Using Big Data to Estimate the Environmental Benefits of Congestion Pricing: Evidence from California

APPAM, 2022

Antonio M. Bento[1,3], Rajat Kochhar[1], Ruozi Song[1] and Andrew Waxman[2]

[1]University of Southern California

[2]University of Texas at Austin

[3]NBER

Motivation and Key Question

- Large body of evidence of negative health impacts of local air pollution (LAP) such as NO_x, ozone, particulate matter

  Knittel et al. (2016), Currie and Walker (2011)

- Lots of work on behavior of point LAP sources (power plants, refineries) and good deal known about non-point sources from the transportation sector at an aggregated level (20-50 miles, weekly)

  Jacobsen, Sallee, Shapiro, van Benthem, 2021; Parry & Small, 2005; Bento, et al., 2009

- Criteria pollutants regulated at the air basin level under NAAQS, but could finer levels of adjustment, particularly for non-point sources have a big impact?

- This paper - empirical: How much does LAP vary at more granualar levels of time and space conditional on weather and air chemistry?

- This paper - counterfactual analysis: Would rearranging traffic across time or space have meaningful impacts on LAP concentrations?

Context

Los Angeles: well known for worst congestion in the US

- 104 hours excess traffic delay each year
- 13 working days/year
- $9.7 billion dollars/year - social cost of congestion ($2,408 per driver)

- LA's METRO board approved congestion pricing feasibility study in Feb 2019
- Despite electrification of new passenger vehicles, emissions from used cars remains a concern

The Link between Traffic and Pollution

- Tailpipe emissions from vehicles
  - Particulate Matter (PM): primary production from tires, secondary production from air chemistry: atmospheric ammonia plus NO_x or SO_2
  - NO and NO_2 are formed rapidly
  - NO can be short-lived (around 1 minute)
  - Ozone: NO_x + VOCs + Sunlight

Determinants of Pollution Levels

- Air Chemistry
- Vehicle Composition and Fuel Type
- Weather
- Background pollution, dispersion of fixed sources of pollution and nearby freeways
- Vehicular Speeds (Four Possible Channels)

Determinants of Pollution Levels

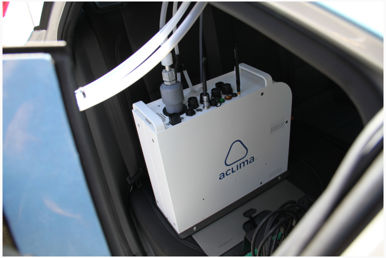
- Air Chemistry
- Vehicle Composition and Fuel Type
- Weather
- Background pollution, dispersion of fixed sources of pollution and nearby freeways
- Vehicular Speeds (Four Possible Channels)
  - Engine efficiency: Passenger vehicles operate more efficiently around speeds of 40-50MPH
  - Vehicle throughput: More cars per hour $\Rightarrow$ more pollution
  - Travel time: As road approaches capacity, each car is on roadway longer
  - Acceleration: More vehicle has to accelerate, less efficient (hills, wind, higher speeds or start & stop traffic

Determinants of Pollution Levels

- Air Chemistry
- Vehicle Composition and Fuel Type
- Weather
- Background pollution, dispersion of fixed sources of pollution and nearby freeways
- Vehicular Speeds (Four Possible Channels)
  - Engine efficiency: Passenger vehicles operate more efficiently around speeds of 40-50MPH
  - Vehicle throughput: More cars per hour $\Rightarrow$ more pollution
  - Travel time: As road approaches capacity, each car is on roadway longer
  - Acceleration: More vehicle has to accelerate, less efficient (hills, wind, higher speeds or start & stop traffic
    $\Rightarrow$ we develop flexible interaction of capacity and demand to allow for non-linear effects of density on pollution ▸ eq

Pollution: ACLIMA Google Car Monitors

- ACLIMA - Real-time pollution readings from monitors installed in two Google cars
  - August to October 2016; 6AM-6PM weekdays; freeways & local roads
    - On local roads, car slows down to take pictures
  - Readings of NO, NO_2, ozone, PM2.5 location and speed
  - Unit of observation: pollution level average over 5 minutes



