**Title**

Induced Travel Estimation Revisited

**Author**

Manville, Michael

**Publication Date**

2024-07-01

RESEARCH REPORT

# Induced Travel Estimation Revisited

Michael Manville
Professor of Urban Planning
July 2024

UCLA Institute of Transportation Studies

## About the UCLA Institute of Transportation Studies

The UCLA Institute of Transportation Studies supports and advances cutting-edge research, the highest-quality education, and meaningful and influential civic engagement on the many pressing transportation issues facing our cities, state, nation and world today. The institute is part of the University of California Institute of Transportation Studies, a four-campus consortium that includes UC Berkeley, UC Davis and UC Irvine. UCLA ITS is also a proud partner of the Pacific Southwest Region 9 University of Transportation Center, a federally funded research network with seven other universities.

## Acknowledgments

I thank all the panel members, as well as Annie Nam, Warren Whiteaker, Eric Sundquist, Tarek Hatata, Joan Walker, and Jamey Volker for helpful conversations. Ariege Besson provided research assistance.

The Institute of Transportation Studies at UCLA acknowledges the Gabrielino/Tongva peoples as the traditional land caretakers of Tovaangar (the Los Angeles basin and So. Channel Islands). As a land grant institution, we pay our respects to the Honuukvetam (Ancestors), ‘Ahiihirom (Elders) and ‘Eyoohiinkem (our relatives/relations) past, present and emerging.

## Disclaimer

The opinions expressed here are the author’s alone.

This report was funded by SCAG™

# Induced Travel Estimation Revisited: Findings from an Analysis Based on an Expert Panel

## Expert Panel Members:

Gilles Duranton
Bruce Griesenbeck
Susan Handy
Donald Hubbard
Chris McCahill
Ronald Milam
Matthias Sweet
Brian Taylor

# TABLE OF CONTENTS

# Table of Contents

**Introduction and Summary** ........ 8

**I. The VMT Effects of Adding Free Capacity** ........ 18

- 1.1 The Conceptual Challenge of Measuring Induced VMT ........ 19
- 1.2 The Empirical Challenge of Measuring Induced Travel ........ 22
- 1.3 Convergence on Unit Elasticity: The New Generation of Induced Travel Studies ........ 26
- 1.4 A Skeptical Look at Unit Elasticity ........ 29

**II. The VMT Effects of Managed Lanes** ........ 42

**III. What VMT Does and Doesn't Tell Us** ........ 48

- 3.1 VMT is a Common Metric for Environmental Impact ........ 49
- 3.2 VMT's Costs Cannot be Separated from Its Benefits ........ 49
- 3.3 VMT Actually Has No Benefits ........ 52
- 3.4 VMT Suppresses Use of Other Modes ........ 54

**IV. Conclusion** ........ 59

**References** ........ 60

## List of Figures

Figure 1. Example of Induced VMT Calculator....9
Figure 2. Map of Location of 24 Freeway Projects Across Nine California Counties ....31
Figure 3. Speed vs Flow on the 405 Freeway. From Marshall (2018)....45
Figure 4. VMT. GDP and Transportation Carbon Emissions ....48
Figure 5. VMT and VMT per Capita, 50 Large UZAs, 2021. Source: Federal Highway Administration ....49
Figure 6. Carbon Monoxide Emissions and Vehicle Travel, 1980-2022 ....51
Figure 7. Traffic Fatalities and Motor Vehicle Use, 1913-2021 .... 52
Figure 8. Driving and Transit Use, 2000-2019....54
Figure 9. VMT and Transit Ridership, 2000-2020, Los Angeles (2000=100) .... 55
Figure 10. Barrier Freeway Effect .... 57

## List of Tables

Table 1. Estimated Sources of Induced VMT .... 28

# CONTENTS

# Introduction and Summary

Suppose a government adds two lanes to a highway in both directions, but then restricts access to those lanes. The lanes will be "managed"—open only to people in carpools, or to people in single-occupancy vehicles who pay a toll. The toll will vary in response to demand, with the goal of avoiding congestion and keeping vehicles moving.

Will the total amount of driving in the region rise, fall, or not change?

This question lacks an intuitive answer. On the one hand adding road capacity should induce more travel, causing the total amount of driving to rise. On the other hand restricting access, especially via pricing, might reduce demand for these lanes. This could make driving fall. A third possibility is that these two countervailing forces (more supply and higher prices) will cancel each other out, leaving driving largely unchanged. There are intermediate possibilities as well. The fact that the new lanes are priced, for example, might make total driving rise, but rise by less than would be the case if the new lanes were free. This result, however, could depend on the mix of vehicles in the lanes: on how many are carpools, and how many are paying tolls.

Given these ambiguities, theory cannot offer a clear prediction of how any given managed lane will or (won't) induce new travel. Normally, when theory falls short, we can lean on empirics. But empirics in this instance also offer little guidance. No empirical research examines the specific question of whether and by how much new managed freeway lanes induce new travel. This is so largely because managed lanes themselves are quite rare. A large empirical literature examines induced travel, but that literature draws on evidence from existing roads, and most existing roads, in the United States and elsewhere, are free. Little evidence, as a result, speaks directly to the question of what happens when a government builds new managed lanes, be they High Occupancy Vehicle (HOV) lanes, High Occupancy/Toll (HOT) lanes, or lanes that are purely tolled.[1]

In California, however, the question has become important, for two reasons. First, much of the new highway capacity being proposed in the state is HOT lanes. The Southern California Association of Governments (SCAG), for example, has proposed an expansive network of HOT lanes throughout its six-county region. Second, the state of California has changed the way it evaluates transportation's environmental impacts. Specifically, the state has, in response to Senate Bill (SB) 743, issued an order saying that transportation impacts will now be measured in terms of Vehicle Miles Traveled (VMT), rather than Level of Service (LOS). The state, in essence, now treats new VMT as equivalent to climate damage or environmental degradation. To quote a Caltrans document explaining the policy, the state treats "traffic as the pollutant."[2] Under this interpretation of SB 743, any project likely to induce new VMT must either mitigate that VMT—through projects that will reduce driving by a comparable amount—or face difficulty being approved.

