (A_{it}) are the same except that average daily traffic is not included, as this information is not available from the earlier time period. The availability of the ADT data, the opportunity to provide out-of-city control group estimates with Dallas camera intersections, and a better propensity score overlap, are the reasons why our preferred Houston sample is selected using 2008-2010 characteristics.

29

the actual Houston camera program treatment effect is zero (we estimate a statistically imprecise removal estimate of -4% in our preferred sample), then matching on pre-period trends is able to account for about half of the effect of mean reversion when analyzing the endogenous introduction of the cameras.

Table 4 panel C uses the same alternative Houston sample as in panel B, except returns to our preferred approach that focuses on the exogenous removal of the cameras. The point estimate for the effect on total accidents, -3%, is nearly identical to the point estimate from our main Houston sample (-4%). Neither of the point estimates are statistically different from zero.

Table 4 panel D shows estimates from our main Houston model (i.e. using a logit model and 2008-10 intersection characteristics), except that we use a different propensity score trimming rule to select the final sample. In panel D, we drop all intersections that either have estimated propensity scores less than 0.01, or fall into a five unit bin (0.00 to 0.05, 0.051 to 0.10, etc.) which does not have both treatment and control intersections. This trimming rule results in a much larger group of control intersections. As with the results in panel C, the estimated impact on total accidents is similar to our preferred sample. However, the effects on angle and non-angle accidents are smaller in magnitude.

Table 5 shows the same point estimates and standard errors from our main Houston and Houston-Dallas samples (manuscript Table 3), as well as a second measure of statistical precision from running a series of permutation tests. For each sample and outcome, we randomly allocate the intersections into the treatment and control groups (conditional on the maintaining the original group sizes) and calculate the estimated coefficient using our model. We repeat this process 10,000 times for each sample and outcome, and use these results to construct confidence intervals and probability values. The table displays the probability value for the null hypothesis that each coefficient is equal to zero, and for the null hypothesis that the angle and non-angle coefficients are equal. The permutation-based probability values suggest that the statistical significance of the estimated coefficients may be overstated. For example, the clustered standard errors imply that the estimated coefficient for angle

accidents (panel A, column 1) from the Houston sample is statistically different from zero at the 5% significance level, whereas it would fail to be significant at the 16% significance level using the probability value from the permutation test. However, the permutation tests still indicate that the angle and non-angle coefficients are statistically different from each other at (at least) the 5% significance level.

3.2 Houston Frontage Sample

This section provides supporting documentation for the frontage sample. Most of the intersections that receive a camera in Houston are on frontage roads. Our main estimation samples are not balanced. In our Houston sample, 75% of the camera intersections are on frontage roads, while just 4% of the non-camera (control) intersections are on frontage roads. In our Houston-Dallas sample, the difference is smaller, but still unbalanced. 75% of the Houston camera intersections are on frontage roads, while 38% of the Dallas intersections are on frontage roads. (See manuscript Table 2). To help address this concern, we also estimate our model on a sample of Houston intersections that are all located on frontage roads. Overall, we estimate a very similar effect on total accidents (-6% using the Houston frontage sample and -4% in our main Houston sample).

We construct this sample by visually identifying all intersections with frontage roads alongside freeways (or similar roads) in Houston. This is done in Google Maps. We then apply filters for being located within the city and not within a half mile of a treated intersection.

Table 6 shows yearly means for accident and intersection characteristics for camera and non-camera intersections in the frontage road sample before and after trimming with the propensity score. The means are taken over the pre-referendum years 2008-2010. We do not calculate mean values for ADT, as we do not have ADT information for most of the intersections in the frontage sample.

Figure 4 shows the distribution of the propensity scores in the Houston

frontage road sample. There is reasonable overlap in propensity scores between the treatment and control groups, as there are just two bins with any observations that fail to have both treatment and control intersections. Overall, the treatment intersections appear to be roughly uniformly distributed, while the control intersections are more likely to have lower propensity scores.

Figure 6 shows the yearly average number of angle, non-angle, and injury accidents for camera (treatment) and non-camera (control) intersections in the frontage sample. There are very similar trends between the treatment and control groups for the three pre-referendum years (2008-2010). The yearly number of accidents is slightly higher for camera intersections for each accident type in the years before the referendum, and mostly lower after the referendum.

Figure 7 panel A shows the pre-referendum level of angle (y-axis) and non-angle (x-axis) accidents for each Houston frontage sample camera intersection. Panel B plots the pre-referendum to post-referendum shift in the level of angle and non-angle accidents for each camera intersection, after accounting for yearly accident trends and fixed intersection characteristics.

4 References

The impact of red light cameras (photo-red enforcement) on crashes in virginia. 2008.

Christopher Abbess, Daniel Jarrett, and Catherine C. Wright. Accidents at blackspots: Estimating the effectiveness of remedial treatment, with special reference to the "regression-to-mean" effect. *Traffic Engineering and Control*, 22(10), 1981.

Joshua D. Angrist and Jorn-Steffen Pischke. Undergraduate econometrics instruction: Through our classes, darkly. *Journal of Economic Perspectives*, 31(2), 2017.

Lint Barrage. Online appendix for optimal dynamic carbon taxes in a climate-economy model with distortionary fiscal policy. *Working Paper*, 2016.

L. J. Blincoe, T. R. Miller, E. Zaloshnja, and B. A. Lawrence. The economic and societal impact of motor vehicle crashes, 2010 (revised). Technical report, May 2015.

James Bonneson, Srinivasa Sunkari, and Michael Pratt. Traffic signal operations handbook. Technical report, March 2009.

Mark Burkey and Kofi Obeng. Technical report.

A. Colin Cameron and Pravin K. Trivedi. *Microeconometrics*. Cambridge Press, 2005.

Alicia Carriquiry and Michael Pawlovich. From empirical bayes to full bayes: Methods for analyzing traffic safety data. *White Paper, Iowa State University*, 2004.

City of Houston. Traffic counts, 2017. URL http://data.ohouston.org/dataset/traffic-counts.

