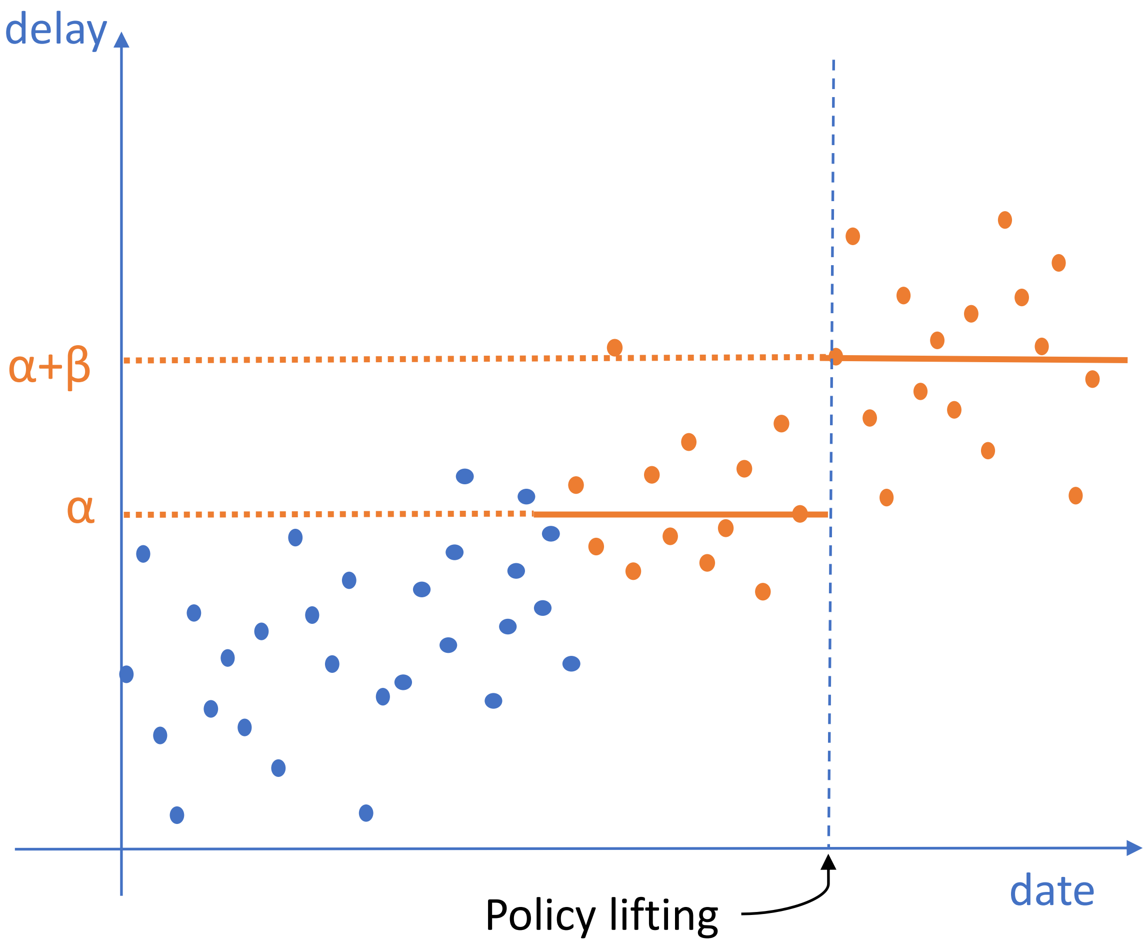
‘Identification’ of causal effect (review)

- What if the timing of event is intended to coincide with the changes in outcomes?
  - As opposed to the changes being caused by the event?
  - Assumption: Event is uncorrelated with trends in outcomes
- What would outcomes have looked like in the absence of the policy?
  - Would the average delay have stayed at α?
  - Assumption: ‘Treated’ observations would resemble ‘control’ observations in the absence of the event