In these circumstances, a lot rides on how induced travel is estimated. One of the state's preferred estimation methods is the National Center for Sustainable Transportation's (NCST's) Induced Travel Calculator.[3] The calculator is based on a recent body of academic research, kickstarted by Duranton and Turner's (2011) seminal article in the *American Economic Review*, about how new roads induce travel. This literature uses regression analysis to estimate induced VMT

---

[1] These lanes go by different names. HOT lanes are sometimes referred to as "Express Lanes" or "priced managed lanes." I use these terms interchangeably. I also use "HOV lanes" and "carpool lanes" interchangeably.

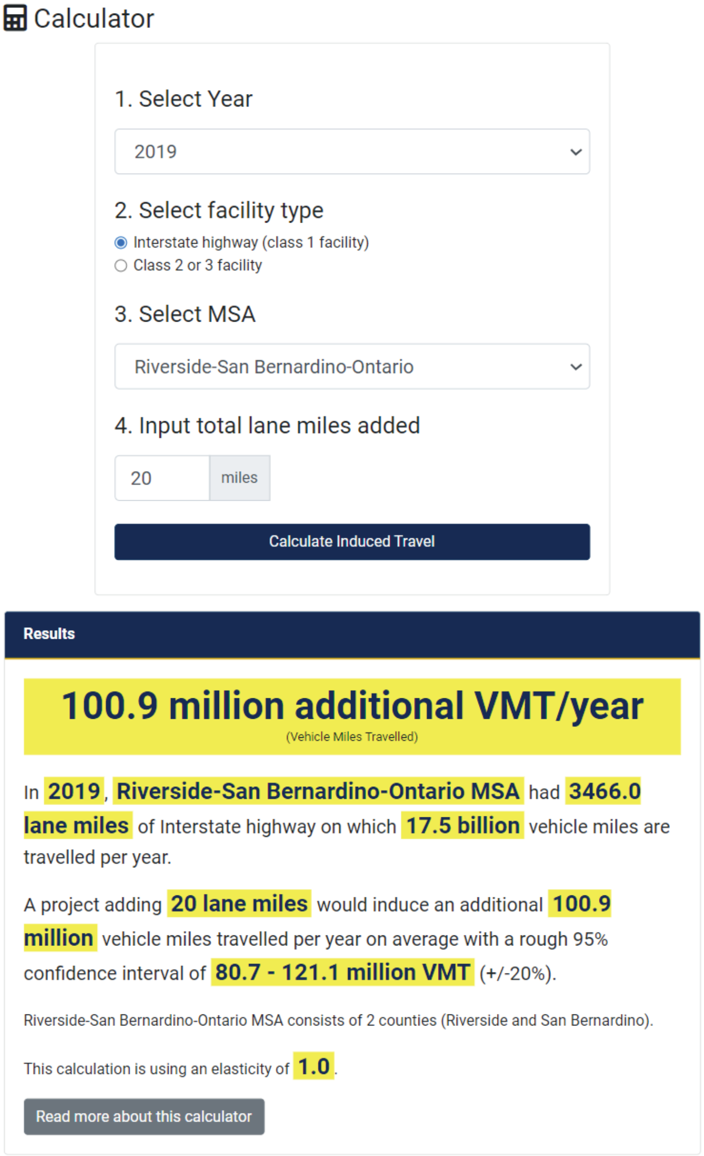
[2] Caltrans, no date.

[3] The calculator can be found at https://travelcalculator.ncst.ucdavis.edu/

and present it as an elasticity. An elasticity is a metric that captures the proportional responsiveness of one variable to another: an elasticity of 0.1 suggests that for every 10 percent increase in a variable *x*, there would be a one percent increase in a variable *y*.

A consensus finding in the literature is that freeway expansions have an elasticity of 1. An elasticity of 1 is called *unit elasticity*, and implies that every proportional increase in freeway lane mileage will yield an equally proportional increase in VMT. A ten percent increase in lane mileage, for example, will cause a ten percent increase in driving. The NCST calculator combines that elasticity with data on the existing amount of VMT and road capacity in a given region, and uses it to estimate the amount of induced VMT a given project will cause. In essence, the calculator converts the main finding of the academic literature into a simple tool for practitioners. Figure 1 below shows some calculator output for a hypothetical road expansion (of 20 new lane miles) in the Riverside-San Bernardino MSA:

**Figure 1. Example of Induced VMT Calculator**



Calculator

1. Select Year

2019

2. Select facility type

● Interstate highway (class 1 facility)
○ Class 2 or 3 facility

3. Select MSA

Riverside-San Bernardino-Ontario

4. Input total lane miles added

20 miles

**100.9 million additional VMT/year**

(Vehicle Miles Travelled)

In **2019**, **Riverside-San Bernardino-Ontario MSA** had **3466.0 lane miles** of Interstate highway on which **17.5 billion** vehicle miles are travelled per year.