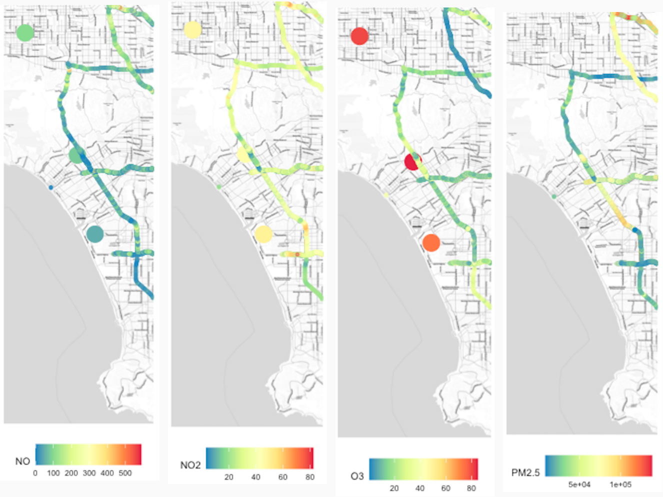
Traffic: California Performance Measurement System (PeMS)

- Speed and flow data from induction-loop detectors on LA freeways
- 1,871 detectors collecting over 5 minute intervals, average spacing .5 mile
- Variable of interest: Roadway vehicular density ($Cars/Mile$) at 5-minute intervals.

Other Data

- Use EPA AQS Ground Monitors to validate ACLIMA
- CHP accident & vehicle census counts from PeMS
- Weather: Northeast Regional Climate Center 9 stations hourly temps, wind speed & dir., sky cover, precipitation, humidity, and pressure

Aclima and EPA (Hourly Max) Pollution Oct 2016: West LA



(a) NO **(b)** NO_2 **(c)** O_3 **(d)** $PM_{2.5}$

Econometric Estimation

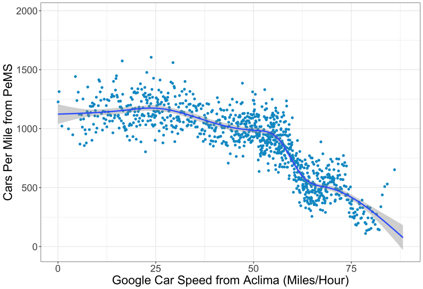
$$y_{ict} = \sum_{b=1}^{B} \alpha_b Cars/Mile_{it} \times \mathbf{1}\{Speed_{ict} \in S_b\} + \boldsymbol{\beta X}_{it} + \tau_t + \gamma_c + \phi_f + \epsilon_{ict}$$

- y_{ict}: Log of pollution level (parts per billion) from Google car c at time t, at location i. Separate regressions for pollution from NO, NO_2, ozone and PM2.5
- $Cars/Mile_{i,thdm}$: Log of number of vehicles per mile (density)
- $\mathbf{1}\{Speed_{ict} \in S_b\}$: Indicator: speed of Google car within specified bin. For results below: $S = \{< 20mph, 20 - 50mph, > 50mph\}$.

Econometric Estimation

$$y_{ict} = \sum_{b=1}^{B} \alpha_b Cars/Mile_{it} \times \mathbf{1}\{Speed_{ict} \in S_b\} + \beta \boldsymbol{X}_{it} + \tau_t + \gamma_c + \phi_f + \epsilon_{ict}$$

- y_{ict}: Log of pollution level (parts per billion) from Google car c at time t, at location i
- $Cars/Mile_{i,thdm}$: Log of number of vehicles per mile (density)
- $\mathbf{1}\{Speed_{ict} \in S_b\}$: Indicator: Google car speed within bin. $S = \{< 20mph, 20 - 50mph, > 50mph\}$



Econometric Estimation

$$y_{ict} = \sum_{b=1}^{B} \alpha_b Cars/Mile_{it} \times \mathbf{1}\{Speed_{ict} \in S_b\} + \beta \boldsymbol{X}_{it} + \tau_t + \gamma_c + \phi_f + \epsilon_{ict}$$

- y_{ict}: Log of pollution level (parts per billion) from Google car c at time t, at location i. Separate regressions for pollution from NO, NO_2, ozone and PM2.5
- $Cars/Mile_{i,thdm}$: Log of number of vehicles per mile (density)
- $\mathbf{1}\{Speed_{ict} \in S_b\}$: Indicator: speed of Google car within specified bin. For results below: $S = \{< 20mph, 20 - 50mph, > 50mph\}$.
- $\boldsymbol{X}_{i,hdm}$: hourly weather controls (second order polynomial of temperature, relative humidity, wind speed, as well as wind direction)
- τ_t, γ_c, ϕ_f: time, Google car and freeway fixed effects.
- Standard errors are clustered at hour and freeway level