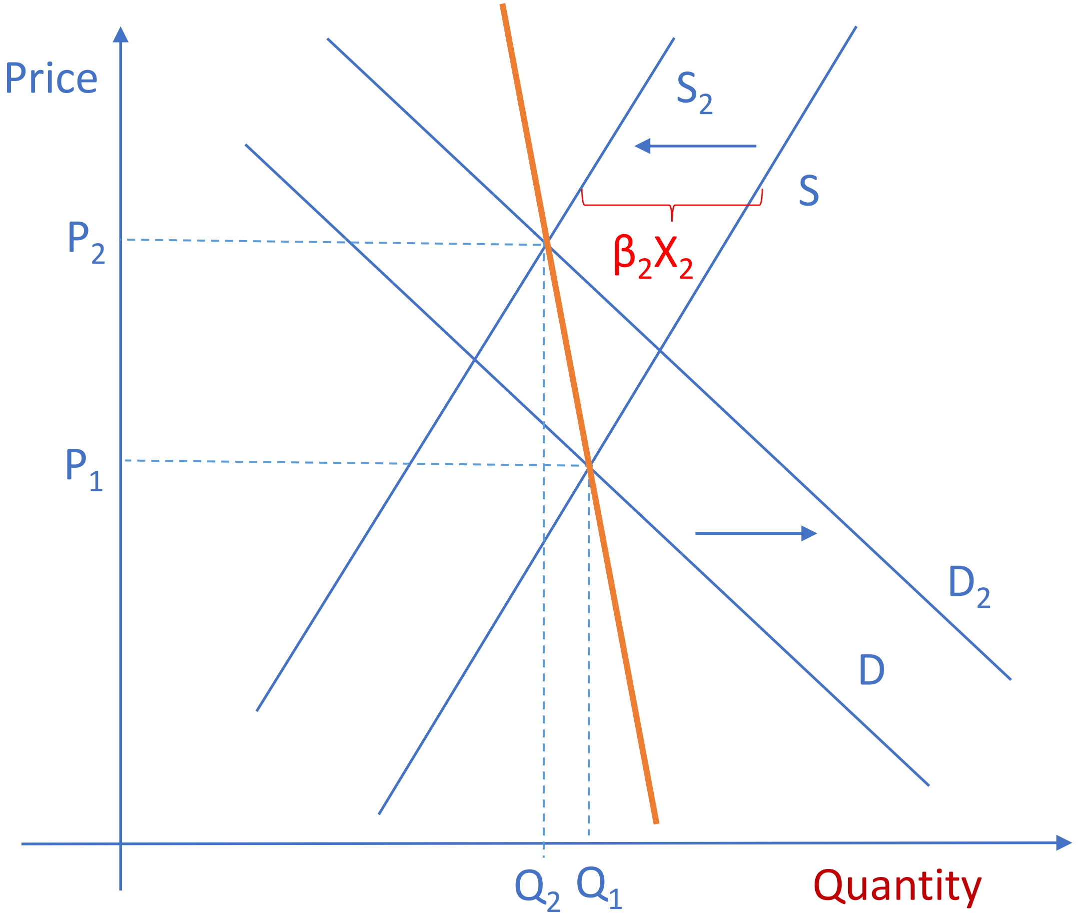


Omitted Variable Bias (review)

e.g., if supply shift is caused by X_2: conditional on the effect of X_2 on quantity, the relationship between price and quantity lets us estimate the slope of the supply curve:

$$Ln(Quantity) = \beta_0 + \beta_1 Ln(Price) + \beta_2 X_2 + \varepsilon$$

Not including the variable X_2 in the regression can **bias** our estimate of β_1.



Hall, Palsson, and Price (2018)

Table 3: Effect of Uber on log transit ridership

	Uber entry				Uber penetration			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
UberX	0.00263	-0.0591**	0.0598**	-0.00190	0.0138***	-0.00483	0.0328***	0.00758
	(0.0143)	(0.0295)	(0.0236)	(0.0364)	(0.00515)	(0.00526)	(0.00652)	(0.00677)
Above median population		0.0666**		0.0665**		0.0228***		0.0343***
× UberX		(0.0294)		(0.0307)		(0.00716)		(0.00796)
Above median ridership			-0.0811***	-0.0811***			-0.0281***	-0.0323***
× UberX			(0.0292)	(0.0292)			(0.00977)	(0.0100)
Observations	71,386	71,386	71,386	71,386	58,015	58,015	58,015	58,015
Clusters	309	309	309	309	227	227	227	227

Hall, Palsson, and Price (2018)

- Causal identification requires Uber's entry choice is unrelated to transit ridership.
- E.g., what if transit ridership change led to Uber entering the city (as opposed to the other way around)?