A project adding **20 lane miles** would induce an additional **100.9 million** vehicle miles travelled per year on average with a rough 95% confidence interval of **80.7 - 121.1 million VMT** (+/-20%).

Riverside-San Bernardino-Ontario MSA consists of 2 counties (Riverside and San Bernardino).

This calculation is using an elasticity of **1.0**.

The example illustrates a more general phenomenon, which has in turn generated controversy. Because many California urban areas already have tremendous amounts of VMT, the calculator suggests that even modest increases in lane mileage will induce large amounts of driving. In this case, adding 20 lane miles of freeway creates an additional 100 million VMT per year. Mitigating that much VMT would be costly. These potential costs have led to concern, on the part of local agencies and other project proponents, that the proposed new lanes will no longer pencil out.

The result has been a robust debate about the state's implementation of SB 743. Critics of the implementation have advanced three arguments, which are not mutually exclusive. First, that the unit elasticity estimate in the NCST calculator is simply wrong, and thus inappropriate to use for individual road projects of any sort. Second, that the unit elasticity estimate is inappropriate for HOT lanes in particular, because HOT lanes differ substantially from free general purpose lanes. And third, that treating VMT as the negative impact of transportation—rather than Greenhouse Gas emissions (GHGs) or some other metric of harm—is inappropriate.

The debate has been carried out in various forums: in written arguments defending or criticizing the use of the elasticity (Fehr and Peers 2019; Volker and Handy 2019; Milam et al 2022; Systems Metric Group 2023) and in various panels and discussions convened by different stakeholders. In 2020, for instance, Caltrans assembled an expert panel and asked it to review different approaches for estimating induced travel. That panel, in its findings (Deakin 2020), arrived at some high-level agreement, but much of the agreement was aspirational. The panel agreed, for instance, that it would help if transportation demand models improved, if more and better data became available, and so on. These conclusions are hard to contest, but they also steer mostly around the *current* argument: what should we do now, given that our existing data, methods and models are imperfect?

It was against this backdrop that SCAG, in 2023, assembled its own expert panel, and charged it with reviewing whether an elasticity of 1 was appropriate for evaluating HOT lanes. The panel was a combination of academics and transportation practitioners, and it met twice, although not every member attended both meetings. I facilitated the panel, and between and after the meetings I circulated comments and writeups, and solicited further input from the members. The hope was that this panel would progress beyond the groundwork laid in earlier discussions.

That was not to be the case. The obstacle, in part, was the nature of the question. Without empirical research to examine, the panelists needed to rely on theory and conjecture. This got us only so far. As discussed above, one can, from first principles, mount a superficially plausible case in either direction. This created a situation that limited how much progress could be made. A larger problem, however, was that consensus overall was hard to come by. A number of panel members were simply skeptical of the unit elasticity finding, not just for HOT lanes but for any road expansion project. As one panelist noted in a written comment, "The elasticity in the NCST calculator is not correct and should not be generally applied."

It was this area, in fact—the overall validity of the unit elasticity estimate—where *most* of the disagreement occurred. There was relatively little comment about the correct elasticity for HOT lanes (again, perhaps, because of the dearth of data), but ample discussion and disagreement about the NCST calculator. The panelists skeptical of it suggested that regression-derived elasticities could not account for the fact that VMT has been falling in many places. They worried that the NCST was insufficiently sensitive to local conditions, and contended that Travel Demand Models should play a larger role in estimating induced VMT. They observed that an older research literature about induced travel—not used in creating the NCST calculator—often found smaller induced travel effects, and wondered why those findings did not inform the NCST estimates.

All this disagreement, it should be noted, broke down along stark disciplinary lines. The academics were more receptive to the academic research that produced the unit elasticity estimate, and practitioners were more skeptical of it.

Fewer panelists weighed in on whether the VMT was the appropriate metric for the state to focus on—this was not in the panel's explicit charge—but the topic did come up, and a number of panelists expressed reservations using VMT as a measure of impact. This skepticism, unlike the skepticism toward the unit elasticity, did not break down along disciplinary backgrounds. The people uneasy about considering VMT a pure cost included both academics who supported the unit elasticity estimate (including one panelist who has been instrumental in producing it) and practitioners who opposed it. These panelists variously observed that VMT had benefits alongside its costs, and that it was neither a measure of welfare nor of transportation system performance.

One point where panelists *did* generally agree related to the source of the current controversy: the trouble was arising because local agencies wanted to build new capacity and manage it, and that problem could be avoided if those agencies could instead legally convert existing lanes to managed lanes. I polled participants on this question at the end of the second session and all present concurred. I double-checked that result in a follow-up document circulated by email to all panelists, and no one contested it.

*

To summarize: the panel's main takeaways mirror the larger debate that prompted the panel's formation. This result is unsatisfying; we risk leaving off in the same place we started. To avoid that outcome, in the remainder of this report I will try to adjudicate some of the areas of disagreement. I will do so by reviewing the relevant literature, and highlighting areas where it may fall on one side or the other of the argument.

I am under no illusion that this exercise will suddenly deliver consensus. Nevertheless, I believe it is valuable. At least some of the conflict I observed stemmed, in my opinion, from misunderstandings about what the research literature on induced VMT actually says. In other instances I think the panel's conversation stalled out because it remained overly abstract, and it remained overly abstract for lack of a shared understanding of terms, assumptions, and methods. Thus one goal I hope to accomplish—which I think would ameliorate both these problems—is to simply explain the induced VMT literature with enough depth to address areas of contention, but also in the plainest terms possible, in a manner accessible to academics, modelers, practitioners and interested laypeople. Obviously not everyone will agree with my choice of literature, or my choice of assumptions. Even with such disagreement, however, this exercise could still get the argument to a more specific place. Doing so may, if nothing else, clarify the areas of contention, and help future panels and stakeholders make more progress.

