HEIDELBERG UNIVERSITY
DEPARTMENT OF ECONOMICS

UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

Air Quality Alerts and Don't Drive Appeals: Cautionary Evidence from Germany

Alexander Dangel

Timo Goeschl

AWI DISCUSSION PAPER SERIES NO. 718
June 2022

Air Quality Alerts and Don't Drive Appeals: Cautionary Evidence from Germany

Alexander Dangel*[a,b] and Timo Goeschl†[a,b,c]

[a]Alfred Weber Institute for Economics, Heidelberg University
[b]Heidelberg Center for the Environment, Heidelberg University
[c]ZEW – Leibniz Centre for European Economic Research, Mannheim

June 9, 2022

Abstract

We study an air quality alert program that informs the public of high ambient air pollution levels and broadcasts a Don't Drive Appeal (DDA) to encourage motorists not to drive on poor air quality days. We use fixed effects panel models and a rigorous sub-sampling method to analyze 28 months of traffic data from Stuttgart, Germany and evaluate whether DDAs reduce driving. We find DDAs inadvertently increase driving by up to 2% in Greater Stuttgart. This overall effect is driven by heightened weekend and periphery traffic during DDAs. Notably, DDAs successfully reduce city center traffic on some weekdays and for the first five days of DDA events. However, estimated traffic reductions never exceed 5% of daily traffic flows, suggesting that high switching costs and dynamic norm factors may deter most motorists from choosing the DDA's desired response. These results provide cautionary evidence about implementing DDAs to reduce driving.

Keywords: information-based regulation; voluntary policies; air quality alerts; prosocial behavior; transportation choice

JEL Classification: D91, Q52, Q53, R40

*Corresponding author. Email address: alexander.dangel@awi.uni-heidelberg.de. Address: Alfred Weber Institute for Economics, Heidelberg University, Bergheimer Str. 20, 69115 Heidelberg, Germany.
†Email address: timo.goeschl@awi.uni-heidelberg.de

1 Introduction

Air quality alerts (AQAs) have become common policy instruments in urban areas for informing the public of heightened air pollution levels and appealing for behavioral changes. Growing evidence suggests that individuals, particularly those from air-pollution sensitive populations, respond to AQAs by avoiding or rescheduling commutes (Saberian et al., 2017; Welch et al., 2005), abstaining from strenuous outdoor activity (Noonan, 2014; Ward and Beatty, 2015), forgoing leisure in outdoor recreational spaces (Graff Zivin and Neidell, 2009), and investing in protective face masks (Liu et al., 2017). However, evidence that combining AQAs with Don't Drive Appeals (DDAs) reduces private car use is sparse (Cutter and Neidell, 2009). In this paper, we investigate the impact of DDAs in a novel setting.

Previous evidence comes exclusively from North American cities, where DDAs are largely ineffective in abating car use on poor air quality days (Noonan, 2014; Sexton, 2012; Cummings and Walker, 2000) and have even inadvertently increased driving (Tribby et al., 2013).[1] Despite such shortcomings, policy-makers may still rationalize the use of moral appeals (Ito et al., 2018; Ferraro et al., 2011; Cutter and Neidell, 2009; Reiss and White, 2008) for targeted driving reductions. To model this thinking, we draw from existing modal switching models (Cutter and Neidell, 2009; Sexton, 2012; Basso and Silva, 2014) and introduce a theoretical framework for DDAs that predicts driving reductions and incorporates dynamic social norm effects.

We test this model empirically in a European metropolitan setting seemingly well-suited to DDAs due to an abundance of modal substitutes and widespread environmental preferences in its target population.[2,3] From January 2016 to April 2020, local authorities in Stuttgart, Germany raised a Particulate Matter Alert (*Feinstaubalarm*, henceforth PMA) on days with a limited atmospheric interchange capacity.[4] When local authorities

[1] These findings correspond with first-order expectations under the assumption of self-interested, utility-maximizing agents. Motorists, who pollute the air and thereby impose a negative externality on others, optimize their private well-being (including private health costs) when deciding how much to drive but do not factor in the social cost of their choices. In aggregate, this leads to a socially-inefficient pollution surplus. Policy-makers attempt to solve this collective action problem using moral levers (i.e. DDAs) or congestion management policies (i.e. transit fare subsidies) to make driving relatively more costly and shift private driving choices towards the socially-optimal level. However, we would not expect self-interested, utility-maximizing agents to be swayed by an appeal for collective benefits at a private cost, beyond its direct effect on private well-being.

[2] Stuttgart has an extensive public transportation network consisting of seventeen regional train lines, seven suburban train lines, nineteen light-rail lines, and 390 bus lines.

[3] A coalition led by the Green party has governed the state of Baden-Württemberg since 2011, Germany's first Green party state Minister-President was elected in Baden-Württemberg in 2011 and reelected in 2016 and 2021, and a Green party politician has held office as Stuttgart's Mayor since 2013.

[4] Days with a limited interchange capacity have high air pollution (PM_{10}) concentrations and tend to have

activate the PMA, they inform the public of high ambient air pollution levels, temporarily reduce public transit fares, and widely broadcast a DDA encouraging motorists to stop driving cars and to switch to riding public transit, cycling, walking, or otherwise abstaining from driving.

Our analysis of a 28-month panel of Stuttgart traffic data shows that vehicle flows across the city and at its periphery increase, on average, between 0.1% and 1.9% on days when authorities implement DDAs. We employ a dynamic linear fixed effects regression model with a robust set of controls to show this adverse DDA effect is primarily driven by weekend traffic increases and heightened periphery traffic. However, when disaggregating our analysis, we find city center traffic levels do respond as intended to DDAs on Mondays and Fridays and over the first five days of a DDA event. We use a novel regression-discontinuity-like approach to rigorously sub-sample our data and validate our findings. These results contribute to a growing pool of evidence that DDAs can be ineffective (Noonan 2014; Sexton 2012; Cummings and Walker 2000) or even counter-productive (Tribby et al. 2013) in abating driving overall, but they also provide evidence that spatially and temporally heterogeneous alert effectiveness may be obscured in aggregate analyses. Our disaggregate analysis finds that for certain times and locations, the DDA can reduce traffic by up to 5% compared to non-DDA days.

2 Background

2.1 Stuttgart's Particulate Matter Alert Program

On January 1, 2016, Stuttgart city officials introduced the PMA program as part of a multi-policy air quality plan targeting compliance with EU air quality standards.[5] During the PM season,[6] the PMA program notified residents in the greater Stuttgart

low rainfall, low wind speed, nighttime ground inversions, and low daytime atmospheric mixing layers. In these conditions, particulate matter pollution can easily accumulate to higher levels. The program targeted collective environmental benefits from emissions reductions related to to driving reductions. See Background for more details.

[5]Under EU Air Quality Directive 2008/50/EC, daily average ambient PM_{10} concentrations are not to exceed 50 $\mu\ g/m^3$ more than 35 times per calendar year. From 2004 through 2017, daily ambient PM_{10} concentrations at the Neckartor air quality monitor in central Stuttgart annually exceeded this legal threshold. The city government, under the auspices of the state government, implemented an air quality improvement plan which included establishing a low emissions zone and corresponding vehicle bans, upgrading public transit and bicycle infrastructure, investing in cleaner public transit fleets, expanding Park-and-Ride parking lots, lowering speed limits on busy streets, banning wood burning stoves during PMAs, reducing public transit fees, increasing street cleaning, and incentivizing employers to recruit employees to purchase monthly public transit tickets.

[6]Stuttgart authorities can call a PMA during the particulate matter (PM) season from October 15th to April 15th, when PM levels are typically highest.

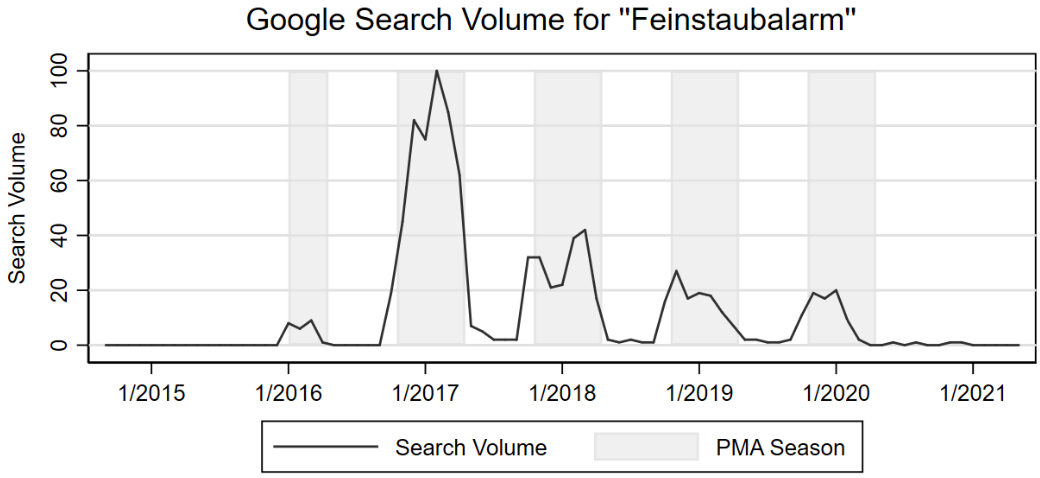

Figure 1: Google Trends search interest for "*Feinstaubalarm*" (Particulate Matter Alert) in Baden-Württemberg from January 2015 through May 2021. Search volume is relative to maximum (=100) in February 2017.

metropolitan region of upcoming and ongoing poor air quality episodes via electronic road signs, radio, television, social media, and newspapers. The PMA program's DDA encouraged motorists not to drive and instead use less-polluting transportation. In contrast to health-oriented air quality alert programs in other cities, local authorities did not explicitly warn Stuttgart residents about the negative health effects of air pollution exposure; the PMA program focused on the collective environmental benefits or so-called "quality-of-life improvements" that could result from a widespread temporary switch away from cars.[7] In early 2020, local authorities announced plans to abandon the PMA program after April of that year, citing its success in reducing air pollution in the city.[8]

Based on commuting statistics from the German Federal Employment Agency and the Baden-Württemberg State Statistical Office, we estimate that roughly 382,000 commuters (73% of individuals employed in the city) travel by car or motorcycle into and out of or within the city of Stuttgart on a given workday, compared to 66,000 (13%) who take public transit and 75,000 who walk or bike (14%).[9] In two telephone sur-

[7]Residents may certainly have acknowledged negative health impacts of air pollution exposure *ex ante*, may have become informed of them through adjacent media programming or may have inferred them from the nature and language of the program.

[8]*Stuttgarter Zeitung*. 2020. Bessere Luft in Stuttgart: Feinstaubalarm wird im April abgeschafft. January 17, 2020.