Identification Strategy: Threats to Identification

- Endogeneity of traffic
  - -
- Unobserved point pollution sources & vehicle composition
  - -
  - -
- Chemical reactions between NO_x and ozone
  - -
  - -
- Weather conditions interact with traffic and pollution
  - -
- Less variation at lower speed
  - -
  - -

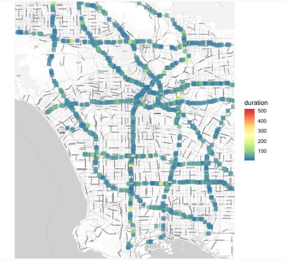
Identification Strategy: Threats to Identification

- Endogeneity of traffic
  - Instrument traffic with accidents
- Unobserved point pollution sources & vehicle composition
  - Hour-by-road segment fixed effects
  - Controls for proportion of heavy trucks by road-segment and hour from Caltrans vehicle Census
- Chemical reactions between NO_x and ozone
  - Control for previous pollution accumulation through pollution lags
  - Temporal granularity to capture short-lived pollutants
- Weather conditions interact with traffic and pollution
  - Interaction between cars per mile, weather, and speed bins
- Less variation at lower speed
  - Construct a lower bound based on estimates at higher speed
  - Compare with engineering literature

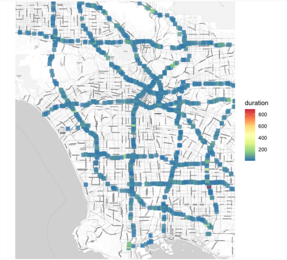
IV Strategy: Accidents

- Relevance: Accident acts as a supply shock, reducing the effective capacity of the freeway. As drivers re-route, nearby capacity is also affected.
- Validity: Congestion & speeds correlated with accident rates, want to use nearby accidents, but no easy rule for instrument set construction: 1st-stage LASSO following Belloni, et al. (2012) and Derenoncourt (2019)

- 42,449 accidents from August to October in 2016.
- 77.2% lasted longer than 5 minutes; 59.8% last longer than 15 minutes; 44.0% longer than 25 minutes.
- Simultaneous Aclima observation at nearby location, incident =1; otherwise =0.



(a) 6AM-10AM



(b) 10AM-2PM

Duration of Accidents Across Periods

Results: Effect of Freeway Traffic on Pollution - OLS

	OLS			
log of concent. in ppb	$PM_{2.5}$	NO	NO_2	O_3
log(car/mile):speed <20	0.118***	0.096	0.096*	-0.003
	(0.003)	(0.062)	(0.055)	(0.008)
log(car/mile):speed 20-50	0.125***	0.098*	0.098*	-0.05
	(0.003)	(0.059)	(0.053)	(0.007)
log(car/mile):speed 50+	0.134***	0.099	0.098*	-0.001
	(0.003)	(0.060)	(0.054)	(0.007)
Observations	38,335	41,484	43,055	35,604
R^2	0.814	0.642	0.505	0.880

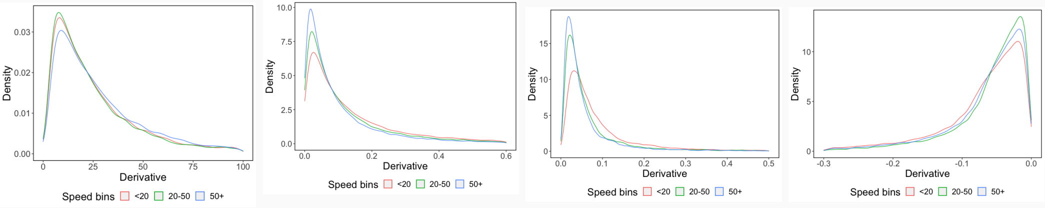
Notes: Includes hour, freeway by day-of week, and Google Car fixed effects. Weather controls include second order polynomial of temperature, relative humidity, wind speed, as well as wind direction. Standard errors are clustered at hour and freeway level. *$p<0.1$; **$p<0.05$; ***$p<0.01$