Table 2: Linear regressions predicting when and whether Uber enters an MSA

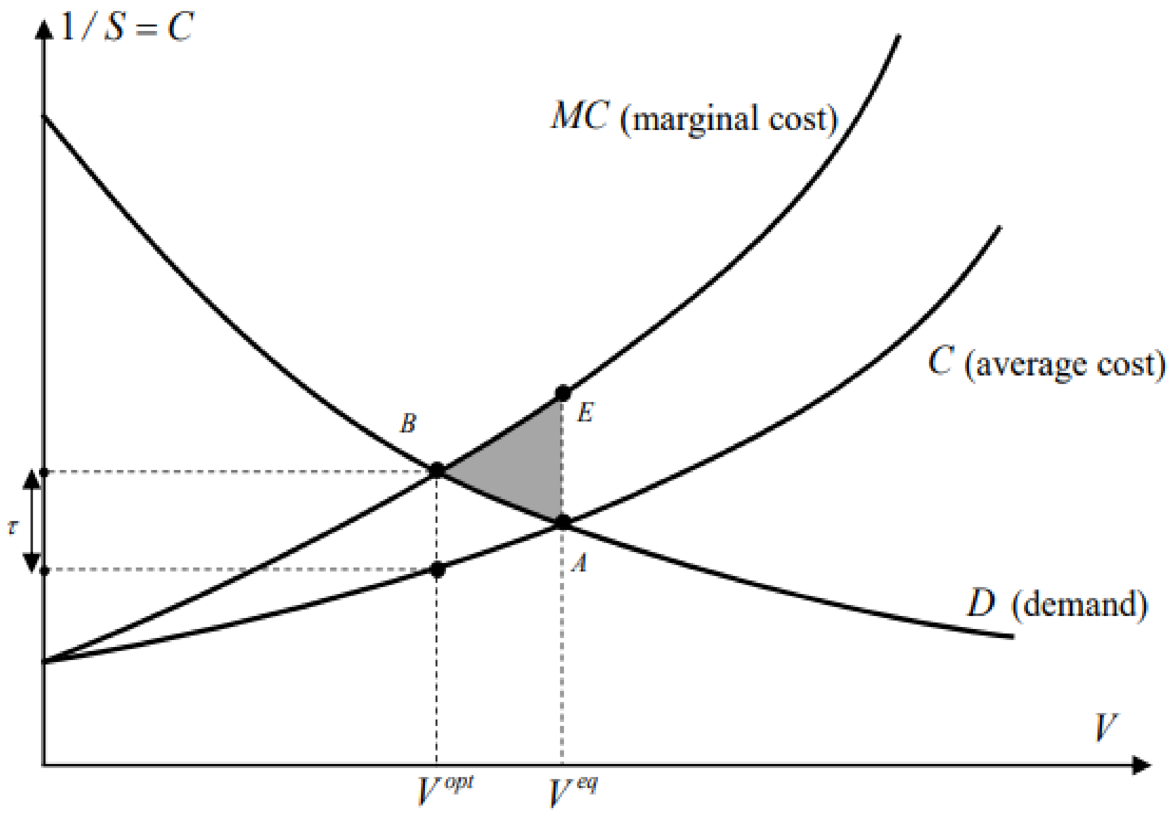
	Date UberX entry (1)	Did UberX enter (2)
Log(population) (σ)	-103.2***	0.256***
	(4.690)	(0.00705)
Percent with bachelor's degree (σ)	-41.77***	0.180***
	(5.108)	(0.00857)
Median age (σ)	30.90***	-0.0518***
	(5.777)	(0.00844)
Median income (σ)	-11.40**	-0.0288***
	(4.905)	(0.00963)
Excess unemployment (σ)	-41.73***	0.0336***
	(4.713)	(0.00789)
Percent work trips transit (σ)	-9.956**	-0.0792***
	(4.968)	(0.00948)
Capital expenditures on public transit (σ)	-4.868	-0.00152
	(4.939)	(0.00698)
Dist from Uber HQ (σ)	11.99***	0.00823
	(4.389)	(0.00662)
Trend in log(population) (σ)	11.46	0.0214*
	(8.225)	(0.0120)
Trend in median income (σ)	1.641	-0.0298
	(13.65)	(0.0204)
Observations	197	386
Adjusted R-squared	0.383	0.394

Anderson (2014)

Anderson, Michael L. 2014. "Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion." *American Economic Review*, 104 (9): 2763-96.