Forrest M. Council, Bhagwant Persaud, Kimberly Eccles, Craig Lyon, and Michael S. Griffith. Safety evaluation of red-light cameras. *Federal Highway Administration*, (FHWA-HRT-05-048), 2005.

Alena Erke. Red light for red-light cameras? a meta-analysis of the effects of red-light cameras on crashes. *Accident Analysis and Prevention*, 41, 2009.

Paul J. Fisher. Green light for rent-seeking? an investigation of chicago's red light camera program. *Undergraduate Thesis*, 2017.

Ezra Hauer. Empirical bayes approach to the estimation of "unsafety": The multivariate regression method. *Accident Analysis and Prevention*, 24, 1992.

Ezra Hauer. *Observational Before-After Studies in Road Safety.* Emerald Group Publishing Limited, 1997.

Ezra Hauer. On the estimation of the expected number of accidents. *Accident Analysis and Prevention*, 18, 2006.

Ezra Hauer, Douglas W. Harwood, Forrest M. Council, and Michael S. Griffith. Estimating safety by the empirical bayes method: A tutorial. *Transportation Research Record*, 1784, 2002.

Alena Høye. Still red light for red light cameras? an update. *Accident Analysis and Prevention*, 55, 2013.

Wen Hu, Anne T. McCartt, and Eric R. Teoh. Effects of red light camera enforcement on fatal crashes in large us cities. *Journal of Safety Research*, 42, 2011.

Myunghoon Ko, Srinivas Reddy Geedipally, and Troy Duane Walden. Effectiveness and site selection criteria for red light camera systems. *Transportation Research Record: Journal of the Transportation Research Board*, 2013.

Myunghoon Ko, Srinivas Reddy Geedipally, Troy Duane Walden, and Robert Carl Wunderlich. Effects of red light running camera systems installation and then deactivation on intersection safety. *Journal of Safety Research*, 62, 2017.

Barbara Langland-Orban, Etienne E. Pracht, and John T. Large. Red light cameras unsuccessful in reducing fatal crashes in large us cities. *Health Behavior & Policy Review*, 1, 2014.

Dominique Lord and Srinivas Reddy Greedipally. Safety effects of the red-light camera enforcement program in chicago illinois. *Technical Report, Chicago Tribune*, 2014.

Robert P. Maccubbin, Barbara L. Staples, and Arthur E. Salwin. Automated enforcement of traffic signals: A literature review. Technical Report 0900610D-02, Federal Highway Administration, 2001.

Hani S. Mahmassani, Joseph L. Schofer, Breton L. Johnson, Omer Verbas, Amr Elfar, Archak Mittal, and Marija Ostojic. Chicago red light camera enforcement: Best practices & program road map. Technical report, Northwestern University Transportation Center, 2017.

Anne T. McCartt and Wen Hu. Effects of red light camera enforcement on red light violations in arlington county, virginia. *Journal of Safety Research*, 48, 2014.

North Central Texas Council of Governments. Historical traffic counts, 2016. URL http://www.nctcog.org/trans/data/trafficcounts/indexcdp.asp.

Bhagwant Persaud and Craig Lyon. Empirical bayes before-after safety studies: Lessons learned from two decades of experience and future directions. *Accident Analysis & Prevention*, 64, 2014.

M Powers and J Carson. Before-after crash analysis: A primer for using the empirical bayes method. tutorial. Technical report, 2004.

Srinivas S. Pulugurtha and Ramesh Otturu. Effectiveness of red light running camera enforcement program in reducing crashes: Evaluation using "before the installation", "after the installation", and "after the termination" data. *Accident Analysis and Prevention*, 64, 2014.

Richard A. Retting and Sergey Y. Kyrychenko. Reductions in injury crashes associated with red light camera enforcement in oxnard, california. *American Journal of Public Health*, 92, 2002.

Kangwon Shin and Simon Washington. The impact of red light cameras on safety in arizona. *Accident Analysis and Prevention*, 39, 2007.

Robert Stein, Ned Levine, and Tim Lomax. Criteria for red light camera intersection selection, 2006.

Texas Department of Transportation. Red light cameras - annual data reports, 2009-16. URL http://www.txdot.gov/driver/laws/red-light/reports.html.

USDOT. Highlights of the 2001 national household travel survey, 2003. URL https://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/highlights_of_the_2001_national_household_travel_survey/index.html.

T. Walden and B. Bochner. Effectiveness of red light cameras-texas statewide evaluation. *ITE Journal*, 81, 2011.

Timothy Wong. Lights, camera, legal action! the effectiveness of red light cameras on collisions in los angeles. *Transportation Research Part A: Policy and Practice*, 69, 2014.

5   Figures and Tables

Figure 1: **Distribution of Propensity Scores: Houston and Houston-Dallas Samples**
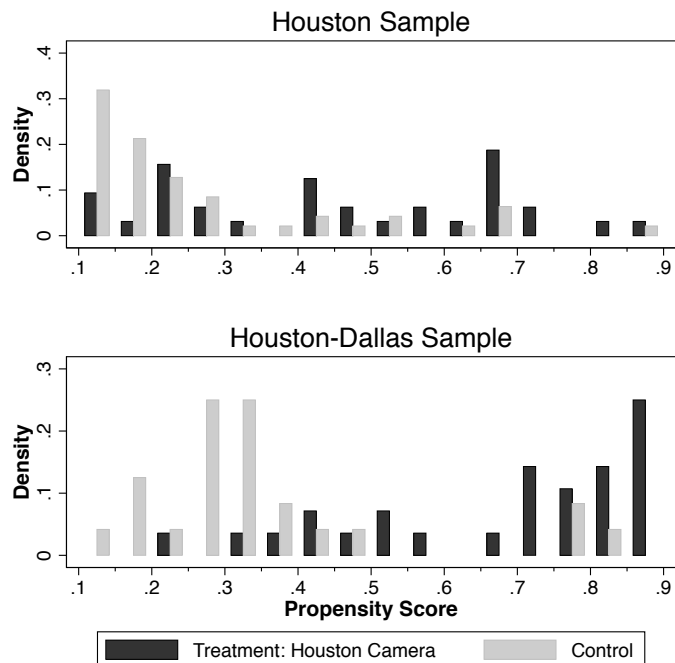
The figure shows the distribution of propensity scores in the Houston and Houston-Dallas samples. The control group of intersections are Houston non-camera intersections in the Houston sample, and Dallas camera intersections in the Houston-Dallas sample. The propensity scores are estimated by logistic regression (see text for details). Each panel plots the fraction of observations in the treatment (black bar) and control (grey bar) groups that fall within five percentage point propensity score bins. The leftmost bin is for observations with propensity scores ranging from 0.10 to 0.15, while the rightmost bin is for observations with scores from 0.85 to 0.90. Data sources: City of Houston, Texas Department of Transportation.