[9]Hence, for each percentage point change in daily car commuters on DDA days, we estimate that about 4,000 car commuters switch their mode of transit or work from home. We anticipate that these are

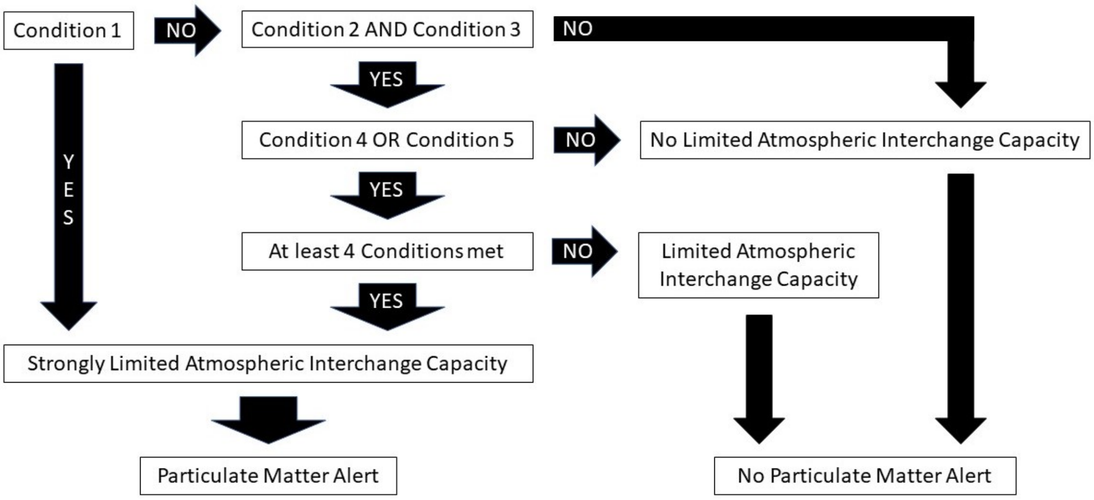

Figure 2: DWD Decision Tree for calling and ending a PMA. The "Particulate Matter Alert" outcome leads authorities to broadcast a Don't Drive Appeal (DDA). Adapted from information from DWD.

veys conducted by the city government in early 2016, 90-92% of respondents (n_1=1,008, n_2=1,004) reported having heard about the PMA program and 15-25% of respondents reported lowered car use on DDA days.[10] The survey results and online search query data (figure 1) confirm that PMA messaging arrives in the general population. However, survey responses were self-reported. Surveyors neither elicited nor observed the actual extent of driving reductions, so findings ought to be interpreted cautiously.

2.2 Particulate Matter Alert Conditions and Dynamics

Stuttgart authorities decide whether to call a PMA and broadcast a DDA using a decision tree (figure 2) based on six binary atmospheric conditions. On each day during the PM season, the German Weather Agency (DWD) takes stock of the following conditions:[11]

- Condition #1: Whether the daily mean PM_{10} concentration at Neckartor monitoring station is over 30 μg/m3 and no rainfall is forecasted until 12am of the first forecast day. [12]

low ballpark estimates for the daily number of vehicles on Stuttgart roads, as our calculations do not include non-employed motorists (e.g. retirees, students, unemployed people, etc.), nor do estimates include other reasons for driving into the city (e.g. business travel, delivery, construction, etc.)

[10]See *Befragung zum Thema Feinstaubalarm in Stuttgart und Umgebung* (Omnitrend, 2016b) and *Befragung zum Thema Feinstaubalarm in Stuttgart und Umgebung im Zeitraum 26.2.2016 bis 28.2.2016* (Omnitrend, 2016a)

[11]See *Schadstoffrelevante Kriterien des Deutschen Wetterdienstes* (DWD, 2020)

[12]Snowfall and sleet are treated as rainless.

- Condition #2: Whether rain is forecasted for both the bridge day and the first forecast day.
- Condition #3: Whether wind blows with an average wind speed over 3 km per hour from 180°-330°.
- Condition #4: Whether there is a nighttime ground inversion.[13]
- Condition #5: Whether there is a low daytime mixing layer.[14]
- Condition #6: Whether average wind speed is below 3 km per hour.

According to the outcome of each binary condition and the corresponding decision rules (figure 2), DWD classifies the atmospheric interchange capacity as either "not limited," "limited" or "strongly limited" with only the latter leading to a PMA. As the primary condition, fulfillment of Condition #1 is sufficient for calling a PMA. If Condition #1 is not fulfilled, Conditions #2 and #3, and either Condition #4 or Condition #5, and at least four criteria overall must be fulfilled for the city to call a PMA.

If local authorities decide to call a PMA, in the early afternoon of the issue day they begin notifying the public of high air pollution levels and about a forthcoming DDA that goes into effect approximately 36 hours later (see figure 3, DDA day: -1). A bridge day (DDA day: 0), when the public continues to be informed about the PMA but the DDA has not gone into effect, follows the issue day. The DDA comes into effect after the bridge day at 0:00 am of the first forecast day (DDA day: 1). The DDA must continue for at least a second day (DDA day: 2) and remains in effect until the DWD forecasts two consecutive days where the atmospheric interchange capacity is not "strongly limited." Local authorities will announce the end of the PMA and DDA two days before messaging subsides.

Importantly, PMA and DDA designation is based on weather forecasts, not *actual* weather conditions on a given day. If authorities raise a PMA, unanticipated meteorological changes between issue day and any subsequent DDA day (first, second, third, etc.) may improve atmospheric interchange capacity to the extent that some PMA conditions may no longer be fulfilled on that DDA day. On these days, a DDA may have been broadcast although it need not have been. By similar logic, actual meteorological conditions may worsen the atmospheric interchange capacity to the extent that, on a

[13] Nighttime ground inversion is defined as an air layer within which temperature increases with altitude. Such an inversion traps particulate matter in the Stuttgart valley.

[14] The mixing layer height indicates the interchange capacity of the low lying air masses. The lower the mixing layer height, the smaller is the interchange capacity. The criterion is fulfilled if the mixing layer height is lower than 500 meters during the day.

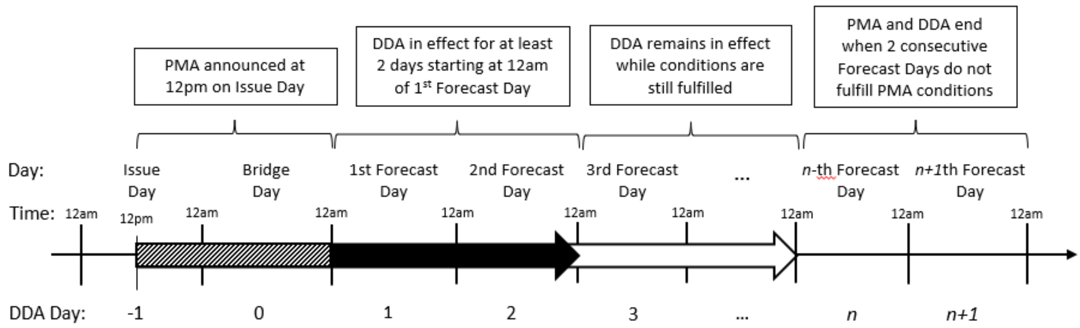

Figure 3: Particulate Matter Alert (PMA) and Don't Drive Appeal (DDA) timing. Information from the City of Stuttgart.

given non-DDA day, a DDA should have been broadcast, even though it was not. At the margin, local authorities may also exercise limited discretion in initiating a PMA event and broadcasting the DDA, specifically in cases when thresholds are just barely met (e.g. a small amount of rainfall may not be deemed sufficient to clear particulates from the air).

3 Theoretical considerations

Stuttgart's policy-makers employ a DDA in the ostensible belief, publicly expressed, that a morally framed request directed at car owners, combined with a public transit subsidy, will reduce driving. To see whether this belief can be rationalized, we develop a plausible mental model that formalizes this thinking. This simple theoretical framework is informed by existing models of modal switching for the *Spare The Air* (STA) program in the San Francisco Bay Area (Cutter and Neidell, 2009; Sexton, 2012) and urban congestion management policies in London and Santiago, Chile (Basso and Silva, 2014). To adapt the framework for the case at hand, we explicitly downplay the individual health aspects at the heart of the Bay Area's STA program, which are not part of Stuttgart's DDA, and instead emphasize its moral appeal considerations.

The literature identifies injunctive and descriptive norms as the main pathways through which a moral appeal can change the behavioral calculus of which action to choose (Bicchieri, 2005). Injunctive norms define how an individual ought to act. They constitute abstract moral absolutes, that is behavioral benchmarks independent of other people's behavior. Descriptive norms, on the other hand, reflect how most other people act. They are observable behavioral patterns in the population. In both cases, the literature has argued, individuals receive emotional rewards or losses from themselves and others as a function of adherence to or deviation from the norm. The associated feelings of righteousness and approval and of shame and guilt enter the utility function and can thus affect decision-making (Battigalli and Dufwenberg, 2007; Zafar, 2011).

Policy-makers are unlikely to be unaware of the subtle distinction between injunctive and descriptive norms. Yet, their mental model of DDAs may well capture the idea of injunctive norms by postulating that a DDA makes people attach positive feelings to deciding not to drive.[15] Descriptive norms could be captured by attaching to driving a negative feeling whose strength depends on the effectiveness of the appeal on others: Guilt and shame are strongest if the individual driver finds himself the only driver on the road, particularly if watched by non-drivers. They do not arise when traffic density during the DDA event is the same (or even higher) than before (Zafar, 2011). Considerations of positive and negative feelings triggered by adhering and deviating from norms would provide policy-makers with a behaviorally informed model of how car owners respond to the introduction of a DDA. They can also be extended to the question of how effective a DDA is likely to be over time. Policy-makers' intuition that the impact of DDAs wears off over a multi-day DDA event and needs time to recover between DDA events accords with well-established findings in psychology. Experimental tests of the theory of "ego depletion of self control" (Baumeister et al., 2000) consistently show that the emotional costs of not complying with norms that require a change from previous behavior decrease over time (Dang, 2018) and require a 'recovery period' between norm activation events (Tice et al., 2007). Considerations of both a static and dynamic nature are therefore likely to populate policy-makers' mental models of how a DDA affects driving.

To give some analytical heft to policy-makers' reasoning, we assume in line with the static congestion model of Basso and Silva (2014) that at any given point in time t, each individual i with access to a car and wishing to travel decides between driving (D) and not driving (ND) to reach their destination.[16] Driving is associated with utility (time arguments suppressed)

$$U_i^D = V_i^D - \tau_i t^D(1 + Q^D) - p^D - \mathbb{1}_A E_i \max\left\{(\overline{Q}^D - Q^D); 0\right\} \tag{1}$$

[15]Equivalently, it could be introduced as a negative feeling attached to driving. Analytically, it leads to the same results.

[16]These model formulations purposefully neglect the extensive margin of deciding not to travel.

while not driving is associated with utility

$$U_i^{ND} = V_i^{ND} - \tau_i t^{ND} - p^{ND}(1 - \mathbb{1}_A \delta) - \mathbb{1}_A G \tag{2}$$

with $\mathbb{1}_A$ an indicator variable that is one if an appeal has been issued and zero otherwise.

Expressions (1) and (2) capture that in the absence of a DDA ($\mathbb{1}_A = 0$), the respective utilities are a function of the intrinsic value that individual i associates with driving D and not driving ND, V_i^D and V_i^{ND}, the expenses of driving and not driving at market prices, p^D and p^{ND}, and the mode-independent[17] opportunity cost of time τ_i multiplied by the mode-specific travel time, t^D and t^{ND}. As in other models, total driving time is approximated as linear in car traffic density, measured by the aggregate demand for driving Q^D, along the entire itinerary, $t^D(1 + Q^D)$.[18] The driving-related air quality impacts that play a central role in the health-messaging models by Cutter and Neidell (2009) and Sexton (2012) are neglected in our representation of the policy-makers' mental model of moral appeals.

When a DDA is issued ($\mathbb{1}_A = 1$), three additional factors in expressions (1) and (2) are activated. First, in (2), the policy-maker reduces the cost of public transit through a discount δ, reducing non-driving expenses to $p^{ND}(1 - \delta)$. Second, also in (2), the policy-maker conveys through the appeal an injunctive norm that foregoing the use of car is the 'right thing to do'. The affective benefits of not driving are captured by a warm glow parameter G associated with norm compliance. Third, in (1), the DDA conveys a descriptive norm about driving: The greater the reduction in traffic densities during the DDA event relative to before, the greater the emotional cost to someone still driving. To approximate this effect, a simple linear formulation captures the emotional costs associated with violating the descriptive norm by driving as $E_i \max \left\{ \overline{Q}^D - Q^D; 0 \right\}$, with $\overline{Q}^D$ denoting aggregate demand for driving outside a DDA event. For traffic densities Q^D at or above pre-DDA levels, the emotional cost of driving is zero; for densities below, it is $E_i(\overline{Q}^D - Q^D)$. In line with the "ego depletion" mechanism, E_i is highest on the first day of a multi-day DDA event ($\overline{E_i}$) and declines to zero over time.[19]

As in Basso and Silva (2014), equilibrium traffic is the aggregate outcome of indi-

[17] Empirical evidence points to mode dependence: Time spent in one's own car has a lower opportunity cost than time spent in public transit. We abstract from this detail here.