- In the US, only 1-2% of travel miles via mass transit. Yet, public transit subsidies are popular in large cities like Los Angeles!
  - In 2008, 67% of LA county voted to allocate $26 billion to transit over 30 years
  - Why? if few voters are transit riders?
- Public transit relieves congestion and benefits more than just riders?
  - But only moving a small fraction of drivers off the streets!
- Paper:
  - Commuters on different roads and times face different levels of congestion
  - Transit attracts commuters who face the worst congestion, who would otherwise drive on the most congested roads at the most congested times.
  - Drivers on heavily congested roads have a much higher marginal effect on congestion
  - So, transit has a large impact on reducing congestion.

Market for road travel



Anderson, Michael L. 2014. "Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion." *American Economic Review*, 104 (9): 2763-96.

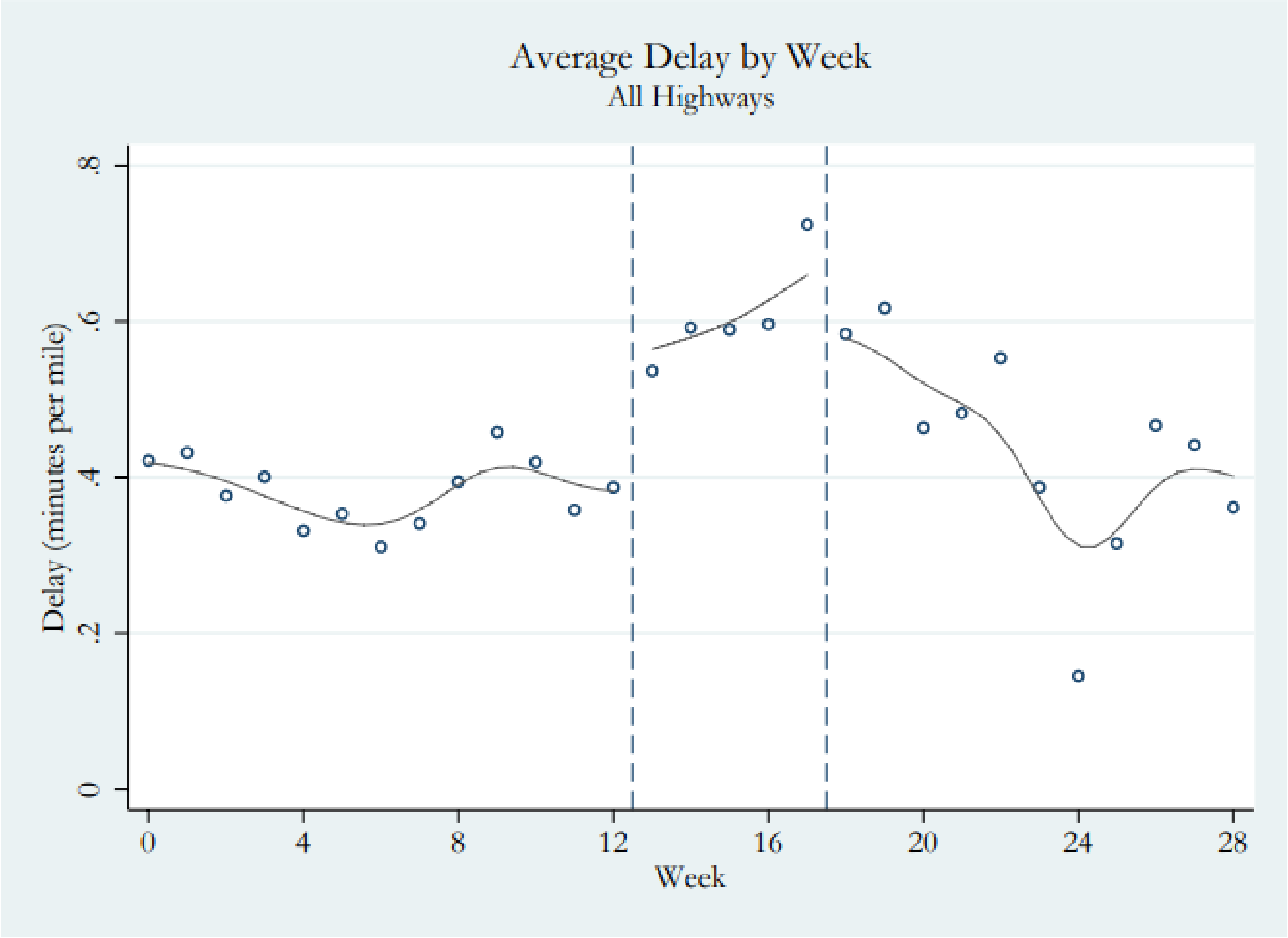
- **Question:** By how much does LA's public transit relieve congestion?