37

Figure 2: **Houston Camera and Non-camera Locations**

The figure plots the 32 camera intersections (triangles) and 47 non-camera intersections (squares) in our main Houston sample. Map source: US Census TIGER/Line USA Major Roads.

Figure 3: **Angle and Non-angle Accidents by Camera Intersection**
Houston-Dallas Sample

Panel A: Level of Pre-Referendum Accidents

| ▲ In-sample, Frontage | △ Not In-sample, Frontage |
| ● In-sample, Not Frontage | ○ Not In-sample, Not Frontage |

Panel B: Post-Referendum Shift in Accidents

| + In-sample | + Not In-sample |
| ■ In-sample Mean | ■ Not In-sample Mean |

Panel A shows the pre-referendum level of angle (y-axis) and non-angle (x-axis) accidents
for each Houston camera intersection based on whether the intersection is included in the
Houston-Dallas estimation sample, and by whether the intersection is on a frontage road.
Panel B plots the pre-referendum to post-referendum shift in the number of angle and non-
angle accidents for each camera intersection, after accounting for yearly accident trends and
fixed intersection characteristics.   39

Figure 4: **Distribution of Propensity Scores**
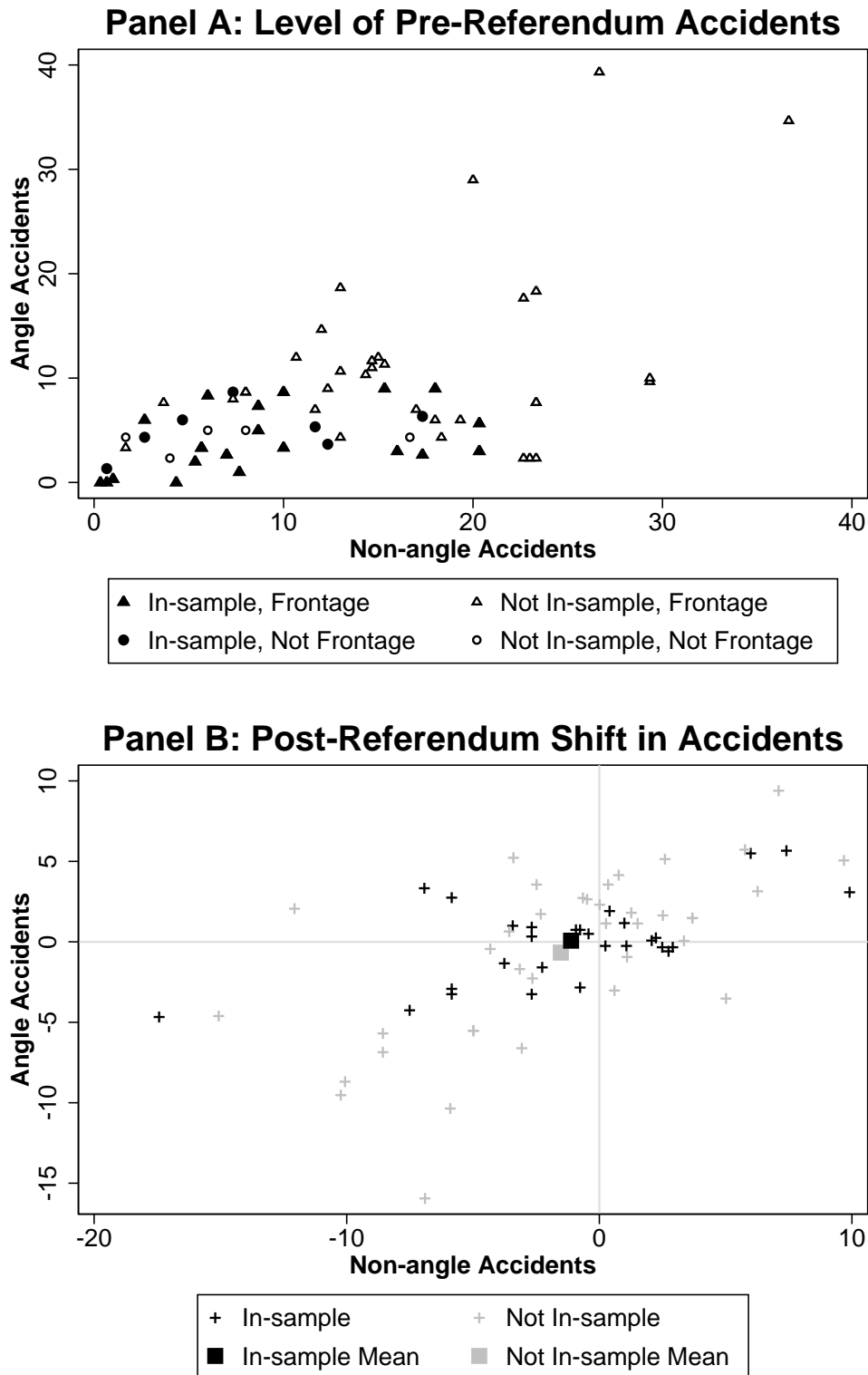

The figure shows the distribution of propensity scores in the Frontage road sample. The control group of intersections are Houston non-camera intersections. The propensity scores are estimated by logistic regression (see text for details). Each panel plots the fraction of observations in the treatment (black bar) and control (grey bar) groups that fall within five percentage point propensity score bins. The leftmost bin is for observations with propensity scores ranging from 0.10 to 0.15, while the rightmost bin is for observations with scores from 0.85 to 0.90. Data sources: City of Houston, Texas Department of Transportation.