[18] Total travel time is t^D when no other car is on the road ($Q^D = 0$) and increases in proportion to use by drivers. The linear approximation overestimates the effect of density on travel time for low levels of density and vice versa for high levels. This will lead to a slight overestimation of the effect of a DDA close to road capacity.

[19] We suppress the time argument in this sketch for notational simplicity.

viduals deciding to drive if $U_i^D - U_i^{ND} > 0$. Across individuals, this leads to aggregate demand for driving of

$$Q^D = \sum_i \mathbb{1}_i^D, \tag{3}$$

with $\mathbb{1}_i^D$ and indicator variable that is one if for individual i, $U_i^D - U_i^{ND} > 0$.

As a result of the congestibility of the road network, there is a demand equilibrium outside DDA events with a simple closed-form solution under the assumption of identical agents of the type

$$\overline{Q}^D = \frac{1}{\tau t^D} \{\Delta V - \Delta p - \tau \Delta t\} \tag{4}$$

with $\Delta V = V^D - V^{ND}$ denoting the difference in intrinsic values, $\Delta p = p^D - p^{ND}$ the difference in expenses, and $\Delta t = t^D - t^{ND}$ the difference in travel time between driving and not driving. Equilibrium traffic density increases in the intrinsic value differential and decreases in the price and travel time differential between driving and not driving. It is scaled down by the effective cost of time of driving τt^D on account of the congestion externality that every driver imposes on all other drivers in the road network.

A few steps of simple algebraic manipulation also yield the equilibrium traffic density during a DDA as

$$Q^D = \overline{Q}^D - \frac{G + p^{ND}\delta}{\tau t^D - E} \tag{5}$$

As intended by the policy-maker, equilibrium traffic density is always lower when a DDA is in effect.[20] The reduction increases in the warm glow of the appeal, G, and in the public transit discount, δ. Their effect size is scaled by the effective cost of driving time, τt^D, net of the emotional cost of driving when others do not, E.[21] It also follows from equation (5) that traffic density is lowest at the beginning of a DDA when $E = \overline{E}$ such that

$$\underline{Q}^D = \overline{Q}^D - \frac{G + p^{ND}\delta}{\tau t^D - \overline{E}} \tag{6}$$

and increases as the emotional costs of non-compliance fall with a continuing DDA:

$$-\frac{dQ^D}{dE} = \frac{G + p^{ND}\delta}{(\tau t^D - E)^2} > 0 \tag{7}$$

[20] This statement holds for positive traffic densities, which require that $\tau t^D - \overline{E} > 0$.

[21] Incidentally, the static congestion model also highlights the presence of an instrument for inducing a switch from driving that policy-makers did not consider: Increasing travel time t^D through speed restrictions.

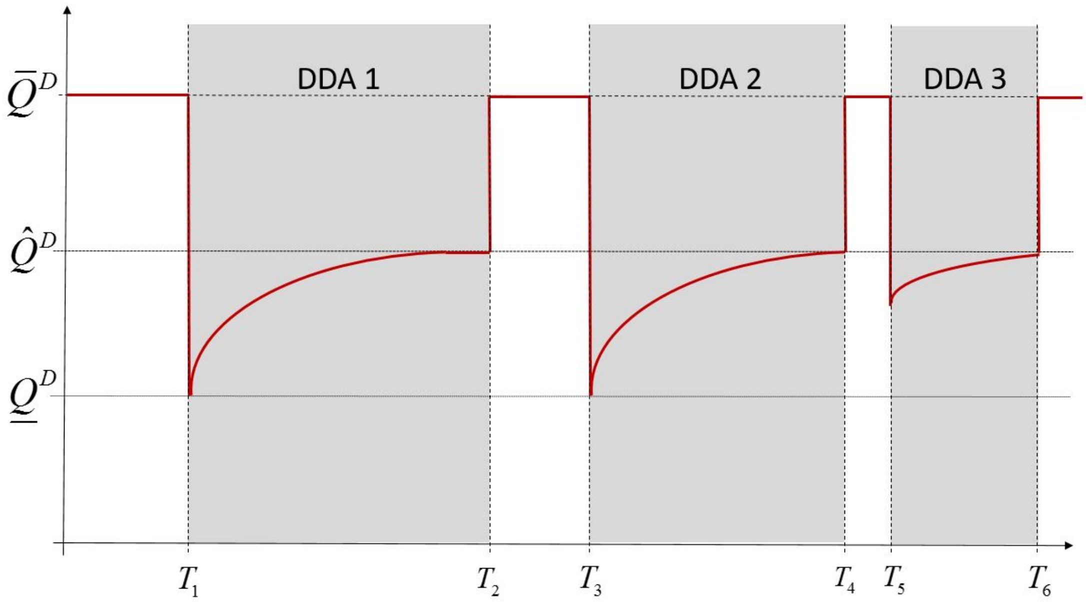

Figure 4: Illustrative evolution of traffic densities for a scenario with three DDAs.

A policy-maker reasoning along the lines sketched in expressions (1) to (5) can therefore conclude that issuing a DDA induces predictable temporal patterns: At the onset of a DDA, when the emotional costs of non-compliance are high ($\overline{E}$), the reduction in traffic density is greatest, leading to a minimal traffic density of $\underline{Q}^D$. As potential drivers progressively care less about non-complying, traffic density increases again and reaches a long-run equilibrium level $\hat{Q}^D > \underline{Q}^D$ given by $\lim_{E \to 0} Q^D = \overline{Q}^D - \frac{G+p^{ND}\delta}{\tau t^D} = \hat{Q}^D < \overline{Q}^D$.

As an illustration, figure 4 shows the evolution of traffic densities associated with a fictional scenario in which three DDAs are called between time T_1 and T_6: Baseline traffic density starting at time 0 is $\overline{Q}^D$. The first DDA, called at T_1, initially brings traffic levels down to $\underline{Q}^D$ as implied by expression 6. Over time, the emotional cost of non-compliance E wears off and traffic densities increase to $\hat{Q}$. When the DDA is called off at T_2, traffic returns to $\overline{Q}^D$ and the "ego" can recover in the time interval $[T_2; T_3]$. At time T_3, a second, shorter, DDA is called, followed by a shorter recovery interval $[T_4; T_5]$. As a result of incomplete recovery, the third DDA does not benefit from the same initial effect on traffic density as the two preceding episodes, falling short of traffic reduction $\underline{Q}^D$ at time T_5. From there, traffic again increases before the DDA is suspended at time T_6, just before traffic reaches the long-run equilibrium $\hat{Q}^D$ that prevails when "ego depletion" reduces the emotional cost of non-compliance to zero.

Together, equations (4) and (5) emphasize three aspects. One is that policy-makers can rationalize their belief in the effectiveness of DDAs: Invoking the norm-setting effects of DDAs in a behaviorally informed model provides a causal mechanism for affecting the choice whether to drive or not. The second is that the predicted equilibrium car traffic density under a DDA is strictly below non-DDA levels: The possibility that traffic might increase when a DDA is in force would require the policy-maker to consider a larger set of mechanisms. The third aspect is that the dynamic patterns of driving choices within and between multi-day DDA events make specific empirical predictions: Traffic reduction is expected to be greatest on the first days of a multi-day DDA event before tapering off to below-normal levels and is expected to be negatively affected when DDA events are spaced close together.

While the framework is good at capturing the moral appeal considerations of policy-makers, it probably does injustice to their understanding of the complexity of driving decisions. For example, it neglects issues of expectations and learning that are likely to be particularly important during early phases of the DDA program as car owners closely observe traffic densities. It also neglects problems of intertemporal substitutability of car-based activities (Basso and Silva, 2014), of health-related aspects of driving decisions (Cutter and Neidell, 2009; Sexton, 2012), and of the congestibility of public transit (Basso and Silva, 2014). These complexities can be expected to impact on the success of DDAs – and to be part of the ex-ante assessment undertaken by policy-makers in a more or less systematic fashion.

4 Data

4.1 Traffic Data

We obtain hourly vehicle traffic counts for the five PM seasons from January 2016 to April 2020 for 60 automatic traffic counters (ATCs) operated by the City of Stuttgart's Intregrated Traffic Control Center (*Integrierte Verkehrsleitzentral*, IVLZ) and from January 2016 to December 2019 for twenty ATCs from the Federal Highway Research Institute (*Bundesanstalt für Strassenwesen*, BaSt). Daily, counter-level traffic flows are recorded as the sum of twenty-four hourly counts if data are available for all 24 hours of a day, otherwise they are recorded as missing. We exclude all observations from 2020 due to the unprecedented effect of COVID-19 lockdowns on mobility and restrict our sample to counters that have at least 75% of daily observations during the PM seasons from January 2016 through December 2019 (n=43). Of 31,519 possible counter-day observations spanning 43 counters and 733 particulate matter season days, we observe 27,290 vehicles

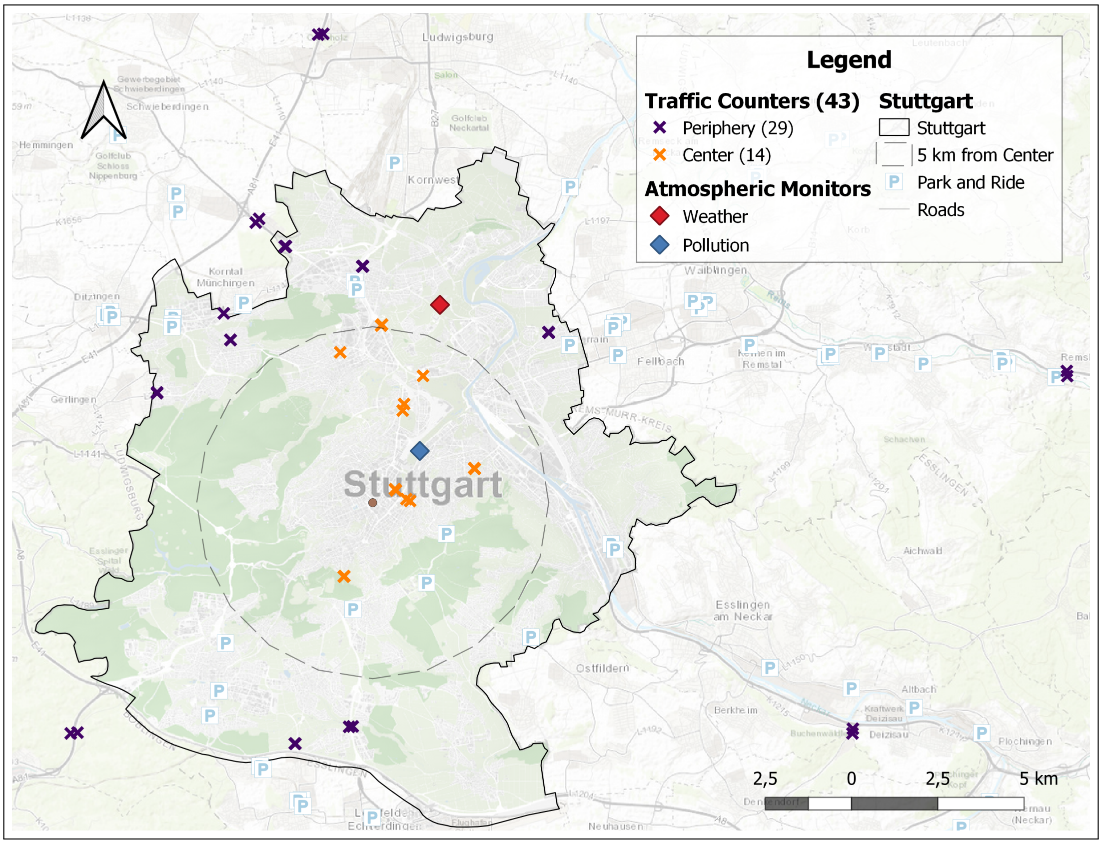

Figure 5: Map of Stuttgart with traffic counter locations by type (center vs. periphery) and weather and pollution monitoring sites.