- Exploit a "natural experiment"
  - October 2003: LA public transit workers began a 35-day strike shutting down bus and rail lines
  - Look at effect on hourly traffic speeds on major Los Angeles freeways
  - Using a Regression Discontinuity Design (RDD)

- Leads to increase in average travel delays of 47% during peak hours
  - Largest effects on freeways that parallel popular transit lines

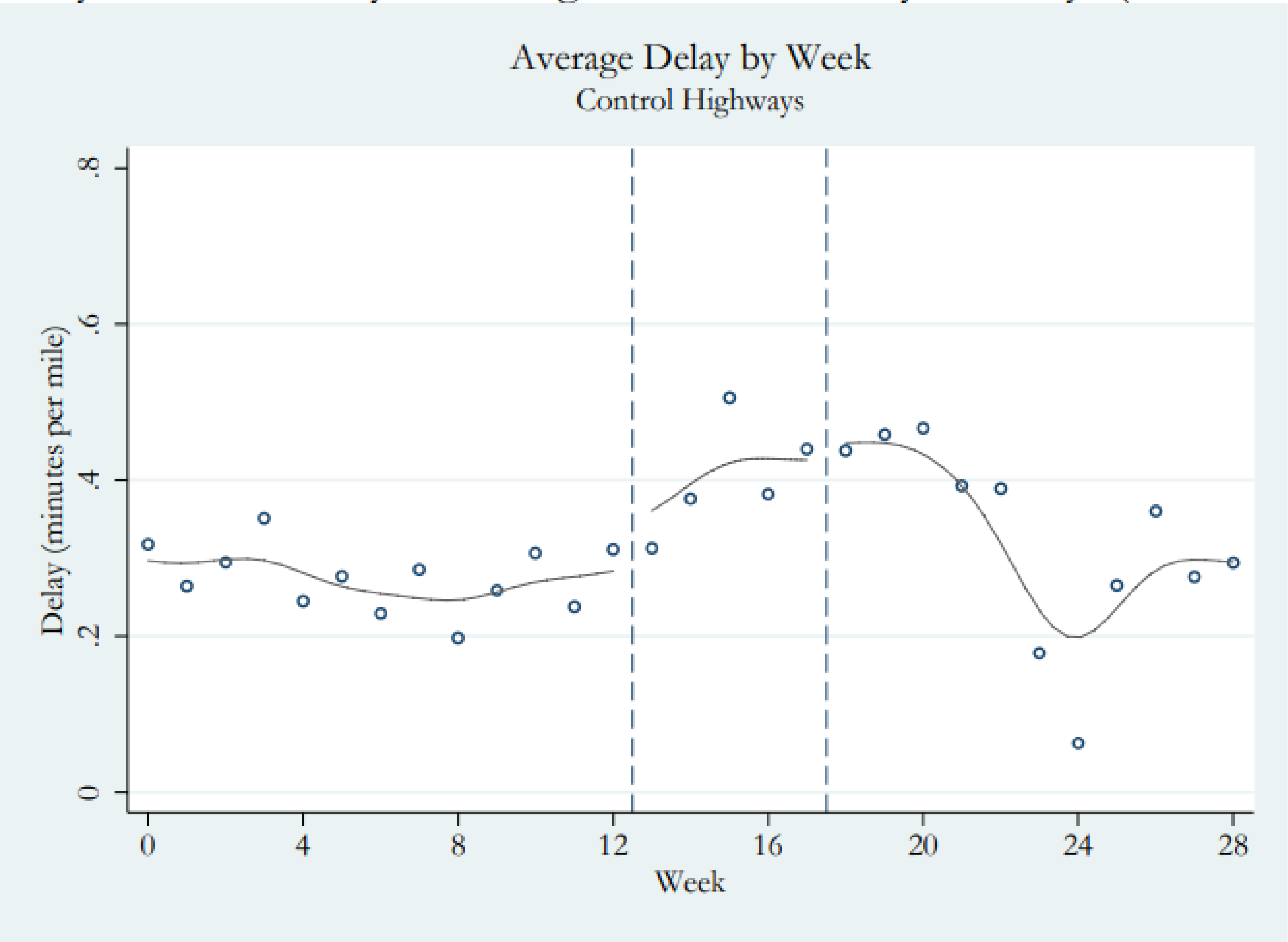
Anderson, Michael L. 2014. "Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion." *American Economic Review*, 104 (9): 2763-96.