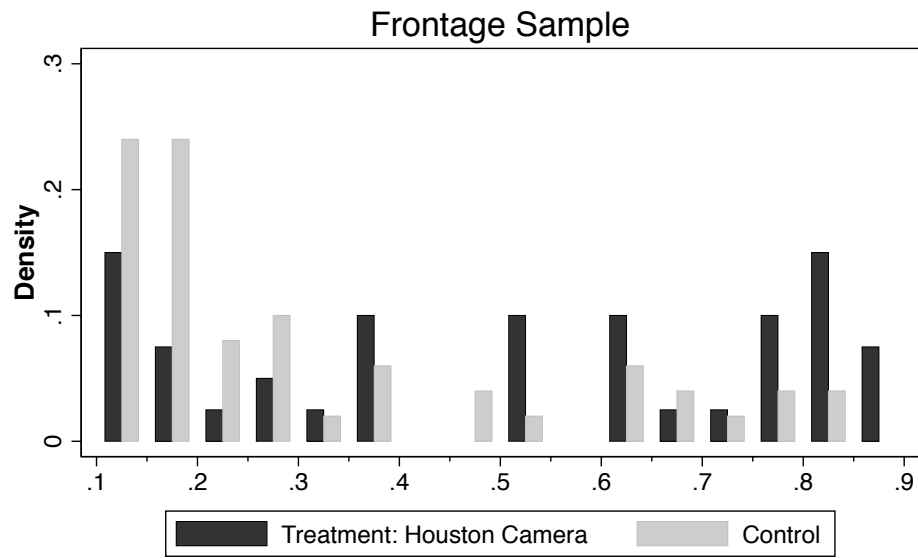Figure 5: **Houston Camera and Non-camera Locations Frontage Road Sample**
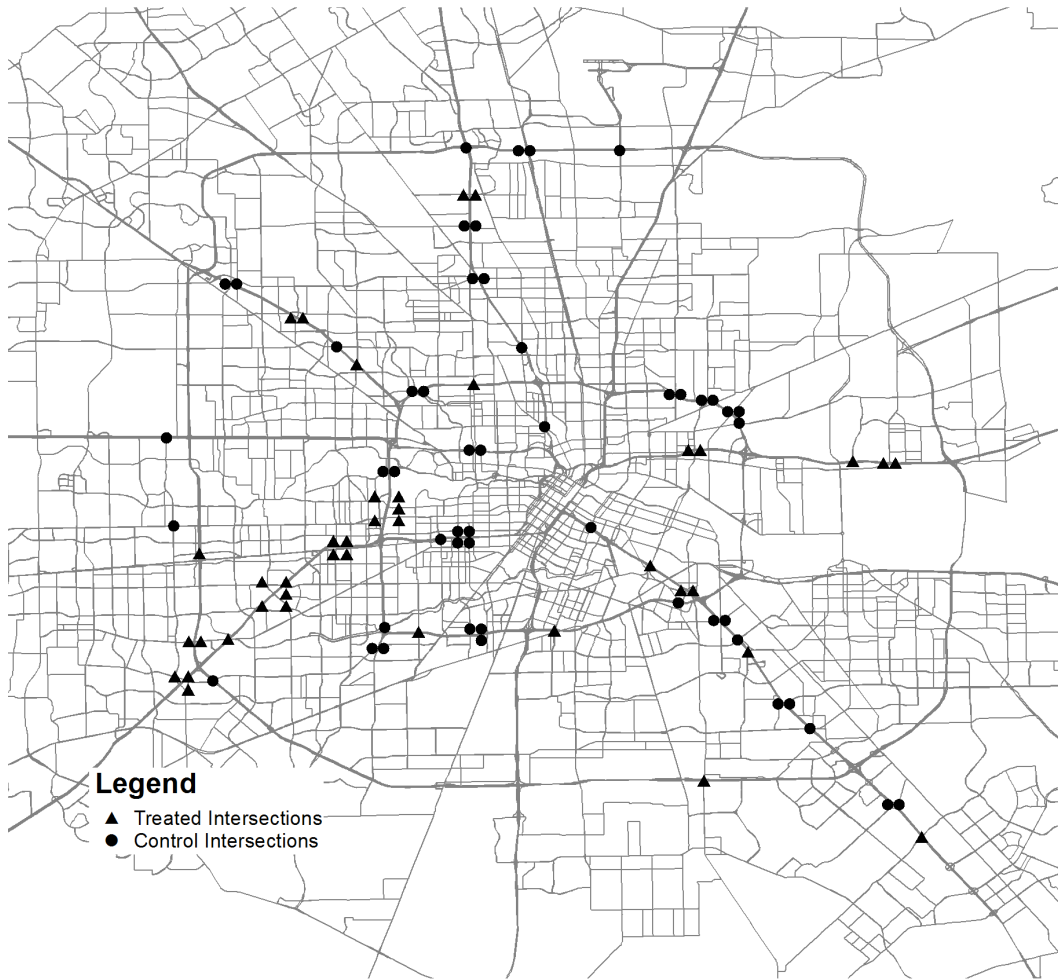
The figure plots the 40 camera intersections (triangles) and 50 non-camera intersections (squares) in our Houston frontage sample. Map source: US Census TIGER/Line USA Major Roads.