Table 1: Summary Statistics: Vehicles per Day

Counter Set	Counters	Obs.	Mean	SD	Min	Max
All	43	27,290	23,340.60	18,588.50	85	89,384
IVLZ	31	18,494	13,714.88	10,076.05	85	52,697
BaSt	12	8,796	43,579.11	15,942.36	10,256	89,384
Periphery	29	18,503	25,856.94	20,781.67	662	89,384
Center	14	8,787	18,041.86	11,060.13	85	52,697

per counter-day observations (86.6%).

On average, 23,341 vehicles pass each counter each day, with traffic increasing moderately (+6%) over the course of the work week before dropping off on Saturdays (-14%) and more considerably on Sundays (-30%) relative to Mondays. Public and school holidays also have considerably lower traffic levels (-19% and -38%, respectively) compared to non-holidays. Aggregate traffic flows are also subject to temporary shocks (e.g. accidents, congestion, and construction sites), seasonal trends, and long-term shifts in road usage (e.g. vehicle bans, road closures, new road infrastructure, transit alternatives, macroeconomic shocks). A visual comparison of traffic count box plots on DDA days vs. non-DDA days (excluding holidays) in figure 6 suggests that traffic levels are similar across DDA status on all days of the week.

We categorize our set of traffic counters into those within 5km of Stuttgart's administrative centroid (n=14) and those at the city's periphery beyond 5km from the centroid (n=29) and map these locations in figure 5. This 5km radius proxies for the city center and captures its topographic setting at the middle of a basin. It also reflects the presence of park and ride infrastructure at the periphery, where parking opportunities are located for car commuters wishing to take public transit to reach the city center. Table 1 shows that average periphery traffic flows are considerably higher (25,857 vehicles per counter-day) than city center traffic flows (18,042 vehicles per counter-day).

Our traffic data limits the scope of our analysis in three ways. First, we analyze aggregate traffic counts and cannot observe the intensive and extensive margins of driving. That is, we cannot decipher between a relatively small set of automobiles on the road being driven more intensively (i.e. high daily vehicle kilometers traveled per car) and a proportionally larger set of automobiles being driven relatively less intensively (i.e. fewer daily vehicle kilometers traveled per car). Second, we are not able to observe individual-level modal switching.[22] Third, our data set consists of traffic flows on a sub-

[22]We have inquired at the city and its public transportation partners about alternative transit data. The city nor its public transportation partners maintain turnstiles at public transit stations that would deliver daily measures of public transit use. Available monthly ticket sales do not have the

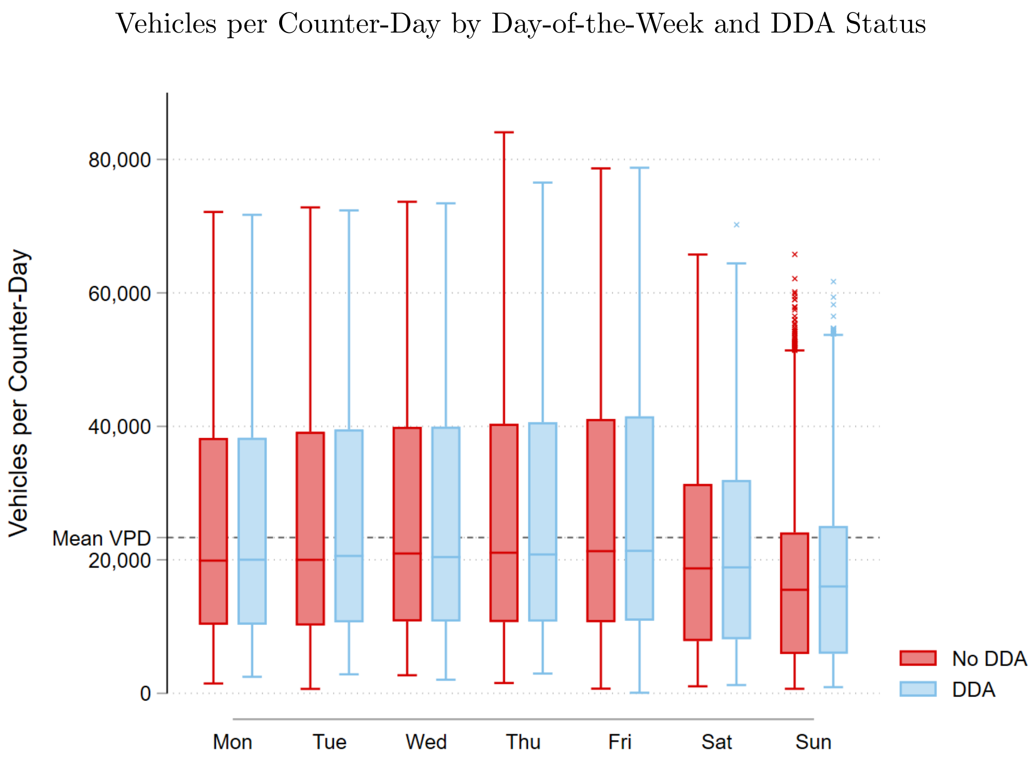

Figure 6: Box plot of vehicles per counter-day. Median, inner quartile range, lower and upper bounds, and outliers (beyond 1.5 × inner quartile range) are depicted by day of the week and DDA status. Holidays are excluded. Mean vehicles per counter-day (VPD) equals 23,341.

set of streets in Stuttgart. The 43 traffic counters we use in our analysis are distributed across 22 sites, which we believe are representative of overall city conditions as they are dispersed across different road types and neighborhoods.

4.2 Weather, Pollution, and DDA Data

We follow the AQA literature to control for daily weather factors which may influence driving such as temperature, precipitation by type, wind speed, and sunshine hours. We retrieve weather data for the Schnarrenberg weather station (See location in figure 5) from DWD Open Data and assume that weather conditions there are the best available measure of meteorological factors that influence motorists. Air pollution data come from the Baden-Württemberg State Institute for the Environment, Survey and Nature

temporal or spatial resolution necessary for our analysis. Stuttgart also collects cycling data at two automatic bicycle counters over the time period of interest, but this data is not rich enough for our analysis.

Table 2: Summary Statistics: Weather and Pollution Variables by DDA Status

	Non-DDA Days					DDA Days				
Variable	Obs.	Mean	SD	Min	Max	Obs.	Mean	SD	Min	Max
DDA	483	0	0	0	0	250	1	0	1	1
Mean Temperature ($^{\circ}C$)	483	6.20	3.99	-7	16.5	250	4.38	5.61	-9	16.2
Rainfall (mm)	483	1.23	2.80	0	20.3	250	0.06	0.28	0	2.5
Snowfall (mm)	483	0.34	1.52	0	17.3	250	0.08	0.42	0	4.1
Sleet (mm)	483	0.34	1.45	0	18.3	250	0.05	0.30	0	2.4
Relative Humidity (%)	483	77.45	9.85	38.38	98.54	250	73.92	12.52	26.92	98
Sunshine Hours	483	2.34	2.81	0	12.31	250	5.13	3.90	0	12.41
Mean Windspeed (km/h)	483	3.37	1.24	0.8	8.3	250	2.57	0.87	0.9	5.9
Daily Mean PM_{10} ($\mu g/m^3$)	466	27.72	17.84	4	202	243	54.22	24.34	17	176

Conservation (*Landesanstalt für Umwelt Baden-Württemberg*, LUBW), which monitors PM_{10} concentrations in the city center (See location in figure 5). We manually input DDA status from a Stuttgart website as a binary variable that equals one on days when a DDA is called and zero otherwise (figure 7).

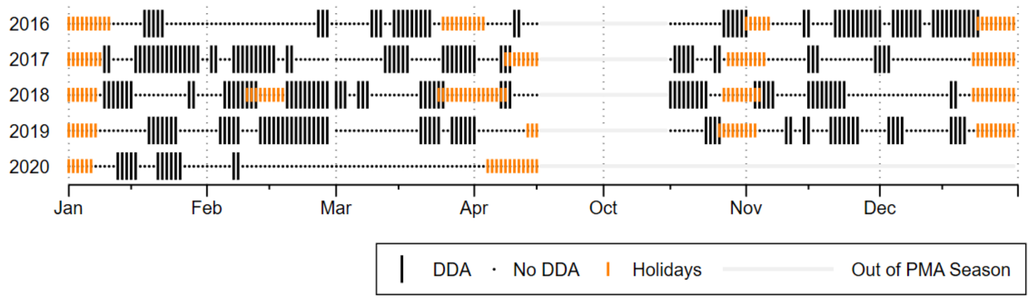

Figure 7: DDA days from January 2016 to April 2020.

In comparison to other DDAs and AQAs studied in the literature, Stuttgart's DDA is implemented very frequently and for long durations.[23] Over 733 possible PMA days from January 2016 through December 2019, Stuttgart authorities broadcast a DDA on 250 days (34%) in 44 multi-day DDA events with the average DDA extending 5.7 days. Due to Stuttgart's PMA design, DDA days are, on average, colder, less windy, sunnier, and more polluted than non-DDA days (table 2). They also experience less precipitation (i.e. rain, snow, sleet) and fewer heavy precipitation events. DDA days are typically preceded by days with similar weather and pollution levels, while the same holds for

[23]For example, in Cutter and Neidell (2009) about 4.5% of days in San Francisco are treated with an *Spare the Air* alert, in Saberian et al. (2017) about 1.3% of days in Sydney experience an ozone alert day, and in Tribby et al. (2013) about 16% of PM season days have either a yellow or red AQA.

Table 3: DDA Days by Day of the Week

DDA Day #	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Total
1	12	7	8	5	2	4	6	44
2	6	12	7	8	5	2	4	44
3	4	2	9	7	7	3	2	34
4	0	3	2	9	6	5	3	28
5	2	0	3	2	9	4	4	24
6	3	2	0	2	2	6	4	19
7	3	2	1	0	2	1	6	15
8	6	2	1	1	0	1	0	11
9	0	5	2	1	1	0	0	9
10	0	0	3	2	1	0	0	6
11	0	0	0	3	1	1	0	5
12	0	0	0	0	2	0	1	3
13	1	0	0	0	0	2	0	3
14	0	1	0	0	0	0	1	2
15	1	0	1	0	0	0	0	2
16	0	0	0	1	0	0	0	1
Total	38	36	37	41	38	29	31	250

non-DDA days. Authorities are also less likely call DDAs on public and school holidays, possibly because they expect lower traffic levels on these days. As figure 7 depicts, only few DDAs fall on public or school holidays (6%, 14 of 250 DDA days) compared to non-holidays (94%, 236 of 250 DDA days). For this reason, we believe that holidays may systematically differ from non-holidays, so we remove public and school holidays from parts of our analysis. Authorities also often announce PMAs on weekends and at the beginning of the week, leading to a large share of DDAs starting on Mondays, Tuesdays, and Wednesdays (table 3). Overall, there is a fairly uniform distribution of DDA days across the working week with weekends being treated with DDAs less often than weekdays.

5 Empirical Framework

5.1 Estimation Strategy

To recover the effect of calling a DDA on traffic, we employ a dynamic panel estimation model and restrict our sample around the DDA trigger in a regression-discontinuity-like approach. Our strategy tunes estimation techniques from adjacent studies (Cutter and Neidell, 2009; Sexton, 2012; Noonan, 2014) to Stuttgart's DDA design. Due to the DDA design, the data-generating process in Stuttgart depends strongly on previous periods. This motivates the inclusion of lagged controls and careful consideration of treatment counterfactuals. We begin by introducing the estimation equation underlying our main regression model before discussing our identification strategy and its accompanying subsampling scheme in the following subsection.