The structure of the report is as follows. Section I reviews the concept of induced VMT, and then focuses on the topic where the panel spent the most time and had the most disagreement—the unit elasticity estimate. Is it correct for any roads at all? This section highlights the challenges of measuring induced VMT, explains how the recent literature tries to overcome those challenges, and then reviews the main findings of that literature. From there I turn a skeptical eye to those findings, and consider arguments against them.

Section II is shorter, and turns to the more specific question of estimating induced VMT for HOT lanes. A HOT lane is a combination of a carpool lane and dynamically tolled lane, so I approach this section by considering how both those components—the carpool restriction and the price—might change induced VMT. Section III turns to the ancillary yet important question of whether VMT, induced or not, is an appropriate measure of negative transportation impacts.

I emphasize that while the opinions that follow were informed immensely by the insights of the panelists, they are my own. I draw three broad conclusions:

**1) An elasticity of 1 is appropriate for general purpose lanes.** In my judgment, abandoning the unit elasticity estimate requires both a compelling case that it is flawed and the presence of an alternative that is obviously better. Right now neither of those criteria are met. To be clear, I understand why some stakeholders are wary of the NCST calculator. VMT in general are hard to measure, and induced VMT are impossible to directly observe. Any tool we employ will, as a result, yield an imperfect estimate, and one that will be impossible to verify with certainty. It's also worth noting that the NCST calculator is built upon an academic literature that was not designed to measure VMT per se. VMT in this literature was instead a proxy for traffic congestion; congestion is the problem that research is concerned with, and the motivation behind its analyses. The state of California could, if nothing else, be more transparent about this fact.

More broadly, I understand why someone might doubt the ability of a single statistical relationship to describe the impact of road expansions large and small in different contexts. Can it really be the case that every increment of new road capacity, in every circumstance, prompts an equally proportional increase in vehicle travel?

On one level, the answer is probably no, simply because every statistical distribution has tails. A sample where the average elasticity is 1 will feature some observations that depart from that average, and where VMT is more or less responsive to new road supply. Suggesting otherwise would be silly.

The fact of this variance, however, does not by itself justify discarding the average. Public policy routinely relies on average treatment effects, and departs from them only when specific compelling reasons exist for doing so, not simply because the average has a variance. The COVID vaccines, for example, had an average treatment effect, and that average effect was used to justify both their approval by the Food and Drug Administration, and the vaccination mandates many institutions issued for their employees. It has always been understood that this average treatment effect has a variance around it, and that the vaccine trials, while thorough, did not capture the world's full range of individuals and their circumstances. Some people will receive more or less protection from the vaccine than the average effect, and for some people the vaccine will be harmful. Few people, however, would endorse refusing to approve or take the vaccine on no grounds other than the documented effect being a mere average. Declining to use vaccine can in some cases be sensible, but a persuasive argument would not merely assert that some people would react to the vaccine differently. It would instead enumerate specific reasons *why* the average treatment effect was wrong, or did not apply in a particular circumstance (such as underlying medical conditions).

To bring the analogy back, it is not enough to note that there will be cases where the unit elasticity is incorrect. There must be specific criticisms that demonstrate, or at least strongly suggest, that the elasticity is *systematically* flawed. The specific criticisms most commonly levied against the unit elasticity estimate, however, do not demonstrate as much. Often these criticisms are not convincing. Those that *are* convincing, moreover (such as the potential unreliability of VMT data) apply not just to regression-derived elasticities but to other methods of calculating induced VMT as well.

Some criticisms of the unit elasticity seem to reflect a misunderstanding of the regression analyses that created it. There are skeptics, for example, who say that these analyses fail to control for variables like topography or land use. Such assertions are simply wrong. It isn't clear how these mistaken statements arise. They may stem from a failure to closely read the literature, or perhaps from an unfamiliarity with concepts like fixed effects. Other skeptics contend that the elasticity estimates should be revised because VMT overall has been falling. This argument is also misguided. It fails to distinguish between two different concepts: the total demand for driving, and the elasticity of demand for driving with respect to lane mileage. A change in one does not automatically imply a change in the other.

In still other instances, critics find ammunition against the unit elasticity estimate by mining an earlier generation of induced travel research. These older papers, often written by prominent and accomplished transportation scholars, did indeed find lower elasticities. But these papers also share some combination of two problems: they do not adequately control for the endogeneity of new road supply, or they estimate corridor-specific rather than regional elasticities (i.e., they measure how much VMT was induced only on the new road itself, rather than across the entire road network). Together these problems render the older papers unhelpful for the current debate. I try in this report to make clear, especially for people whose backgrounds may not be in econometrics, why the newer approaches to controlling for endogeneity matter, and why they cast doubt on the older results. I also, for the sake of thoroughness, take two of the older studies that skeptics cite frequently and examine them in detail. I show that these papers are products of a different time, and their methods are such that we should not use them as guidance in today's debates.[4]

In making these points, I am *not* suggesting that the unit elasticity estimate is an ironclad law of nature. Research on induced travel will continue, and the elasticity estimates may well change. As I discuss, it is possible that the changing nature of capacity expansion—it is now more common to widen existing routes than it is to build entirely new ones—will yield different elasticity estimates. It is also possible that greater Internet connectivity will not just make people drive less, but also make drivers less responsive to new road supply (these are, again, different concepts). Right now, however, those ideas, however plausible they may seem, are untested.