Figure 6: **Treatment and Control Intersection Accident Trends Frontage Road Sample**
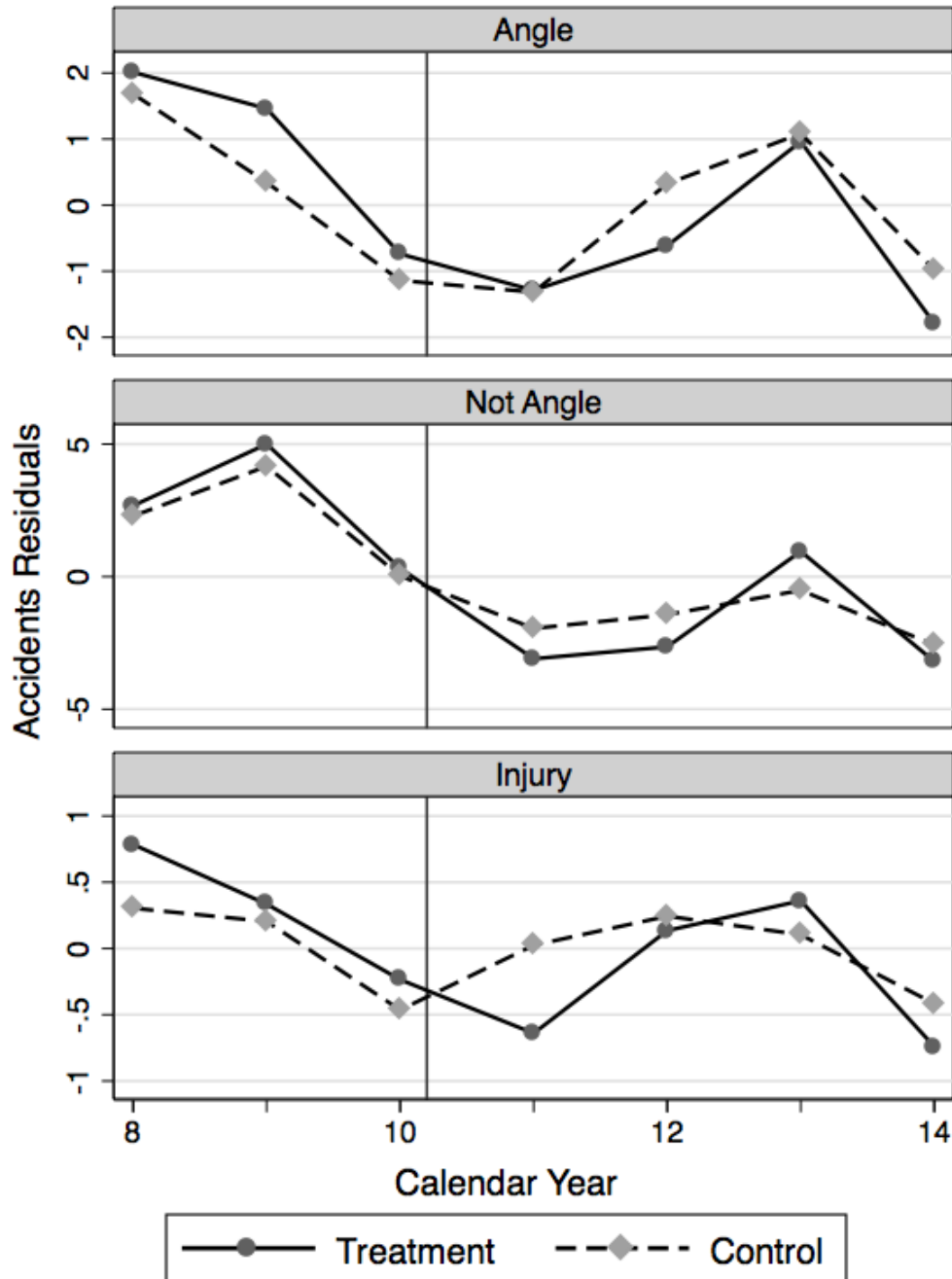
The figure plots yearly accident residuals from an OLS regression of yearly angle (row 1), non-angle (row 2), and injury (row 3) accidents on a vector of intersection fixed effects. The residuals are plotted separately for the control and treatment intersections. Treatment and control intersections are Houston camera and propensity score matched non-camera intersections (2008-2010) located on frontage roads. The accident data from 2010 are multiplied by 6/5 before running the regression, in order to account for 10 months of available data. Data source: Texas Department of Transportation.

Figure 7: Angle and Non-angle Accidents by Camera Intersection Frontage Road Sample
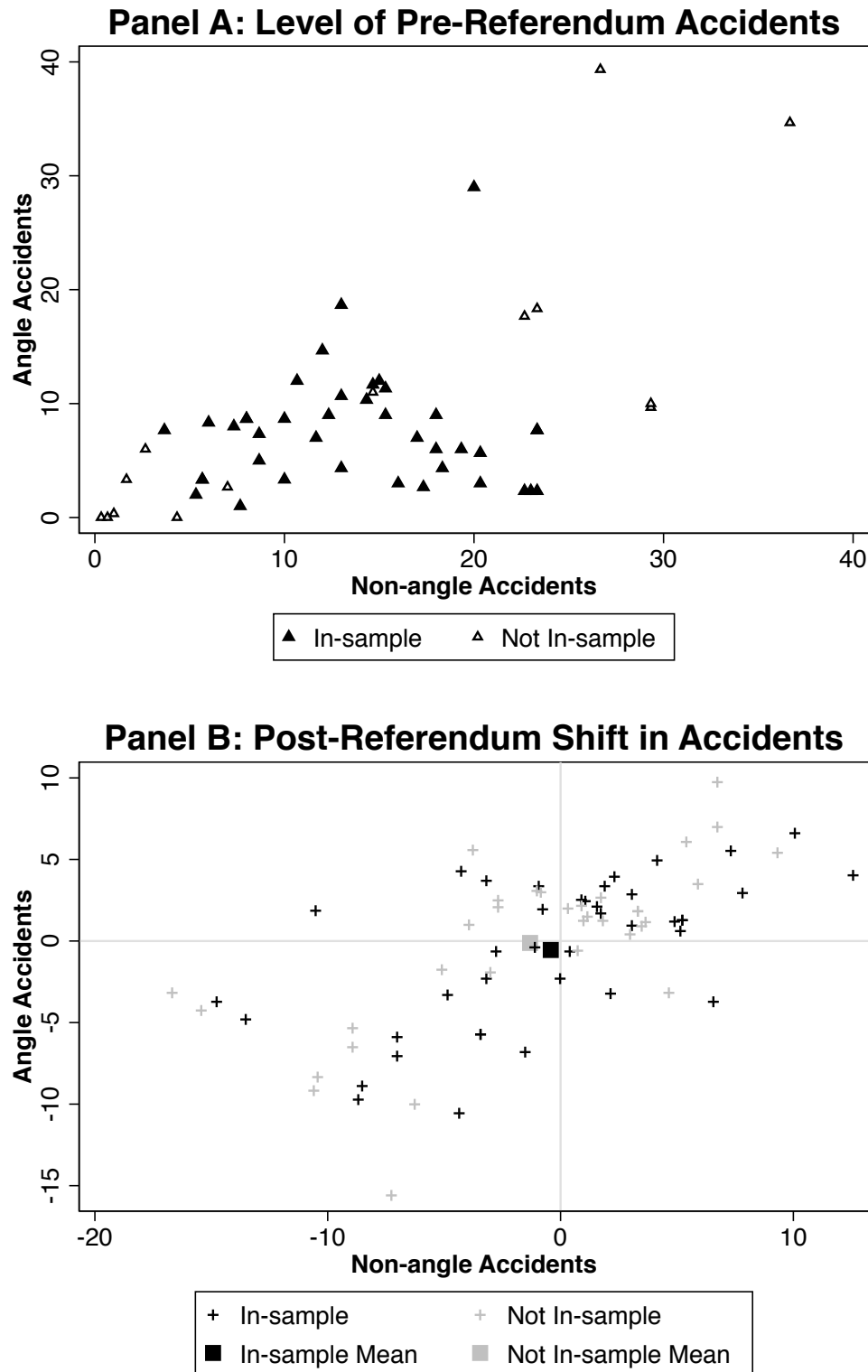
Panel A: Level of Pre-Referendum Accidents