We estimate the impact of Stuttgart's DDA on traffic levels using an ordinary least squares (OLS) regression model described by the following equation:

$$y_{i,t} = \beta_1 DDA_t + \delta_1 y_{i,t-1} + \delta_2 M_t + \delta_3 M_{t-1} + \delta_4 M_{t-2} + \delta_5 M_{t-3} + \gamma_i + \phi_t + \epsilon_{i,t}, \quad (8)$$

where $y_{i,t}$ is the number of vehicles passing counter i on date t, and β_1 estimates the overall DDA effect as the average difference in daily traffic counts between DDA-days and non-DDA days across all counters. The variable of interest, DDA_t, takes on a value of one on DDA-days and zero otherwise. This model uses a lagged dependent variable ($y_{i,t-1}$) to adjust for previous day traffic shocks (e.g. traffic re-routing, construction), contemporaneous and lagged environmental controls ($M_t, M_{t-1}, M_{t-2}, M_{t-3}$) to account for multi-day weather patterns,[24] day-of-the-week and holiday dummies to address intra-weekly traffic trends and traffic shifts during vacation periods, counter-level fixed effects (γ_i) to account for counter-specific traffic levels, and year-month time fixed effects (ϕ_t) to capture seasonal trends in car use and policy discontinuities that might influence overall traffic levels (e.g. varying public transit prices, vehicle bans, new infrastructure, etc.).

Our estimation strategy tests the null hypothesis that the DDA effect is equal to zero ($H_0 : \beta_1 = 0$), or, in other words, that traffic flows do not differ significantly on days when a DDA is broadcast. If, as intended, private car use is lower on DDA days, the DDA effect coefficient must be negative ($\beta_1 < 0$) and differ significantly from zero. In the regression model defined by equation (8), we employ Huber-White standard errors to address heteroscedastic residuals, and we account for serial correlation and spatial auto-correlation by clustering standard errors on traffic counter site.[25] All regressions are also carried out with a logged dependent variable.[26]

We are also interested in evaluating whether DDA effectiveness varies over time and space.[27] To inspect for temporal heterogeneity, we successively augment equation (8)

[24] We follow the literature on air quality alerts and transportation choice in including precipitation, temperature, sunshine, and humidity as control variables. In addition to absolute precipitation by type, we also include squared terms for rainfall (mm^2), snowfall (mm^2), and sleet (mm^2).

[25] Note there are typically two counters at each site with one corresponding to each traffic direction.

[26] Log-scaling the outcome variable approximates differences in the outcome variable as percentage changes.

[27] Previous research on AQAs has highlighted spatial and temporal heterogeneity. For example, Tribby et al. (2013) find evidence of spatial displacement effects where traffic increases at Salt Lake City's periphery. Saberian et al. (2017) and Graff Zivin and Neidell (2009) find evidence of alert fatigue on the second day of ozone alerts.

with day of the week ($DOW_t \times DDA_t$) and year ($YEAR_t \times DDA_t$) interaction terms to evaluate how DDA effectiveness differs within the week and year-by-year, respectively. We also test for treatment heterogeneity by DDA event day (e.g. first day effect vs. second day effect, etc.) by adding a DDA day interaction term ($DDADAY_t \times DDA_t$) to equation (8). To test for spatial heterogeneity, we run our regression models separately for groups of counters at the city's periphery and center, and we fully disaggregate our panel and estimate individual counter-level DDA effects using time-series models synonymous with equation (8).[28]

5.2 Identification Strategy

In our setting, local authorities determine DDA treatment status based on six observed or forecasted atmospheric conditions, and they only lift this assignment once two consecutive days do not fulfill these conditions (figure 2, section 2.2). This multidimensional treatment protocol poses two main challenges for successfully identifying the DDA's effect on car-trip demand.

First, a given day's DDA treatment status is not determined by a single contemporaneous atmospheric parameter (e.g. a PM_{10} threshold value) but is rather a multivariate function of previously-realized atmospheric observations and uncertain weather predictions. This complicates the use of a canonical regression discontinuity design, as implemented by Cutter and Neidell (2009) or Noonan (2014), because multiple atmospheric thresholds must be fulfilled simultaneously and multiple pathways to a DDA exist. There is no single cut-off point we could exploit as a policy discontinuity. Also, due to forecast uncertainty, actual weather conditions can deviate from those outlined in the treatment protocol, jeopardizing whether treated days and untreated days are subject to respectively similar atmospheric conditions. Finally, discretion available to local authorities when evaluating uncertain weather forecasts may also bias whether DDA events are, in fact, initiated or terminated when treatment conditions suggest they should have been. Inspecting treatment conditions and classifying DDA days by their similarity to control and treatment days could help alleviate these potential complications.

Second, we anticipate that DDA determinants (i.e. weather) directly influence transportation demand and may confound our DDA effect estimates. In particular, persistent atmospheric conditions, which are endogenous to the treatment protocol, may correlate with modal switching. For example, some motorists may begin riding public transit or

[28]Here counter-level fixed effects are omitted due to the lacking panel structure. These counter-level models do accept monthly counter-specific time fixed effects, which were otherwise averaged out in the full panel model.

cycling due to prolonged dry, sunny weather, while such conditions also increase the likelihood that a DDA is called. Viewing each day as an exclusively independent observation would ignore strong previous day atmospheric and behavioral dependencies (see section 3), so a meaningfully selected control group of untreated multi-day events would better capture Stuttgart's DDA design features.

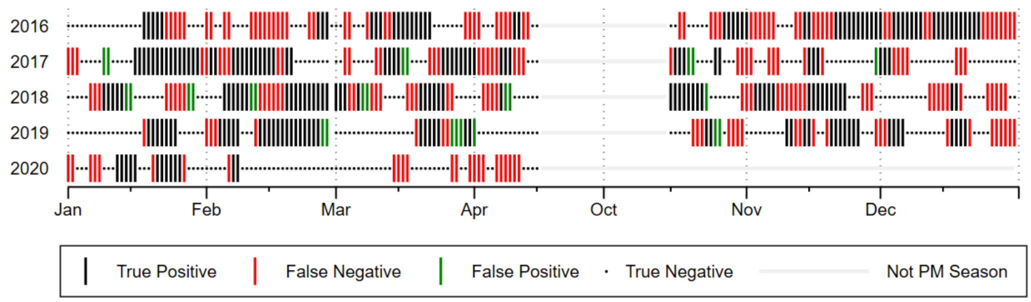

Figure 8: Reclassified DDA days from January 2016 to April 2020.

To address these two concerns, we limit our sample to comparisons of true positive DDA days, when the multi-day DDA conditions were observed and a DDA was broadcast, with false negative "counterfactual" DDA events, when DDA conditions were observed but no DDA was called. We use reported weather and pollution data to reconstruct the DDA conditions, slightly loosen the conditions around the DDA trigger, and thereby identify a set of multi-day non-DDA periods with atmospheric conditions most similar to actual DDA days (figure 8).[29] This approach allows us to "zoom in" on sets of days on either side of the DDA trigger in a regression-discontinuity-like manner and compare sets of days that were treated with a DDA with ones that were not. Using actual weather data rather than weather forecast data abstracts from one empirical aspect, namely that some motorists may switch modes based on multi-day weather forecasts, which are not captured in our observational data. However, we assume local authorities and motorists have similarly accurate weather forecasts at their disposal so that motorists cannot confidently predict DDA policy errors.

A further challenge to identification may arise due to reverse causality between the outcome of interest, car trip demand, and DDA treatment status. Changes in car trip demand could conceivably cause PM_{10} levels to rise above or fall below the threshold of 30 $\mu g/m^3$. However, car trip demand has no influence over the necessary second sub-condition of Condition #1, namely whether rainfall is anticipated or not, nor over the remaining five atmospheric conditions. Consequently, we see it as improbable that

[29] We explain this reclassification scheme in detail in appendix A

car trip demand could cause treatment status to change as treatment status is largely determined exogenously. It is also possible that local authorities' expectations about car use affect their decision to call a DDA. For example, local authorities may be less likely to broadcast a DDA if they anticipate that traffic levels will already be low due to school or public holidays, biasing DDA effect estimates upward. While we cannot observe policy-maker's traffic expectations, we account for this by removing days with traffic level outliers (e.g. holidays) in some specifications.

There are a number of unobservables that could plausibly affect our analysis. First, our data do not allow us to control for same-day traffic shocks (e.g. traffic jams, accidents, large events, etc.). Barring remarkable changes in traffic conditions on DDA days or considerable spatial displacement effects, we think it is unlikely that same-day traffic shocks would differ systematically on DDA days compared to non-DDA days or significantly bias our DDA effect estimates. Furthermore, we expect temporary traffic displacement to average out across nearby counters. Second, we are unable to observe individual motorists' expectations or their PMA and DDA information exposure (e.g. consumption of PMA-adjacent programming, etc.) over time. If such aspects are salient and do influence driving choices during DDAs, we anticipate that they will aggregate systematically in the population and result in detectable differences in DDA effectiveness over time. Finally, the announcement of an upcoming DDA may change motorists' choices until the DDA actually takes effect (i.e. on the issue or bridge day). We cannot observe whether individual motorists take additional trips on issue and bridge days to avoid taking trips during the DDA, but such a scenario would bias our DDA effect estimates downward. We account for these anticipatory effects by removing issue and bridge days from our sample in some specifications.

6 Results

6.1 Overall DDA Effect

Our regression results show that the overall daily DDA effect is positive and of small to negligible magnitude across a variety of different samples and specifications. On average, the number of vehicles passing each counter increases by 0.1% to 1.9% across all counters on DDA days compared to non-DDA days. Our main specification, a dynamic panel model that includes single-day traffic lags, contemporaneous and lagged weather controls, counter-level fixed effects, and year-month fixed effects estimates that traffic increases by 1.02%, or 239 additional vehicles per counter-day, on DDA days in the full sample (table 4, Column 1). This estimate is significant at the 5% significance level. Back-of-the-envelope calculations equate this DDA effect with a net increase of about 3,896 motorists in Greater Stuttgart on DDA days.[30]

Table 4: OLS Regression Results: Overall Daily DDA Effect

	(1) VPD	(2) VPD	(3) VPD	(4) VPD	(5) VPD	(6) VPD
DDA	238.6**	436.1***	120.4**	64.55	105.1*	27.55
	[+1.02%]	[+1.87%]	[+0.51%]	[+0.27%]	[+0.45%]	[+0.12%]
	(69.82)	(111.4)	(38.57)	(57.95)	(44.82)	(82.37)
Full Sample:	Y	N	N	N	N	N
TP & FN Sample:	N	Y	N	Y	N	Y
Holidays Excluded:	N	N	Y	Y	Y	Y
Bridge & Issue Days Excluded:	N	N	N	N	Y	Y
Observations	26,626	11,996	20,899	10,040	16,787	7,641
Counters	43	43	43	43	43	43
Days	733	381	584	320	509	272
DDA Days	250	219	236	212	236	212
Non-DDA Days	483	162	348	108	273	60
Mean VPD	23,341	24,046	24,238	24,586	24,361	24,764

Dependent variable is vehicles per counter-day (VPD). Robust standard errors clustered on 22 counter sites in parentheses. All models include single-day lagged traffic, a full set of weather controls, first, second, and third-day lagged weather controls, counter fixed effects, year-month fixed effects, and day-of-the-week and holiday dummies. Percent change relative to mean VPD in brackets.
*: Significant at 10%, **: Significant at 5%, ***: Significant at 1%.

The magnitude of our overall DDA effect estimate increases to 1.87%, or 436 additional vehicles per counter-day, when we restrict our sample to days that our DDA reclassification scheme categorized as true positive or false negative (table 4, Column 2). To make sure our model compares *similar* DDA and control days, we further restrict our sample by excluding holidays (table 4, Columns 3 and 4) and removing bridge and issue days (table 4, Columns 5 and 6). With these restricted samples, our DDA effect estimates decrease in magnitude and statistical significance but do not switch signs.