At this point a skeptic could take a different tack, and contest the implied onus of the argument thus far. As one panelist observed, it may be unfair to place the burden of proof on the NCST calculator's critics. Why should we assume the calculator is correct and demand specific evidence of its flaws, rather than, for example, assuming that Travel Demand Models are correct and requiring elasticity proponents to find fault with them?

One way to respond to this objection is pragmatically. The state has chosen the NCST calculator, so in point of fact the burden of proof *does* lie with its critics. That response is true, but also an appeal to authority, and like any such appeal it is intellectually unsatisfying. No idea should carry special weight simply because a powerful institution espouses it.

A more persuasive response, however, is that even if we *did* reverse the burden of proof, we most likely wouldn't change the outcome. Suppose we start by assuming we should rely on Travel Demand Models, and agree to drop this assumption only if we find reason to think TDMs err systematically in estimating induced VMT. Would we find such evidence?

We would. Most discussion of how TDMs handle induced travel—and this includes comments from panelists who work with TDMs—acknowledges that TDMs will tend to underestimate it. This underestimate arises, moreover, for specific, persuasive reasons related to how the models manage (or fail to manage) feedback effects. Thus where it is possible that in any given situation an induced travel estimate derived from a regression will be in error, it is *plausible* that in any given situation an estimate derived from a TDM will be too low. This higher probability of error should lead us, all else equal, to favor other approaches.[5]

A TDM proponent could respond by noting that modelers have methods for addressing the underestimation problem. These methods generally involve adjustments that are made to model outputs, based on some combination of results from a second model, the judgment of an expert panel, or the judgment of the modelers themselves. While such post hoc manipulation of model results may sometimes be unavoidable, I think almost any researcher would hesitate to call it desirable. A model, ideally, imposes discipline on researchers. Running a model tests an idea. Creating an environment where models can be altered during or after the analysis, and in a variety of ways that may be hard to document or quantify, at best risks the credibility of the modeling discipline, and at worst opens the door to manipulation. These risks could be mitigated if TDM practice was characterized by a culture of transparency and replication, but right now it is not. Econometric practice offers far more transparency than TDM practice.

---

[4] The deficiencies in these earlier studies are in no way a reflection on their authors. These papers were written when the statistical state of the art was less advanced, and in particular when access to the historical data needed to create high quality instrumental variables was much lower.

[5] In full disclosure, while neither side would likely claim me as a team member, I am much closer to an econometrician than a transportation modeler, and readers should bear that in mind.

TDMs, like econometric methods, do steadily improve, and if new versions are better able to incorporate feedback and capture induced travel, they may well come to supplant the NCST calculator. “I expect the panel would all agree,” as one member said in a written comment, “that the best solution is to improve the models so that they incorporate all of the feedback necessary to accurately estimate induced VMT. But we don’t have those models.” I concur. Given the current state of the practice, calls to abandon regression-derived elasticities because they *may* be wrong, and that leave unmentioned the complications that may arise from off-model adjustments to TDMs, can come across as “isolated demands for rigor”[6]—skepticism selectively applied.

**2) HOT Lanes differ from free lanes in many ways, but probably do not differ in the quantity of regional VMT they induce. For estimating the regional VMT impacts of managed lanes, therefore, an elasticity of 1 is probably appropriate.** Numerous panelists, when discussing how managed lanes might induce VMT, said something to the effect that it just *seemed* like the elasticity should be lower than free lanes. One panelist, after weighing how drivers might trade off between priced and free lanes, and then observing that little is really known about this tradeoff, concluded that he saw little reason “to assume the elasticity is 1.”

I shared this sentiment. While a HOT lane is physically no different from a general purpose lane—asphalt is asphalt—it is *economically* quite distinct. A free road is a commons: a rival but non-excludable good that is intrinsically prone to overuse, congestion and depletion. A HOT lane, in contrast, is rival but excludable, and that excludability can protect it from the problems that afflict a commons.

A HOT lane is, for that reason, a different product than a general purpose lane. The magnitude of that difference depends in part on how, and how well, the lanes are managed (managed lanes can differ substantially from each other). Assuming a HOT lane is well-enforced and designed to maintain throughput, however, it will perform better than general purpose lanes. It will not fall into severe congestion, it will offer more reliable travel times, and the vehicles traveling on it will, as a result, emit less pollution than vehicles on comparable free lanes. Because it is priced, moreover, a newly-built HOT lane may be less likely than a newly-built free lane to carry low-value trips, and more likely to attract people who were already on the road. Someone willing to pay for the reliability the HOT lane offers is probably making a trip that is important to them, and that trip is unlikely to have been induced by the existence of a HOT lane. A new HOT lane, as an example, is more likely to carry a person en route to the airport for a business trip, than someone impulsively deciding to leave the couch because they want peanut butter.

All these differences between HOT lanes and free lanes, however, do *not* suggest that HOT lanes have a different elasticity of induced VMT. This perhaps surprising result arises because the scale at which these roads differ is not the relevant scale for the analysis. The characteristics that make HOT lanes different are specific to the HOT lane corridor itself. The HOT lanes carry higher value trips, have less congestion and probably less pollution. But the state is concerned with how a newly built facility changes *regional* VMT, which means not just the any new driving that occurs on the HOT lane, but also new driving that occurs on other, pre-existing roads *because* the HOT lane was built.