**Panel B: Post-Referendum Shift in Accidents**

Panel A shows the pre-referendum level of angle (y-axis) and non-angle (x-axis) accidents for each Houston frontage sample camera intersection. Panel B plots the pre-referendum to post-referendum shift in the level of angle and non-angle accidents for each camera intersection, after accounting for yearly accident trends and fixed intersection characteristics.

Table 1: **Frequencies of Accident Subtypes for Angle and Non-angle Accidents**

Accident Group	Accident Type	Percent of Group	Percent of Total
Angle	Both Going Straight	0.779	0.163
	Straight - Backing	0.038	0.008
	Straight - Stopped	0.003	0.001
	Straight - Right Turn	0.062	0.013
	Straight - Left Turn	0.103	0.022
	Both Right Turn	0.000	0.000
	Right Turn - Left Turn	0.002	0.000
	Right Turn - Stopped	0.006	0.001
	Both Left Turn	0.002	0.001
	Left Turn - Stopped	0.004	0.001
	Total		**0.209**
Non-angle	Single Vehicle - Going Straight	0.216	0.171
	Single Vehicle - Right Turn	0.008	0.006
	Single Vehicle - Left Turn	0.010	0.008
	Sincle Vehicle - Backing	0.050	0.039
	Single Direction - Rear End	0.214	0.169
	Single Direction - Sideswipe	0.147	0.116
	Single Direction - Straight - Stopped	0.125	0.099
	Single Direction - Straight - Right Turn	0.021	0.016
	Single Direction - Straight - Left Turn	0.030	0.023
	Single Direction - Both Right Turn	0.003	0.002
	Single Direction - Right Turn - Left Turn	0.000	0.000
	Single Direction - Right Turn - Stopped	0.000	0.000
	Single Direction - Both Left Turn	0.005	0.004
	Single Direction - Left Turn - Stopped	0.000	0.000
	Opposite Direction - Both Going Straight	0.014	0.011
	Opposite Direction - Straight - Backing	0.011	0.009
	Opposite Direction - Straight - Stopped	0.001	0.000
	Opposite Direction - Straight - Right Turn	0.000	0.000
	Opposite Direction - Straight - Left Turn	0.067	0.053
	Opposite Direction - Backing - Stopped	0.005	0.004
	Opposite Direction - Right Turn - Left Turn	0.001	0.001
	Opposite Direction - Right Turn - Stopped	0.000	0.000
	Opposite Direction - Both Left Turn	0.001	0.000
	Opposite Direction - Left Turn - Stopped	0.000	0.000
	Parking - Straight	0.015	0.012
	Parking - Right Turn	0.000	0.000
	Parking - Left Turn	0.000	0.000
	Parking - Stopped	0.001	0.001
	Other - Both Parking	0.001	0.000
	Other - Both Backing	0.004	0.003
	Other	0.052	0.041
	Total		**0.791**

The table provides a list of accident causes (types) in Houston from 2003-2005 by frequency of the type of the accident. There are 45 different accident types, 10 of which include the word "angle" and other details (e.g., "Angle: Both Going Straight"). Five non-angle accident types ("OMV other," "other," "not reported," "undetermined," and "reported invalid") are combined into the category "other" listed in the table. Source: Texas Department of Transportation.

Table 2: **The Effect on Accidents from Ending the Camera Program: Non-angle Subgroup Analysis**

Dependent Variable:	(1) Houston	(2) Houston-Dallas
Head On	.152	-.575
	(.638)	(1.152)
Single Vehicle	-.007	-.121
	(.208)	(.232)
Turning	-.139	-0.200
	(.16)	(.155)
Rear End	-.268**	-.289*
	(.123)	(.159)
Other	.265	-.522
	(.208)	(.456)
Treatment Intersections	32	28
Control Intersections	47	24

Our main model estimates pool together all non-angle accidents. Table 1 lists all of the non-angle accident classifications. This table breaks the non-angle accident group into five subgroups we label as: Head On, Single Vehicle, Turning, Rear End, and Other. The table shows coefficient estimates (and standard errors) from estimating our difference-in-differences model using the Houston and Houston-Dallas samples, while limiting the dependent variable to each of the accident subgroups. Standard errors are robust to heteroskedasticity and clustered by intersection, $* < 0.10$, $** < 0.05$, $*** < 0.01$.