Figure 2: Weekly Peak Hr. Delay on Major L.A. Freeways (7/14/03–1/30/04)



Anderson, Michael L. 2014. "Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion." *American Economic Review*, 104 (9): 2763-96.

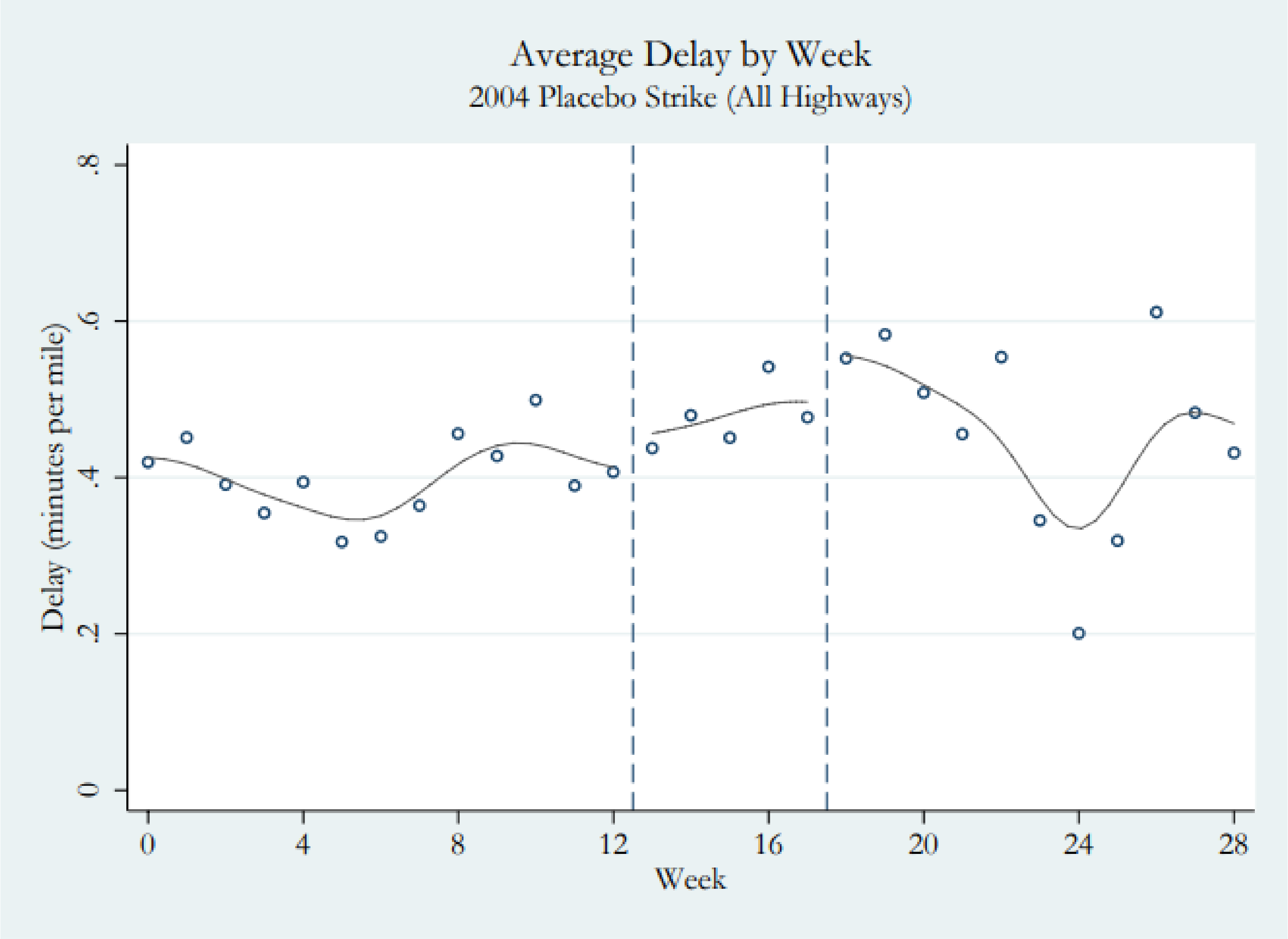
Weekly Peak Hr. Delay on Orange/Ventura County Freeways (7/14/03–1/30/04)