When we expand the analysis to the entire network, HOT lanes and general purpose lanes, at least on the specific question of induced VMT, start to look more alike. Consider first that at a corridor level, a well-managed HOT lane will probably carry *more* VMT than a similar unmanaged lane. Unmanaged lanes regularly break down as a result of heavy congestion, and at peak hours they move substantially fewer vehicles than at other times. This phenomenon is called hypercongestion, and it inflicts costs akin to removing road space. HOT lanes, because they are priced, can prevent hypercongestion, and as such can increase corridor level VMT. This increase in VMT is a sign of improved system performance—it shows that the HOT lane is working.

---

[6] This phrase was first coined by Scott Alexander (2014).

Because the HOT lane carries higher-value travel, however, much of this increased corridor-level VMT may not be *induced* VMT. It will instead represent redistributed VMT. Many vehicles using newly-built HOT lanes would, in the absence of the HOT lane, still be on the road, and perhaps on the free lanes nearby.

This prevalence of redistributed VMT might suggest a lower elasticity for HOT lanes. The catch lies in the fact of construction: when the government builds an entirely new lane to manage it (as opposed to converting an existing lane from general purpose to managed) the traffic that redistributes onto that new lane will open up space on existing unpriced heavily-traveled roads. Opening up this space will have the same impact, economically, as building new general purpose capacity. It will induce new travel on those free lanes. And as discussed above, the best estimate we have for the magnitude of that induced travel on an unpriced space is an elasticity of 1. Right now, then, the most appropriate *regional* elasticity for HOT lanes is probably 1.

Are there cases where that might not hold? Possibly. As one panelist noted, a managed lane could be managed to explicitly reduce VMT. It isn't clear, however, what this would look like, or if in its final form this facility would meaningfully accomplish other transportation goals (it might, for instance, be priced prohibitively high).

Two further remarks are warranted here. First, it's important to reiterate that this area is largely devoid of empirical research. Hopefully that will change. If nothing else, it is probably worthwhile to try and re-estimate induced VMT regressions that explicitly control for HOT lane mileage. There is very little of this mileage relative to the road network, so there is a risk that any differential effect HOT lanes have would be hard to detect, but it is probably worth trying.

Second, as the panel observed, the problem that confronts HOT lanes under the state's induced travel regulations only arises if the HOT lanes are new capacity. This regulatory obstacle to new HOT lanes would largely vanish if the HOT lane was implemented by converting existing free lanes to management. I realize that converting existing lanes is not just politically difficult but may also face administrative hurdles. The state of California has taken steps in recent years to make tolling easier, but the legal path to pricing a currently-free lane remains unclear, and federal rules that apply to federally-funded lanes also make it difficult to toll existing free lanes without adding new untolled capacity. Arguably the impasse around pricing and expansion would be easier to resolve if the legality of tolling existing general purpose lanes was clearer.

Pricing existing capacity is indisputably more efficient than building new capacity to price it. The benefits of a HOT lane come from the faster travel it offers compared to unmanaged free lanes. Those lanes, when they are slower, are slower because they are congested. It thus makes more sense to use prices to remedy the congestion in the free lanes, than it does to build entirely new lanes whose appeal will rely on leaving the problems on the rest of the road unaddressed. To wit: if an old house has three bathrooms with broken plumbing, the optimal solution is to fix that plumbing, not construct a fourth bathroom where the plumbing works and then selectively sell access to it.

**3) VMT is not pollution and classifying it as such makes little sense. At the same time, employing more appropriate metrics for transportation's external costs would probably make it harder, not easier, to construct new road supply.** More than one panelist observed that VMT is neither a good measure of welfare nor a good measure of transportation system performance. A road carrying large amounts of VMT could be working poorly or well, and could be improving or degrading the lives of people near it. VMT can fall because a region builds a high-quality transit network that is superior to driving for many trips, and it can fall because a region plunges into recession that throws many people out of work. One of these scenario implies an improvement in people's well-being, and the other implies immiseration. The mere fact of falling VMT cannot help us differentiate between the two.

I acknowledge that classifying VMT as a cost has practical and intuitive appeal. The state and federal governments have longstanding (if imperfect) ways of measuring VMT, and analysts as a result have decades of VMT data in easily-accessible form. This is not the case for GHGs and many other pollutants. It is also the case that for many development projects, VMT is almost certainly a better metric of externalities than Level of Service, which is the measure it replaced. Perhaps most important, VMT *is* strongly correlated with many negative transportation impacts, and in practice it will be difficult for California to meet its climate goals without reducing VMT. In the context of conversations about pollution and decarbonization, a statement like “we need to reduce VMT” is a perfectly reasonable shorthand. Most people understand the meaning behind it.

The problem arises because *formalizing* VMT as an impact is quite different from *understanding* it as a reasonable *proxy* for impacts. Here we arrive at Goodhart’s Law: a measure that becomes a target ceases to be a good measure. Formally defining VMT as a pure cost ignores the fact VMT, while correlated with many bad outcomes, is also correlated with many good outcomes, which at least suggests it has some benefits. One could argue that in urban California the costs should demand more attention, because additional VMT do more damage in places where VMT are already high (e.g., a car pulling onto a congested road imposes more delay than one pulling onto an empty road). But places with lots of existing VMT tend to be more heavily oriented around driving, and in those places the benefits of VMT can also be inordinately large, precisely because such places offer so little mobility and access in other ways. In most of California, a low-income person acquiring an automobile for the first time will increase total VMT, but many of the miles that person drives will have total benefits that far exceed their costs.