This result is robust to a number of alternative specifications. In appendix B, figure 14 shows the evolution of our overall DDA effect estimate across 90 specifications where we successively build up equation (8) term-by-term. Each panel of figure 14 corresponds to a different sample (see caption, same samples as columns in table 4) and within each panel, from left to right, we separately estimate the DDA effect for each specification, progressively adding year-month fixed effects, site standard errors, mean temperature, rainfall, snow, sleet, humidity, sunshine, mean wind speed, lagged VPD,

[30]This assumes 382,000 motorists on an average work day.

and first, second, and third day control variable lags. The last specification in each panel corresponds to the results shown in table 4. We also estimate our main specifications with logged traffic counts (see table 5 in appendix B). Across the board, the magnitude of these DDA effect estimates is below 2% of mean daily counter-level traffic. None of the models estimates a statistically significant negative DDA effect in line with the DDA's goal of reducing traffic.

6.2 Spatial and Temporal Heterogeneity

Our results in the previous section estimate the overall daily effect of a DDA on traffic levels. As spatially and temporally heterogenous effects are plausible, we disaggregate our model across space and time and estimate separately the daily DDA effect for different locations and time periods.

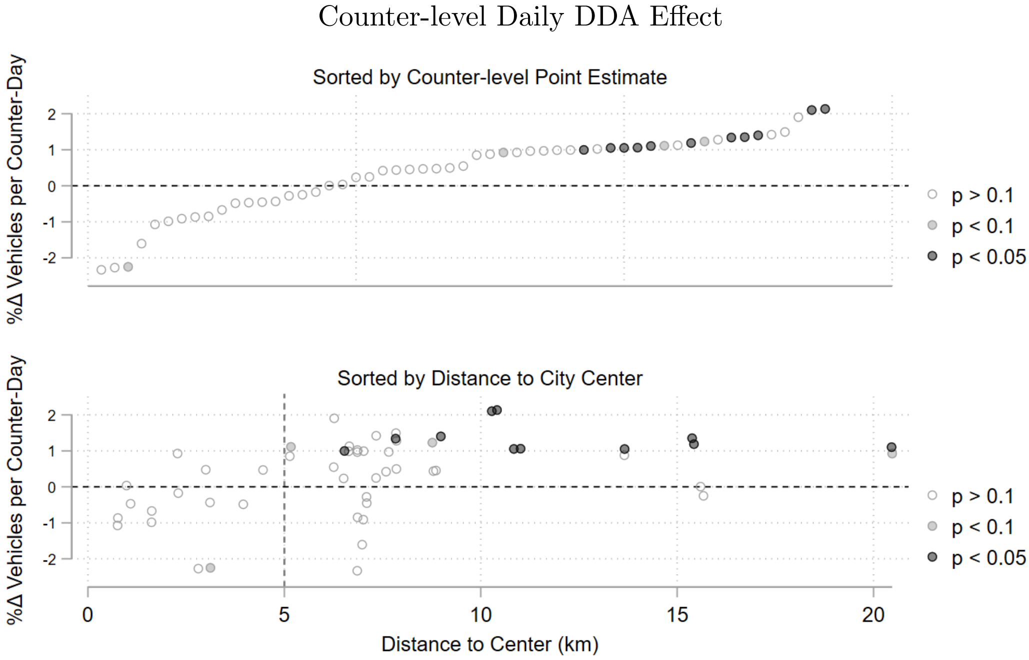

Figure 9: Counter-level daily Don't Drive Appeal (DDA) effect (percent change in vehicles per counter-day) sorted by counter-level point estimate and distance to city center. Percent change relative to mean counter VPD. Note that counters located within approximately 5km of city center are classified as city center counters.

Our fully spatially-disaggregated model (figure 9) shows higher traffic at the majority of counters on DDA days. However, most counter-level estimates are not statistically significant at the 10% significance level. One counter at the city center witnesses a statistically significant traffic decrease (-2%) on DDA days. Traffic does increase significantly about 1-2% compared to non-DDA days at some periphery counters and in particular at counters located furthest from the city. In the city center, most daily DDA effect point estimates range from -1% to +1% although these are with one exception not statistically significant at 10% level.

Daily DDA Effect by Day-of-the-Week, Location, and Sample Subset

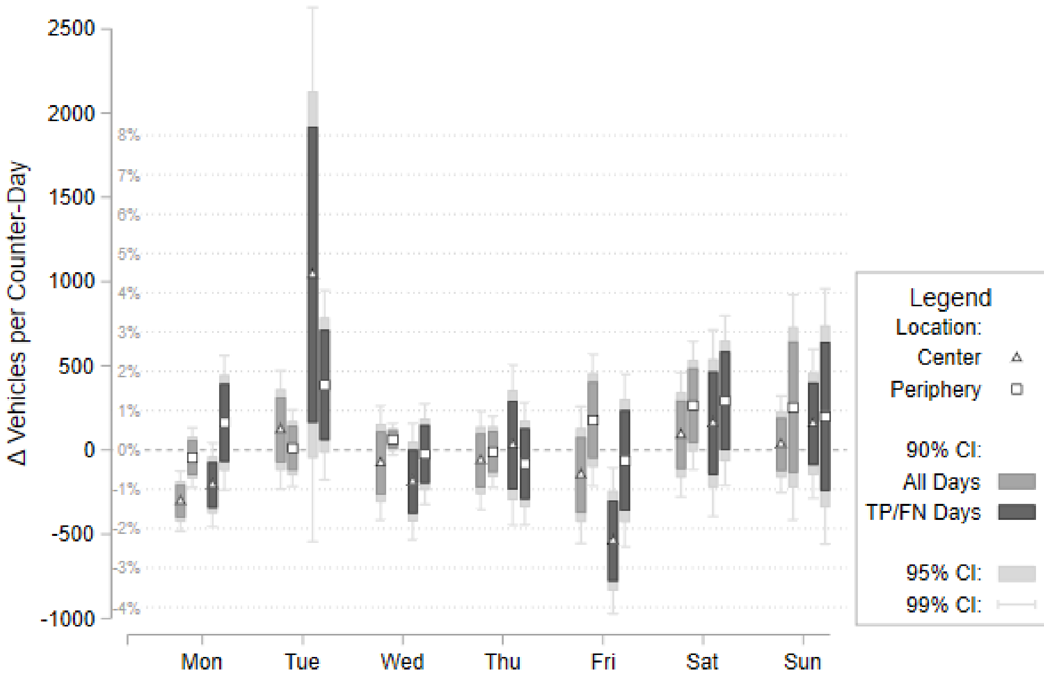

Figure 10: Daily Don't Drive Appeal (DDA) effect point estimates by day of the week, counter location, and sample subset. Percentages relative to average vehicles per counter-day (23,341).

We group counters into those located in the city center and those at the city's periphery to highlight the spatial heterogeneity in DDA effectiveness and to explore temporal heterogeneity at this level. figure 10 displays DDA effect point estimates by day of the week, counter location, and sample subset. On Monday and Friday there is statistically significant evidence for traffic decreases at the city center. However, these results are only robust to both the full sample and our true positive and false-negative sub-sample on Monday. Nevertheless, with the exception of Tuesday, all city center DDA effect estimates are close to or below zero, suggesting that the positive overall DDA effect found in previous sections does not result from increased traffic at the city center.

Instead, we find evidence that periphery traffic increases may outweigh modest decreases at city center locations. In eleven of the fourteen cases depicted in figure 10, periphery DDA effect estimates are greater than their city center counterpart. Periphery effects are most evident on weekends, and many of the periphery point estimates are positive or close to zero. However, periphery estimates are not statistically significant at the 5% level for any day of the week or either subsample. In the vast majority of cases, periphery point estimates do not exceed 1% of mean daily counter-level traffic and never exceed 2% of mean daily counter-level traffic.

Daily DDA Effect by Year, Location, and Sample Subset

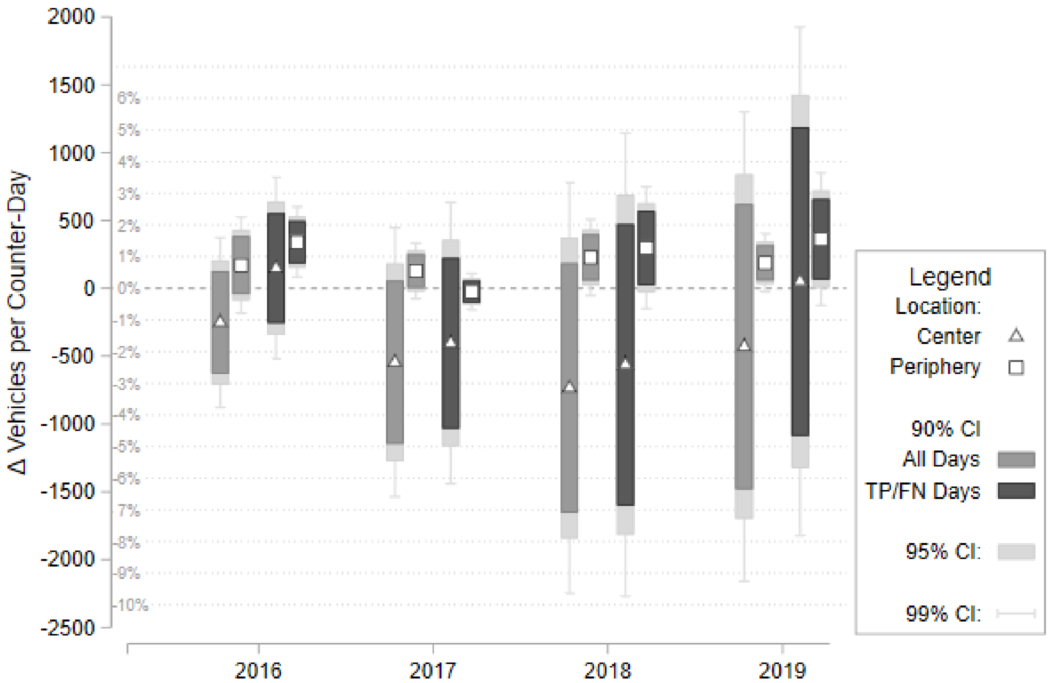

Figure 11: Daily Don't Drive Appeal (DDA) effect point estimates by year, counter location, and sample subset. Percentages relative to average vehicles per counter-day (23,341).

Figure 11 shows variations in DDA effectiveness over the PMA program's lifetime. The estimated DDA effect at the periphery is almost universally between 0% and +2% of mean daily traffic from 2016 to 2019, and these estimates are statistically significant in 2016, 2018, and 2019 for all but one sub-sample (full sample 2016). City center DDA effects are strongest (highest negative magnitude) in 2017 and 2018, with these two years seeing an average traffic reduction between 2% and approximately 3% at the city center on DDA days relative to average overall traffic levels. City center estimates are, however, not statistically significant.