Results: Effect of Freeway Traffic on Pollution - IV

	OLS				IV-LASSO			
log of concent. in ppb	$PM_{2.5}$	NO	NO_2	O_3	$PM_{2.5}$	NO	NO_2	O_3
log(car/mile):speed <20	0.118***	0.096	0.096*	-0.003	0.158***	0.324***	0.302***	-0.461***
	(0.003)	(0.062)	(0.055)	(0.008)	(0.023)	(0.113)	(0.116)	(0.148)
log(car/mile):speed 20-50	0.125***	0.098*	0.098*	-0.05	0.151***	0.253***	0.204**	-0.370***
	(0.003)	(0.059)	(0.053)	(0.007)	(0.020)	(0.096)	(0.086)	(0.141)
log(car/mile):speed 50+	0.134***	0.099	0.098*	-0.001	0.177***	0.205**	0.175*	-0.410***
	(0.003)	(0.060)	(0.054)	(0.007)	(0.017)	(0.093)	(0.094)	(0.138)
Observations	38,335	41,484	43,055	35,604	38,309	41,458	43,029	35,578
R^2	0.814	0.642	0.505	0.880	0.644	0.841	0.429	0.859

Notes: Includes hour, freeway by day-of week, and Google Car fixed effects. Weather controls include second order polynomial of temperature, relative humidity, wind speed, as well as wind direction. Standard errors are clustered at hour and freeway level. *$p<0.1$; **$p<0.05$; ***$p<0.01$

Results: Derivative of Pollution with Respect to Density



(a) $PM_{2.5}$ **(b)** NO **(c)** NO_2 **(d)** O_3

Notes: The four panels plot the smoothed kernel density of the derivative of pollution in parts per billion with respect to vehicle density. Each colored line reflects an individual speed bin, and each observation is the reduced-form impact of traffic on pollution, based on the particular conditions for that observation.

Results: Effect of Freeway Traffic on Pollution - Coefficient Decomposition

Higher pollution from slower speeds can come from the fact that there are more cars on freeway for longer, OR

Vehicle engines are operating at less efficient speeds, which produce more pollution per mile traveled

$$\text{Pollution from 1 vehicle} = \underbrace{\frac{\text{Pollution from 1 vehicle}}{\text{Travel time}}}_{\text{Engine efficiency effect}} \times \text{Travel time}$$

Results: Effect of Freeway Traffic on Pollution - Coefficient Decomposition

	NO	NO_2	$PM_{2.5}$	O_3
log(car/mile):speed_bin<20	0.324	0.302	0.158	-0.461
Pollution contribution of 1 vehicle	0.106	0.052	18.527	-0.055
Engine inefficiency effect	0.026	0.013	4.620	-0.014
log(car/mile):speed_bin20-50	0.253	0.204	0.151	-0.370
Pollution contribution of 1 vehicle	0.074	0.032	15.542	-0.039
Engine inefficiency effect	0.043	0.018	8.932	-0.022
log(car/mile):speed_bin50+	0.205	0.175	0.177	-0.410
Pollution contribution of 1 vehicle	0.080	0.035	23.969	-0.060
Engine inefficiency effect	0.083	0.036	24.710	-0.062
Travel Time Effect	0.042	0.018	8.664	-0.022

Notes: Engine efficiency and travel time effects are calculated with the travel time effect is a bounding case based on speeds in the second bin.

Further Robustness

- Alternative fixed effects, use of lagged dependent variable
- Varying speed bin cutoffs
- Spliting sample by region of LA
  - Differential effects due to weather interactions, background industrial pollution, variation in extent of VOC-limited ozone production
- Time-of-day variation
- IV using smaller instrument set

Conclusions & Next Steps

- Granular estimates of the effect of vehicle density and speed on local air pollution in Los Angeles: removing 1 vehicle can reduce as much as 24.0 ppb of $PM_2.5$, 0.11 ppb of NO, 0.052 ppb of NO_2, ozone more challenging to interpret
  - By comparison, NO_2 reductions between 1990-2018 were 28 ppb in DTLA (EPA NAAQS Trends)

- Next Steps:
  - Characterize pollution exposure from alternative counterfactual vehicle allocations: how low can pollution go if cars are moved around in time and space
  - Effects on local roads
  - Decompose local pollution exposure by demographic groups
  - Local pollution impacts of freeway tolling

References

[1] Currie, Janet and Walker, Reed Traffic congestion and infant health: Evidence from E-ZPass *American Economic Journal: Applied Economics.3. (2011), 65-90*.

[2] Belloni, Alexandre and Chen, Daniel and Chernozhukov, Victor and Hansen, Christian Sparse models and methods for optimal instruments with an application to eminent domain *Econometrica.6. (2012), 2369-2429*.

[3] Derenoncourt, Ellora Can you move to opportunity? Evidence from the Great Migration *American Economic review.2. (2022) 369-408*.

[4] Knittel, Christopher R and Miller, Douglas L and Sanders, Nicholas J Caution, drivers! Children present: Traffic, pollution, and infant health *Review of Economics and Statistics.98. (2016), 350-366*.