To make the same point in a more general way: the costs and benefits of a given vehicle mile of travel can vary tremendously depending on when and where the travel occurs, and who is in the vehicle. An electric vehicle carrying four people, driving on an uncongested freeway in the middle of the night, imposes far fewer social costs than an internal combustion vehicle carrying one person that traverses the same segment at rush hour. A low-income person using that freeway segment to access a job that keeps her household barely afloat, similarly, implies many more benefits (for both the traveler and society) than a high-income person cruising along for the sheer thrill of motion like a latter day Mariah Wyeth. Yet total VMT, in all these very different scenarios, is the same.

One could argue that I am highlighting unusual events and exceptions. Low-income drivers and nighttime electric trips do not describe *most* new VMT, and so the average new mile driven will have more costs than benefits. That may be true. Here, however, is a case where there are compelling reasons to question reliance on an average, and better alternatives as well. It is possible, after all, to use policies that separate VMT’s costs from its benefits. Taxes on gasoline and prices on congestion can do so. Adopting or strengthening such policies would better align state policy with the canonical approach to externalities, which is to preserve the benefits of productive activity while minimizing its harmful byproducts. Taxes and prices are admittedly unpopular, but the fact of their unpopularity does not turn VMT into a pure cost.

Here’s a different way to think about the issue. Suppose the argument above is wrong, and all new VMT is a social cost that warrants mitigation. If that’s so, the state should look less favorably on a proposal to build two new freeway lanes and price them, and more favorably on a proposal to build those lanes and drop boulders across them.[7] While one can argue that every set of standards is bound at some point to yield counterintuitive results, the boulder scenario should at least give us pause.

A potential response, of course, is that no one would ever build a road just to drop boulders across it, and that as such this demonstrates the *effectiveness* of calling a VMT a cost, because it stops governments from building roads. That reasoning, however, quietly changes the policy goal. We have gone from not wanting more VMT to not wanting more roads. If blocking new road capacity is the goal (and there are strong arguments, in most of California, in favor of such

---

[7] I borrow this imagery from Brian Taylor.

a goal), classifying VMT as a pure cost is a needlessly roundabout way to achieve it, and one that opens the door to circumvention. Local governments right now, for example, can expand roads if they sufficiently "mitigate" the induced VMT, but many of the actions that count as mitigations probably do not reduce VMT at all. Agencies are credited with reducing VMT if they invest in light rail or affordable housing. Little evidence, however, suggests that increases in light rail, in California or elsewhere, reduce VMT, and while building housing of any kind in dense, transit-adjacent areas can reduce VMT, it is market-rate housing, not affordable housing, that is likely to reduce VMT the most.[8] Obviously there are other strong justifications for transit and affordable housing investments, but tying them to road proposals could distort their allocation. The state could find itself in the undesirable situation of seeing more roads built, at greater time and expense, and also seeing its rail and housing resources expended not in ways that do the most good, but in ways that make the roads look most appealing. There are large risks involved in turning housing and rail into mitigations for roadways, rather than important amenities in their own right.

Suppose we took a different approach, and observed that on the margin, new automobile infrastructure often has costs that exceed its benefits, and that among those costs is a built environment that makes it harder to move around in non-driving ways. This sensible observation could give us a strong presumption against building new roads, but do so without tying ourselves in logical knots. After all, consider what it means when we say that new road construction often makes driving or walking harder. This statement implies that expanding auto infrastructure is costly because it increases the relative *benefits* of VMT. And that conclusion is intuitive. In auto dependent regions, residents who cannot drive suffer. Driving is less useful in Midtown Manhattan than in suburban Los Angeles, and being carless is easier in Manhattan than Los Angeles, for the simple reason that LA is built around driving and Manhattan is not. This fact alone should suggest that VMT, especially in places built for cars, has benefits. Those benefits might be conditional on past decisions we as a society made that we now wish we hadn't, but since we did make those decisions, the benefits are real.

If we find this situation troubling (I do) the policy remedy should involve Los Angeles changing its built environment, not Los Angeles using any means necessary to slow or reverse the growth of VMT. If Los Angeles confiscated every third automobile but kept adding freeways and parking lots, it would reduce VMT while impose huge costs on its residents. If, in contrast, it changed its built environment to become even a bit more like New York, it could reduce VMT while improving its residents' lives, since the different built environment would reduce driving by reducing the relative benefits of driving.

One panel member observed that the state could more efficiently advance its goals if, rather than focusing on total VMT (or even total induced VMT) it simply declared that no freeway could be expanded until a local agency demonstrated that all existing lanes of that freeway were operating at full capacity at peak hours. This criterion, of course, could only be met if those lanes were dynamically priced. Such pricing may well increase VMT in the short run, but it would also likely reduce pollution and GHGs, help the transition toward a landscape more hospitable to non-auto modes, and let the state deliver a better performing transportation system.

I support this idea, with the acknowledgment again that pricing existing lanes can be politically difficult. The more immediate point, however, is that the problems with classifying VMT as a pure cost, while real, are for our purposes academic. The controversy over SB 743 has arisen because some local agencies want to build new road capacity. The criticism they have levied at SB 743 has merit, but taking it seriously would make it harder, not easier, to increase the supply of freeways. The optimal way to implement SB 743 is through direct pricing of auto travel's externalities, not by constructing new capacity.

---

[8] For the simple reason that higher-income people drive more to begin with. See Chatman et al (2019).

# I. The VMT Effects of Adding Free Capacity

One of the oldest ideas in modern transportation planning is that adding to the supply of roads will, all else equal, encourage more driving. "A newly-opened or widened street," a Los Angeles transportation official observed in 1928, "immediately become[s] glutted by the access of cars that hitherto have reposed more in their garages than they utilized the streets." [9] That, in essence, is the idea of induced travel. More space attracts more vehicles.