Neighboring counties unaffected

Anderson, Michael L. 2014. "Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion." *American Economic Review*, 104 (9): 2763-96.

Weekly Peak Hr. Delay on Major L.A. Freeways 1 Year Later (7/14/04–1/30/05)



Delay is not a seasonal effect

Anderson, Michael L. 2014. "Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion." *American Economic Review*, 104 (9): 2763-96.

Commuter *i* maximizes utility function:

$$U_i = X_i - T(s_i(R_i), a_i(R_i), w_i(R_i), m)$$

By choosing to ride rail (*R*=1) or drive (*R*=0). Subject to an income/budget constraint:

$$s.t. \;\; Y_i = X_i + m \cdot (p_r R_i + p_d(1 - R_i))$$

Value of travel time (VOT) varies across commuters and by activity:

- For more on VOT: Small, K., and E. Verhoef. 2007. *The Economics of Urban Transportation*. New York: Routledge.

Solves an optimization problem:

$$U_i = X_i - v_i[R_i(\frac{m}{s_r} + c(a_{ri} + w_r)) + (1 - R_i)(\frac{m}{s_d} + c(a_d + w_{di}))]$$

$$s.t. \;\; Y_i = X_i + m \cdot (p_r R_i + p_d(1 - R_i))$$

Anderson, Michael L. 2014. "Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion." *American Economic Review*, 104 (9): 2763-96.

$$U_i = X_i - v_i[R_i(\frac{m}{s_r} + c(a_{ri} + w_r)) + (1 - R_i)(\frac{m}{s_d} + c(a_d + w_{di}))]$$

$$s.t. \;\; Y_i = X_i + m \cdot (p_r R_i + p_d(1 - R_i))$$

Choose rail if:

$$[c(a_{ri} + w_r) + \frac{m}{s_r}] - [c(a_d + w_{di}) + \frac{m}{s_d}] \leq \frac{m}{v_i}(p_d - p_r)$$

Say, rail access times a_{ri} vary with some known probability distribution. Then:

$$P(R_i = 1) = P[c \cdot a_{ri} - \frac{m}{v_i}(p_d - p_r) \leq c(a_d + w_d - w_r) + \frac{m}{s_d} - \frac{m}{s_r}]$$

Anderson, Michael L. 2014. "Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion." *American Economic Review*, 104 (9): 2763-96.

$$P(R_i = 1) = P[c \cdot a_{ri} - \frac{m}{v_i}(p_d - p_r) \leq c(a_d + w_d - w_r) + \frac{m}{s_d} - \frac{m}{s_r}]$$

- Calibrate model parameters (on the right) to match observations from data and other studies.
- Vary unobserved parameters (e.g., rail access times a_{ri}) until we can match share of commuters choosing rail.

Can add more individual heterogeneity (in experienced congestion delay w_{di}):

$$P(R_i = 1) = P[c(a_{ri} - w_{di}) - \frac{m}{v_i}(p_d - p_r) \leq c(a_d - w_r) + \frac{m}{s_d} - \frac{m}{s_r}]$$

- Vary w_{di} to match average congestion delay w_d.

Data sources

- **Google Trends**
- **Helsinki travel time matrix**
- **Google Maps APIs**
- **Satellite night lights**
- **US census (nhgis.org)**
- **IPUMS**
  - **International**
- **Digitraffic Traffic Monitoring System data**
- **Travel surveys**
  - **e.g., NHTS**
- **Open Street Maps**
- **American Time Use Survey**
- **Federal Housing Finance Agency**
- **US Bureau of Labor Statistics**
- **General Transit Feed Specification**
- **…**