[5] Parry, Ian W. H. and Bento, Antonio Revenue Recycling and the Welfare Effects of Road Pricing *The Scandinavian Journal of Economics.103. (2001), 645-671*.

Appendix Slides

Hypothetical relationship between pollution & density

Back

Pollution concentration as a function of lag level, emissions, wind, photochemistry

$$p_{i,t} = p_{i,t-1} + e_{i,t} + \Delta w_{i,t} + \Delta pc_{i,t}$$

Determinants of emissions

$$\begin{aligned} e_{i,t} &= \sum_{c \in C_{i,t}} E_c\left(speed'(t), ee_c(speed)\right) \\ &\approx V_{i,t} \cdot \bar{\varepsilon}_{i,t}\left(\text{speed}'(t), \text{ee}(\text{speed})\right) \\ &= \kappa_i \cdot density_{i,t} \cdot speed_{i,t} \cdot \bar{\varepsilon}_{i,t}\left(\text{speed}'(t), \text{ee}(\text{speed})\right) \end{aligned}$$

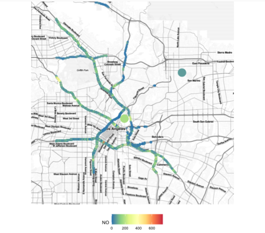
Hypothetical relationship between pollution & density

Back

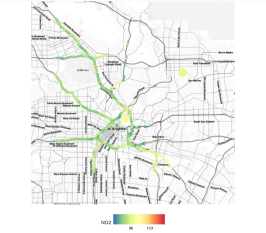
$$
\begin{aligned}
\frac{\mathrm{d}p_{i,t}}{\mathrm{d}density_{i,t}} &= \kappa_i \cdot speed_{i,t} \cdot \bar{\varepsilon}_{i,t}(\cdot) && \textit{density effect} \\
&+ \kappa_i \cdot density_{i,t} \cdot \bar{\varepsilon}_{i,t}(\cdot) \frac{\mathrm{d}speed_{i,t}}{\mathrm{d}density_{i,t}} && \textit{travel time effect} \\
&+ \kappa_i \cdot density_{i,t} \cdot speed_{i,t} \cdot \sum_{c \in C_{i,t}} \frac{\mathrm{d}speed'_c(t)}{\mathrm{d}density_{i,t}} \cdot \frac{\partial \bar{\varepsilon}_{i,t}(\cdot)}{\partial speed'_c(t)} && \textit{acceleration effect} \\
&+ \kappa_i \cdot density_{i,t} \cdot speed_{i,t} \cdot \sum_{c \in C_{i,t}} \frac{\mathrm{d}ee_c}{\mathrm{d}density_{i,t}} \cdot \frac{\partial \bar{\varepsilon}_{i,t}(\cdot)}{\partial ee_c} && \textit{engine efficiency effect}
\end{aligned}
$$

Aclima and EPA (Hourly Max) Pollution Oct 2016: DTLA

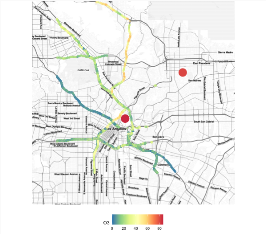
Back



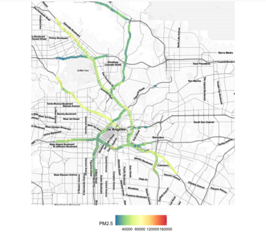
(a) NO



(b) NO_2



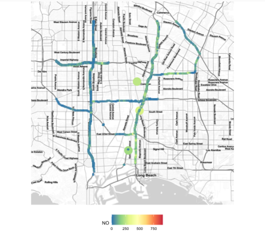
(c) O_3



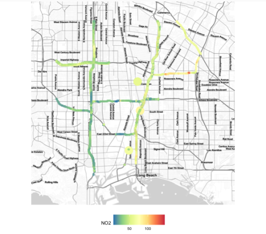
(d) $PM_{2.5}$

Aclima and EPA (Hourly Max) Pollution Oct 2016: Long Beach

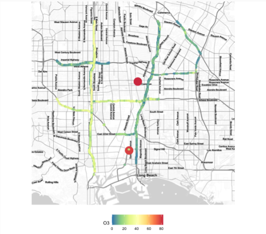
Back



(a) NO



(b) NO_2



(c) O_3

(d) $PM_{2.5}$

Optimal Congestion Pricing - Pollution Externality