6.3 Dynamic DDA Effects

Daily DDA Effect by Day-of-the-DDA and Sample Subset

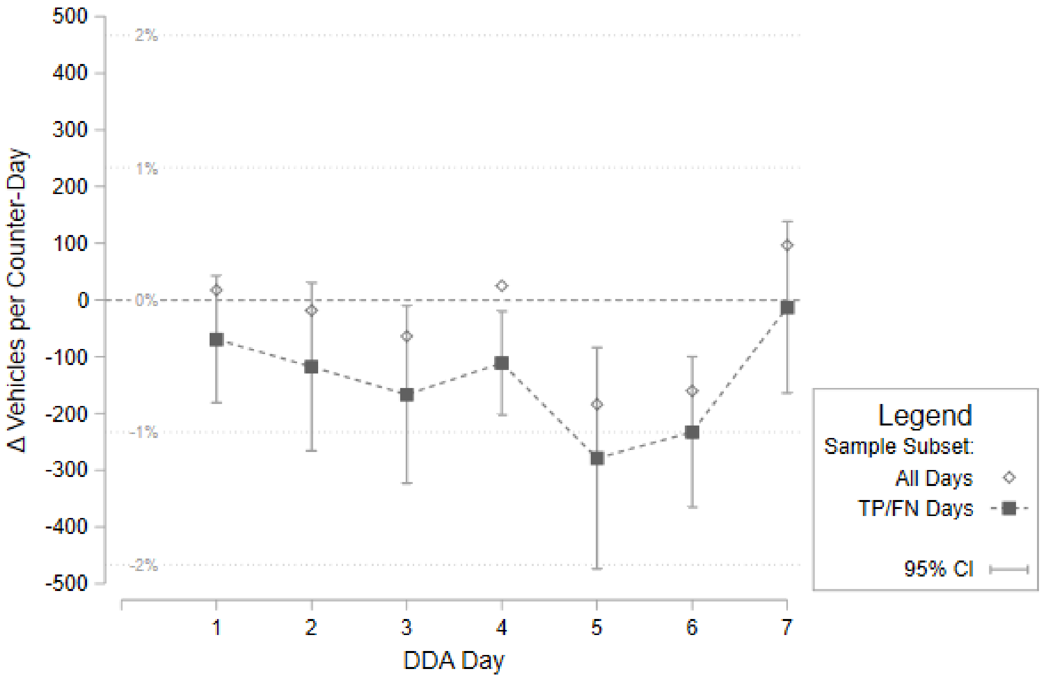

Figure 12: Don't Drive Appeal (DDA) effect point estimates over DDA duration by sample subset. Percentages relative to average vehicles per counter-day (23,341).

A plausible mental model of the effects of a DDA on driving decisions includes possible dynamic effects (see section 3). To explore these, we interact DDA day terms (e.g. first day, second day, etc.) with the DDA effect variable in equation (8), remove lagged traffic volumes, and then estimate the daily DDA effect over the DDA duration. Figure 12 displays DDA effect point estimates for each day of a DDA. In the true positive and false negative sub-sample, which compares the most similar DDA and non-DDA days in terms of atmospheric conditions, we find that the DDA increasingly reduces traffic over the first six days, before traffic rebounds to normal levels after the sixth day. Effect sizes are very modest at or below 1% of daily mean counter-level traffic and, for the full-sample, typically of smaller magnitude compared to the sub-sample point estimates. These results indicate that a small share of drivers may begin to shift away from driving as a DDA continues, in particular when DDAs extend for several days. Traffic recovers to average levels after the sixth day suggesting that drivers' willingness to persistently shift their transportation choice during a DDA is limited.

Daily DDA Effect by Day-of-the-DDA and Location

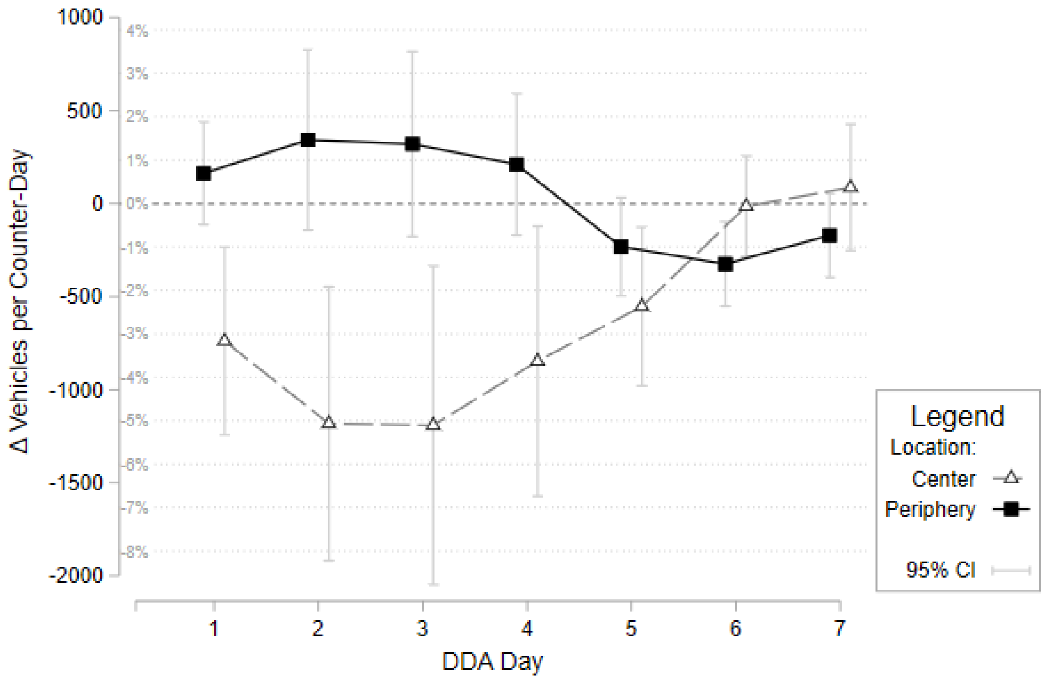

Figure 13: Don't Drive Appeal (DDA) Effect over DDA Duration. Models use true positive and false negative subset of days. Percentages relative to average vehicles per counter-day (23,341).

However, we find this dynamic DDA effect to evolve differently for locations at the city center compared to the periphery. In figure 13, we split the DDA effect by location (city center vs. periphery) and estimate considerable, statistically significant decreases in traffic flows at the city center for the first five days of a DDA while traffic at periphery counters does not differ significantly from non-DDA days over this time period. Over the first five DDA days, the magnitude of the DDA effect ranges between -2% to -5% of mean daily traffic for the city center and hovers between +2% and -1% of mean daily traffic at the periphery. On DDA days six and seven the city center DDA effect diminishes to near zero, while the size and statistical significance of periphery estimates suggests that after DDA day four there may be modest to negligible decreases at periphery locations (approximately -1% of mean daily traffic).

These dynamic patterns capture some of our theoretical hypotheses from section 3.

City center traffic does appear to exclusively experience traffic reductions or negligible effects, which likely aligns with policy-makers' expectations for the DDA. Further, the dynamic tapering of the DDA effect at the city center provides some suggestive evidence of self-control depletion or other norm-based dynamics. However, unlike in figure 4, we find evidence that the DDA reaches peak effectiveness in the city center on the second and third DDA day, whereas we hypothesized that effects would be strongest immediately after broadcasting a DDA.

7 Conclusion

Officials implement air quality alert programs to disseminate air pollution information, promote avoidance behavior in sensitive populations, and appeal for pollution reductions. With this paper, we contribute to a growing literature on air quality alert effectiveness by investigating an alert that encourages commuters not to drive cars on poor air quality days.

The results of our analysis provide new evidence about the effectiveness of combining air quality alerts with Don't Drive Appeals (DDAs) from a well-suited European metropolitan setting with widespread green political support and a dense public transit network. We find that the prediction that DDAs reduce driving on DDA days can be rationalized by appealing to a behaviorally informed model of car owners, but fails an empirical test: On average, the DDA increases traffic on DDA days by 0.1%-1.9%, contrary to the program's overall objective. We do find two important spatial and temporal nuances to this result. First, the DDA increases traffic primarily at the city's periphery and on weekends. Second, we find the DDA reduces city center traffic on certain weekdays (Mondays and Fridays), possibly to a greater extent during 2017 and 2018, and over the first five days after a DDA has been broadcast. Importantly, these city center effects are not universal and always modest (between 0% and -5% of mean daily traffic). Only in limited cases do they approach -5% mean daily traffic. Our overall DDA effect and periphery results echo the findings of Tribby et al. (2013), who find Salt Lake City, USA's particulate matter alert inadvertently increases traffic in the city by 3%-4%. Our city center findings are situated between Sexton (2012)'s no effect result and Cutter and Neidell (2009)'s finding of a 2%-3% traffic reduction on *Spare the Air* days in San Francisco, USA. Our analysis, in particular, highlights how spatial and temporal traffic displacement may be concealed in overall DDA effect estimators.

In this paper we provide two methodological contributions that may inform future research in this domain. First, we use atmospheric data to create a sub-sample of multi-day non-DDA events that are most similar to multi-day DDA events that fulfill DDA conditions and then compare regression results from this sample to the full sample throughout. This approach builds confidence in our choice of a “control” group of non-DDA days, in particular in the context of Stuttgart’s complex, multi-factor DDA design. Second, we derive insights about DDA effectiveness by conducting a spatially and temporally disaggregated analysis. The divergence of our disaggregated DDA effect findings from our overall findings shows that programs with geographic restrictions, temporal designs, and norms-based messaging like Stuttgart’s DDA may have important heterogeneities in effectiveness.

Our findings may also caution policymakers interested in combining air quality alerts with Don‘t Drive Appeals. Air quality alerts are generally considered ineffective policy for achieving driving reductions, and our study does not provide resounding evidence that these policies are persistently effective. Our study also establishes that, even if city center traffic does not inadvertently increase, alerts combined with DDAs may displace traffic to the periphery. It is not clear in Stuttgart’s scenario whether modest traffic decreases at the city center and modest traffic increases at the periphery effectively reduce air pollution exposure in the target population. However, urban policymakers might value traffic (and emissions) reductions at city centers, where population density is likely highest, more than moderate increases at the periphery. Our study is limited by its use of traffic count data. Future research could investigate individual-level responses to DDAs but would require individual-level commuting data and information about individual DDA information exposure. Such analyses might also be able to shed light on socioeconomic dimensions of DDA effectiveness and, with an eye to an equitable mobility transition, inform policymakers how different groups respond to norms-based messaging.

References

Basso, L. J. and Silva, H. E. (2014). Efficiency and substitutability of transit subsidies and other urban transport policies. *American Economic Journal: Economic Policy*, 6(4):1–33.

Battigalli, P. and Dufwenberg, M. (2007). Guilt in games. *American Economic Review*, 97(2):170–176.

Baumeister, R. F., Muraven, M., and Tice, D. M. (2000). Ego depletion: A resource model of volition, self-regulation, and controlled processing. *Social cognition*, 18(2):130–150.

Bicchieri, C. (2005). *The grammar of society: The nature and dynamics of social norms*. Cambridge University Press.

Cummings, R. G. and Walker, M. B. (2000). Measuring the effectiveness of voluntary emission reduction programmes. *Applied Economics*, 32(13):1719–1726.

Cutter, W. B. and Neidell, M. (2009). Voluntary information programs and environmental regulation: Evidence from 'spare the air'. *Journal of Environmental Economics and Management*, 58(3):253–265.

Dang, J. (2018). An updated meta-analysis of the ego depletion effect. *Psychological Research*, 82(4):645–651.

DWD (2020). Schadstoffrelevante kriterien des deutschen wetterdienstes.

Ferraro, P. J., Miranda, J. J., and Price, M. K. (2011). The persistence of treatment effects with norm-based policy instruments: Evidence from a randomized environmental policy experiment. *American Economic Review*, 101(3):318–322.

Graff Zivin, J. and Neidell, M. (2009). Days of haze: Environmental information disclosure and intertemporal avoidance behavior. *Journal of Environmental Economics and Management*, 58(2):119–128.

Ito, K., Ida, T., and Tanaka, M. (2018). Moral suasion and economic incentives: Field experimental evidence from energy demand. *American Economic Journal: Economic Policy*, 10(1):240–267.

Liu, T., He, G., and Lau, A. (2017). Avoidance behavior against air pollution: evidence from online search indices for anti-PM2.5 masks and air filters in chinese cities. *Environmental Economics and Policy Studies*, 20(2):325–363.

Noonan, D. S. (2014). Smoggy with a chance of altruism: The effects of ozone alerts on outdoor recreation and driving in atlanta. *Policy Studies Journal*, 42(1):122–145.

Omnitrend (2016a). Befragung zum thema feinstaubalarm in stuttgart und umgebung. Online.