Table 3: **The Effect on Average Daily Traffic from Ending the Camera Program**

	(1) Houston Sample	(2) Houston-Dallas Sample
Panel A: OLS		
*After Removal * Treated*	2,308	7,271
	(10,129)	(11,604)
Percent Change	4	18
Panel B: OLS, IPS Weighted		
*After Removal * Treated*	4,719	2,970
	(10,985)	(12,095)
Percent Change	9	7
Treatment	26	22
Control	10	19

This table shows the coefficient of interest from estimating our difference-in-differences model using OLS on the (2008-2014) Houston and Houston-Dallas samples. The dependent variable is the average daily traffic (ADT) at each intersection. Intersection ADT values are not available for each year, nor for every intersection. The intersections included in the models have one observation before the program (measured between 2008 and 2010), and one observation after the program (measured between 2011 and 2014). The intersections included in the analysis are a subset of those intersections in our complete Houston and Houston-Dallas samples. Panel B uses inverse propensity score weighting. Standard errors are robust to heteroskedasticity and clustered by intersection, $* < 0.10$, $** < 0.05$, $*** < 0.01$.

Table 4: **Red Light Camera Program Robustness Specifications: Turning Cameras On, Pre-program Sample, Different Propensity Score Trimming**

Dependent Variable:	(1) Angle	(2) Non-angle	(3) Total	(4) Injury
Panel A: Entry Houston Sample				
*After Removal * Treated*	-.610***	-.273***	-.392***	-.335**
	(.132)	(.094)	(.082)	(.159)
Equality of Angle and Non-angle, p-value	0.030			
Treatment Intersections	32	32	32	32
Control Intersections	47	47	47	47
Panel B: Entry 2003-5 Sample				
*After Removal * Treated*	-.111	-.090	-.087	-.298*
	(.126)	(.119)	(.091)	(.156)
Equality of Angle and Non-angle, p-value	0.902			
Treatment Intersections	25	25	25	25
Control Intersections	40	40	40	40
Panel C: Removal 2003-5 Sample				
*After Removal * Treated*	.099	-.093	-.028	.072
	(.136)	(.118)	(.108)	(.186)
Equality of Angle and Non-angle, p-value	0.137			
Treatment Intersections	25	25	25	25
Control Intersections	40	40	40	40
Panel D: Common Support	-.007	-.153**	-.096	-.372***
*After Removal * Treated*	(.090)	(.078)	(.075)	(.126)
Equality of Angle and Non-angle, p-value	0.001			
Treatment Intersections	57	57	57	57
Control Intersections	390	390	390	390

The table shows four robustness specifications. The estimates in this table should be compared to our main Houston referendum-based removal estimates from the manuscript: Table 3 panel A, and manuscript Table 4 panel A column (1). Panel A of this table estimates the effect of turning on the cameras using our main Houston sample. Panel B estimates the effect of turning on the cameras using a propensity score trimmed sample based on intersection characteristics from 2003-5. Panel C uses the same sample as Panel B, except estimates the effect of removal. Panel D shows estimation results from a Houston sample created via an alternative propensity score trimming rule. Standard errors (in parentheses) are robust to heteroskedasticity and clustered by intersection, * < 0.10, ** < 0.05, *** < 0.01. Source: Texas Department of Transportation.

Table 5: **Red Light Camera Program Robustness Specifications: Statistical Precision Using Permutation Tests**

Dependent Variable:	(1) Angle	(2) Non-angle	(3) Total
Panel A: Permutation (Houston)			
*After Removal * Treated*	.258**	-.194*	-.038
	(.131)	(.104)	(.098)
Permutation (10,000 samples), p-value	0.161	0.219	0.805
Equality of Angle and Non-angle, p-value	0.004		
Treatment Intersections	32	32	32
Control Intersections	47	47	47
Panel B: Permutation (Houston-Dallas)			
*After Removal * Treated*	.015	-.312**	-.192
	(.176)	(.137)	(.125)
Permutation (10,000 samples), p-value	0.941	0.151	0.312
Equality of Angle and Non-angle, p-value	0.042		
Treatment Intersections	28	28	28
Control Intersections	24	24	24

This table shows the same point estimates and standard errors from our main Houston and Houston-Dallas samples (manuscript Table 3), as well as a second measure of statistical precision from running a series of permutation tests. For each sample and outcome, we randomly allocate the intersections into the treatment and control groups (conditional on the maintaining the original group sizes) and calculate the estimated coefficient using our model. We repeat this process 10,000 times for each sample and outcome, and use these results to construct confidence intervals and probability values. The table displays the probability value for the null hypothesis that each coefficient is equal to zero, and for the null hypothesis that the angle and non-angle coefficients are equal.

Table 6: **Sample Accident Intersection Characteristics**
Frontage Road Sample

	All Intersections			All Intersections, Trimmed		
	(1)	(2)	(3)	(4)	(5)	(6)
	Treatment	Control	Difference/SD	Treatment	Control	Difference/SD
Panel A: Houston Control (2008-2010)						
Accident Characteristics						
Total	22.46	5.52	1.68	21.48	17.49	0.40
Angle	8.46	2.09	1.35	7.60	6.06	0.33
Non-angle	13.99	3.43	1.64	13.88	11.43	0.38
Injury	2.04	0.64	0.95	1.83	1.57	0.18
Red-light Running	6.99	1.59	1.31	6.12	4.82	0.31
Engineering Characteristics						
Lanes	7.35	6.46	0.49	7.05	7.40	-0.19
Speed Limit	41.31	38.59	0.58	41.08	39.75	0.28
Divided	1.00	1.00	.	1.00	1.00	.
Number of Intersections	54	440		40	50	

The table shows the means for accident and intersection characteristics for the frontage road sample camera and non-camera intersections before and after propensity score trimming. The means are taken over the years 2008-2010. Data sources: City of Houston, Google Maps, North Central Texas Council of Governments, Texas Department of Transportation.