Omnitrend (2016b). Befragung zum thema feinstaubalarm in stuttgart und umgebung im zeitraum 26.02.2016 bis 28.02.2016. Online.

Reiss, P. C. and White, M. W. (2008). What changes energy consumption? prices and public pressures. *The RAND Journal of Economics*, 39(3):636–663.

Saberian, S., Heyes, A., and Rivers, N. (2017). Alerts work! air quality warnings and cycling. *Resource and Energy Economics*, 49:165–185.

Sexton, S. E. (2012). Paying for pollution? how general equilibrium effects undermine the "spare the air" program. *Environmental and Resource Economics*, 53(4):553–575.

Tice, D. M., Baumeister, R. F., Shmueli, D., and Muraven, M. (2007). Restoring the self: Positive affect helps improve self-regulation following ego depletion. *Journal of experimental* *social psychology*, 43(3):379–384.

Tribby, C. P., Miller, H. J., Song, Y., and Smith, K. R. (2013). Do air quality alerts reduce traffic? an analysis of traffic data from the salt lake city metropolitan area, utah, USA. *Transport Policy*, 30:173–185.

Ward, A. L. S. and Beatty, T. K. M. (2015). Who responds to air quality alerts? *Environmental and Resource Economics*, 65(2):487–511.

Welch, E., Gu, X., and Kramer, L. (2005). The effects of ozone action day public advisories on train ridership in chicago. *Transportation Research Part D: Transport and Environment*, 10(6):445–458.

Zafar, B. (2011). An experimental investigation of why individuals conform. *European Economic Review*, 55(6):774–798.

A DDA Reclassification Scheme

In order to create a set of "control" days to compare with DDA days, we classify days by reconstructing DWD's DDA classification algorithm and slightly loosening the DDA trigger conditions. We use DWD meteorological data from DWD Open Data and LUBW air quality data for each day during the PMA seasons from January 1, 2016 to December 31, 2019. Like DWD's official alert designation, we use daily mean PM10 concentrations from the Neckartor monitor (Condition 1). As historical DWD forecast data are unavailable, we use actual daily precipitation levels by type (Conditions 1 and 2), hourly wind speed and direction (Conditions 3 and 6) and radiosonde data (Conditions 4 and 5) from the DWD Open Data database.

Out of 733 PMA season days, we find 350 days where the DDA conditions were fulfilled compared to 250 DDA days by the DWD. We then compare our set of DDA days with the actual DWD DDA days and reclassify days into true positive (TP), false positive (FP), false negative (FN), and true negative (TN) according to the following conditions:

- True Positive (TP): Both our classification scheme and DWD classify a given day as a DDA day.
- False Positive (FP): Our classification scheme does not classify a given day as a DDA day while DWD does.
- False Negative (FN): Our classification scheme classifies a given day as a DDA day while DWD does not.
- True Negative (TN): Neither our classification scheme nor DWD classify a given day as a DDA day.

Each of these classes of days is recorded as a dummy variable that equals one when its conditions are fulfilled and zero otherwise.

We identify 219 true positive DDA days (30%) when a DDA had been broadcast and, according to our classification scheme, the DDA conditions were met, 31 false positive DDA days (4%) when a DDA was broadcast but, according to our classification scheme, the DDA conditions were not met, 162 false negative DDA days (22%) when a DDA was not called and, according to our classification scheme, the DDA conditions were met, and 321 true negative days (44%) when no DDA was called and the conditions were not met, according to our classification scheme. From January 1, 2016 to December 31, 2019, the city issued a DDA on 250 of 733 possible DDA season days. Overall, we classify 540 of 733 PMA season days (74%) in alignment with actual DDA status (either true positive or true negative).

We slightly relax several DDA conditions in our reclassification scheme. As DWD does not provide precipitation forecast data to the public, we use actual rainfall data for all days and, rather than projecting rainfall, use actual future day rainfall as the projected rainfall amount. This is unlikely to causes issues as rainfall forecasts, particularly of larger rainfall amounts, are fairly accurate. Furthermore, we consider days with less than 0.5mm of rainfall as rainless, as we deem this amount of rain as insufficient for clearing air pollutants from the air. TP, FN, FP, and TN classifications are mutually exclusive.

In order to reconstruct the conditions of the DWD DDA algorithm outlined in section 2, we first construct six daily dummy variables, each according to one of the following criteria:

- Criterion 1 equals one if the Neckartor PM10 concentration is greater than or equal to 30μg/m3, zero otherwise.
- Criterion 2: equals one if total daily rainfall is less than 0.5mm, zero otherwise. Snowfall and sleet are treated as rainless.
- Criterion 3: equals one if less than two-thirds of a day's hourly wind direction measurements are between 180°and 330°and daily mean wind speed is less than 3 km per hour, zero otherwise.
- Criterion 4: equals one if the nighttime inversion height is over 100 meters from the ground, zero otherwise. Using data from the 12am radiosonde flight, we calculate the height of the night time inversion level as the height at which air temperatures have risen at least 1°C compared to the air temperature at the ground.
- Criterion 5: equals one if the daytime mixing layer height is under 500 meters from the ground, zero otherwise. Using data from the 12pm noon radiosonde flight, we calculate the height of the daytime mixing layer using the 5 lowest altitudes at which the radiosonde measures an increase in temperature with increasing altitude. This criterion is met when at least 3 of the 5 lowest altitudes at which the radiosonde measures increasing temperatures are below 500m.
- Criterion 6: equals one when the average of a day's 24 hourly wind speed measurements is less than 3 km per hour, zero otherwise.

These six criteria are analogous to the six DWD DDA conditions. We then evaluate two possible paths to calling a DDA on a given day, as depicted in figure 3:

- Path 1: is fulfilled when, on a given day during the PMA season, Criterion 1 is met and Criterion 2 is met on that day (Issue Day) and the following day (Bridge Day). This corresponds to the left-most branch of the DWD Decision Tree in figure 3 where the primary condition, Condition 1 is met.
- Path 2: is fulfilled when, on a given day during the PMA season, Criterion 2 is met on the following day (Bridge Day) and the day thereafter (First Forecast Day) and Criterion 3 is met, while at least one of Criterion 4 or Criterion 5 is met, and at least 4 Conditions are met overall.

If either of these paths are fulfilled, we classify the day as a DDA day according to our algorithm. We then remove isolated DDA days (i.e. single DDA days with neither a DDA before or after a given classified DDA day) and add DDA days when there were single day gaps between two groups of DDA days of more than one day (i.e. the necessary condition to lift the DDA was not fulfilled).

B Additional Regression Results

Table 5: OLS Regression Results: Overall DDA Effect

	(1) log(VPD)	(2) log(VPD)	(3) log(VPD)	(4) log(VPD)	(5) log(VPD)	(6) log(VPD)
Don't Drive Appeal	0.00772*	0.0197***	0.00632	0.00321	0.00414	-0.00455
	[+180.2]	[+473.7]	[+153.8]	[+78.9]	[+100.9]	[-112.7]
	(0.00330)	(0.00501)	(0.00369)	(0.00693)	(0.00450)	(0.00652)
Full Sample:	Y	N	N	N	N	N
TP & FN Sample:	N	Y	N	Y	N	Y
Holidays Excluded:	N	N	Y	Y	Y	Y
Bridge & Issue Days Excluded:	N	N	N	N	Y	Y
Observations	26,626	11,996	20,899	10,040	16,787	7,641
Counters	43	43	43	43	43	43
Days	733	381	584	320	509	272
PMA Days	250	219	236	212	236	212
Non-PMA Days	483	162	348	108	273	60
Mean VPD	23,341	24,046	24,238	24,586	24,361	24,764
Mean log(VPD)	9.70	9.74	9.75	9.77	9.76	9.77

Dependent variable is log of vehicles per counter-day (VPD). Robust standard errors clustered on 22 counter sites in parentheses. All models include single-day lagged traffic, a full set of weather controls, first, second, and third-day lagged weather controls, counter fixed effects, year-month fixed effects, and day-of-the-week and holiday dummies. Absolute change relative to mean VPD in brackets.
*: Significant at 10%, **: Significant at 5%, ***: Significant at 1%.

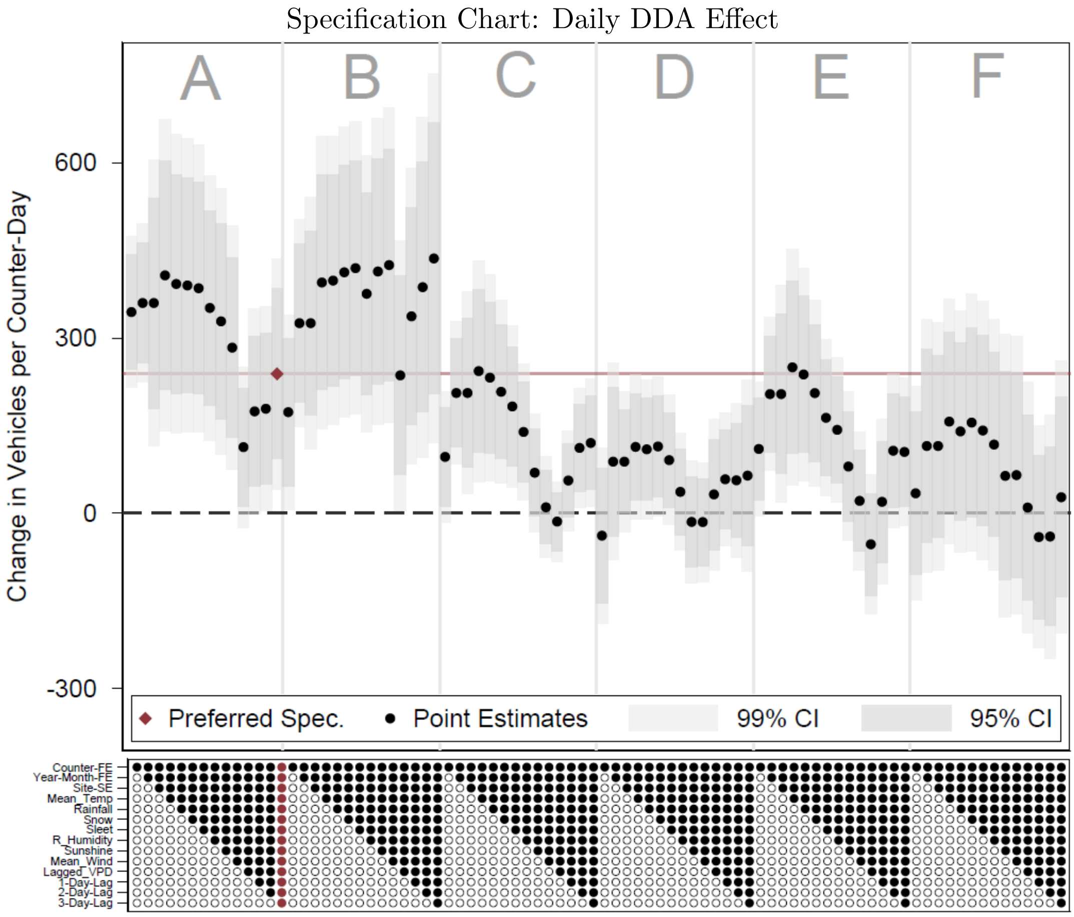

Figure 14: Specification chart depicting DDA effect point estimates for 90 OLS regressions. **Panel A**: Full sample, **Panel B**: True positive and false negative subsample, **Panel C**: Full sample without holidays, **Panel D**: True positive and false negative subsample without holidays, **Panel E**: Full sample without holidays, bridge, or issue days, **Panel F**: True positive and false negative subsample without holidays, bridge, or issue days. Bottom panel tracks specification additions. In each panel from left to right: counter fixed effects, year-month fixed effects, site standard errors, average daily temperature, rainfall amount, snow amount, sleet amount, relative humidity, sunshine hours, average daily wind speed, lagged traffic flows, single day lag of all variables, second day lag of all variables, and third day lag of all variables.