Table 7: **Welfare Model Statistics for Baseline (Conservative) Calculation**

Statistic	Value
Annual cost per camera [r]	85,006
Population age 18-65 [n]	1,331,812
Average wage [w]	28.30
Wage multiplier [σ]	0.50
Minutes delayed per capita per year [m]	0.0082
Accident injury risk per capita per year, multiplied by 100,000 [φ]:	
Fatality	0.15
Incapacitating	0.64
Non-incapacitating	5.33
Possible	20.69
No Injury	28.76
Accident injury costs (1,000's $) per person [$k_j$]:	
Fatality	8,860
Incapacitating	1,001
Non-incapacitating	276
Possible	128
No Injury	42
Accident elasticity estimates [ε_j] (point estimate, 90% CI):	
Incapacitating and Fatality	0.41 [-0.16, .98]
Non-incapacitating	0.12 [-0.26, 0.49]
Possible	0.01 [-0.21, 0.24]
No Injury	-0.01 [-0.19, 0.17]
Yearly expected accident cost [C]	73.01

The statistics in the table can be used to evaluate the social welfare of the camera program using equation 6 in the paper. The table shows the statistics used in our conservative baseline calculation. All dollar estimates in the table are in 2010 $. We multiply our injury outcome difference-in-difference coefficient estimates by -1 to make the elasticity estimates more intuitive, since we estimate the response to a reduction in cameras (i.e., the end of the program). Sources: American Community Survey, Bureau of Labor Statistics, National Highway Traffic Safety Administration, Texas Comptroller, Texas Department of Administration, Texas Transportation Institute, US Department of Transportation.

Table 8: **Monte Carlo Simulation of the Impact of Red Light Cameras**

(1)	(2)	(3)	(4)	(5)
	Empirical Bayes		Difference in Differences	
ADT Dependence	Installation	Removal	Installation	Removal
0.1	-41.21	-0.63	-58.02	0.52
	(3.23)	(8.76)	(3.34)	(9.93)
0.2	-40.00	-0.62	-57.54	0.57
	(3.22)	(8.50)	(3.38)	(9.64)
0.3	-38.32	-0.84	-56.40	0.34
	(3.24)	(8.17)	(3.50)	(9.28)
0.4	-36.22	-0.89	-54.88	0.37
	(3.30)	(7.63)	(3.68)	(8.68)
0.5	-33.55	-1.07	-53.14	0.29
	(3.39)	(7.13)	(3.91)	(8.08)
0.6	-30.03	-1.23	-51.59	0.25
	(3.49)	(6.52)	(4.20)	(7.38)
0.7	-24.50	-1.45	-50.25	0.22
	(3.74)	(6.03)	(4.62)	(6.81)
0.8	-13.29	-1.88	-48.07	0.20
	(4.31)	(5.48)	(5.39)	(6.25)
0.9	-5.89	-2.62	-39.05	0.13
	(2.06)	(4.53)	(6.82)	(5.18)

The table shows Monte Carlo treatment effects from a placebo camera program using simulated intersection accident data. We generate accident data for one thousand intersections over six years, under different assumptions for how well the safety performance function (SPF) captures the underlying variation in accidents. Column (1) shows the percent of the variation in total accidents explained by the SPF. We assign a placebo camera to the 50 intersections in the sample that have the highest level of accidents in the pre-camera period for each SPF (ADT dependence) simulation. The 50 intersections with the next highest level of accidents are selected as the control intersections. We define the first two years of the generated panel data as the pre-camera period, the next two years as having an active camera program, and the final two years as post-program. Please refer to Appendix Section 1.3.1 for details. Columns (2) and (3) show the estimated percent change in total accidents and standard error (in parentheses) for an empirical bayes analysis based on the camera installation and camera removal. Columns (4) and (5) show the same estimates using a difference-in-differences model.

Table 9: **Empirical Bayes Model Accident Estimates from Starting and Ending the Houston Camera Program**

Dependent Variable Accid	(1) All	(2) Installation Angle	(3) Non-Angle	(4) All	(5) Removal Angle	(6) Non-Angle
Panel A: Propensity Score Trimmed Houston Sample						
Safety Index, Θ	0.95*	0.84***	1.10***	0.92	1.20	0.80***
	(0.03)	(0.06)	(0.04)	(0.06)	(0.13)	(0.06)
Panel B: Non-Trimmed Houston Sample						
Safety Index, Θ	0.70***	0.57***	0.79***	0.93	1.10	0.87***
	(0.04)	(0.07)	(0.06)	(0.04)	(0.07)	(0.05)

The table shows empirical bayes model estimates from starting (installation) and ending (removal) the camera program in Houston. Standard errors are in parentheses. We estimate the empirical bayes model on the same trimmed Houston sample as we use in the text, as well as the untrimmed sample (see manuscript Table 2). Empirical bayes model estimates are usually presented as a weighted ratio of actual accidents to estimated accidents. This ratio is sometimes referred to as the safety index. A ratio of less than one is interpreted as a reduction in accidents. Please refer to Section 1.3.2 for more details. Standard errors are in parentheses, $* < 0.10$, $** < 0.05$, $*** < 0.01$ indicate that the estimated safety index is statistically significantly different that 1.

Table 10: **Statistical Techniques in the Red Light Camera Literature**

| | | Primary Method | | Level of Analysis | | Policy Change | |
| | | Before/After with Controls | Empirical Bayes | City | Intersection | Installation | Removal |
Study	Year						
Abdulsalam et al.	2017	X			X	X	
Burkey and Obeng	2004	X			X	X	
Council et al.	2005		X		X	X	
Garber et al.	2007		X		X	X	
Hu and Cicchino	2017			X		X	X
Hu et al.	2011			X		X	
Ko et al.	2013		X		X	X	
Ko et al.	2017		X		X	X	X
Langland et al.	2014			X		X	
Lord and Geedipally	2014		X		X	X	
Mahmassani et al.	2017		X		X	X	
Pulugurtha and Otturu	2014		X		X	X	X
Retting and Kyrychenko	2001	X			X	X	
Shin and Washington	2007		X		X	X	
Walden et al.	2011	X			X	X	
Wong	2014	X			X	X	

The table lists information for 16 recent camera studies. 13 of these studies use data at the intersection-level, while three studies are done at the city-level. Eight of the 13 intersection-level studies use empirical bayes as the primary means of analysis. All of the studies estimate the causal effect of the cameras from the camera installation. Three of the studies also estimate the causal effect when cameras are removed.