A crucial piece of nuance is that new capacity by itself cannot generate VMT, any more than new housing by itself can attract residents. But in an area with unmet (sometimes called latent) demand for driving, new road capacity can generate additional travel by lowering the cost (in time, reliability or stress) of some trips. A wider road might encourage nearby households to make some trips by automobile rather than other modes. It could encourage households to drive somewhere rather than stay home. And it could encourage households to choose more distant destinations for some trips they already make by driving (e.g., going to a preferred grocery store or restaurant that is further away, rather than a merely acceptable one closer by). New road capacity can also generate VMT by opening up access to previously unreachable or hard-to-reach destinations. This new access could encourage households to relocate to outlying places that are further from jobs and other activities, leading to more daily driving.

These two mechanisms—making driving easier on existing, well-traveled routes and opening up new destinations—are related but distinct. Because they are distinct, researchers sometimes differentiate between induced *travel* (or induced VMT) and induced *demand*. The latter is a subset of the former; all induced demand is induced travel, but not all induced travel arises from induced demand.[10]

Today's literature on induced VMT is an outgrowth of Anthony Downs's (1962) original Fundamental Law of Peak Hour Congestion. Downs was concerned primarily with speed. He recognized that the typical anti-congestion policy was motivated by driver dissatisfaction, and that the source of that dissatisfaction was low speed caused by congestion. Drivers complained because they could not move as fast as they wanted to on a particular facility at the time they most wanted to be on that facility (the "moment of maximum convenience" as he put it). Put simply, they didn't like being in stop-and-go traffic on freeways at rush hour, and they complained. Transportation agencies often responded to those these complaints by expanding capacity.

Downs's point was that expanded capacity was unlikely to solve the problem. But his argument wasn't about induced VMT. Indeed, Downs observed that expanded capacity would be counterproductive even if it induced no new travel at all. Downs coined a phrase— "triple convergence" to describe how new road capacity would fill up at peak hours. The new capacity would, in the short run, increase travel speeds, but that speed would itself attract drivers, who would converge on the road at busy times from "other routes, other times and other modes." Triple convergence is elegant in its simplicity: it suggests that efforts to fight congestion with new capacity can be undermined purely through the redistribution of *existing* travel. This redistribution *could* involve increased VMT (e.g. if people stop riding a train and start driving) but did not have to, and given the low transit shares in many urban areas, it mostly would not. Thus while new VMT is mentioned in Downs's article, induced VMT is a corollary to his argument. The primary point was important but narrow: adding capacity to a freeway to *increase speed on that freeway in a peak period* wasn't going to work.

---

[9] Ladd (2012).

[10] There is no strict bright line between these two categories. Over time, widening existing roads in a region might make it easier to reach some outlying destinations.

Downs did not formalize or empirically test his Fundamental Law, but subsequent research has for the most part validated his argument: peak hour expansions tend to have short-lived impacts on peak hour speed. In the medium- to long-run, expanded roads are just as congested, and sometimes more congested, than they were before (e.g. Winston and Langer 2006; Kim 2022; and Metz 2023; Ossekina et al 2023).[11] The expanded road holds more vehicles, but they are not moving any faster.

Downs's article, however, was also an incomplete view of how capacity could impact congestion, both because it largely ignored longer-run land use impacts (e.g. if the new road encouraged sprawling development) and because it dealt only glancingly with the second-order effects of drivers reacting to the changing conditions on the expanded road. If new capacity attracted enough traffic to result in slower speeds than before, as Downs suggested it could, some of that resulting delay might be felt throughout the network, not just on the expanded capacity itself (Fosgerua 2015).[12]

Testing these predictions about new capacity was more complicated, in part because speed data have long been difficult to assemble for entire networks. VMT data, in contrast, are broadly available. VMT, moreover, is a defensible proxy for congestion (although as we will see, not a perfect proxy). After Downs, tests of how new capacity affects congestion across a network began using VMT as a dependent variable. This gave rise to the literature on induced VMT.

## 1.1 The Conceptual Challenge of Measuring Induced VMT

Assuming some demand exists, the extent to which new capacity induces new VMT will depend on two factors: the relevant time horizon, and the nature of the new capacity. The time horizon is straightforward: induced travel effects are larger over the longer-term (usually defined as 10-20 years). This is so largely because demand induced by new access (i.e. changes in land use) takes longer to manifest than changes in existing travel (Volker and Handy 2022). Widening a road to the exurbs, for example, might initially just increase the number of people driving between existing origins and destinations on that road. As time passes, however, the widening might encourage developers to build housing near the roadway's exurban segments. As that new housing gets built and occupied, driving will continue to increase.

The relationship between induced VMT and the type of new capacity is more complex. New capacity can take multiple forms. New roads that open access to greenfields are not the same as new roads that connect existing roads within a developed area, and both in turn are different from expansions that add new lanes to existing roads. That last category—road widening—is probably the most common form of new road capacity, and within it we can draw a further distinction: between adding lanes to one-lane roads, and adding lanes to roads that already have two lanes

---

[11] The specifics of these finding tend to vary. Winston and Langer show that a dollar of federal road spending reduces congestion costs by only 11 cents. Kim examines roadwork (a proxy for road expansions) in urban California and finds that speeds increase at first and then revert to pre-expansion levels after about a year. Metz examines a highway expansion in England and finds a similar result. He also observes that convergence may be faster now than in years past as a result of satellite navigation programs. In Downs's initial formulation, convergence was driven by a minority of optimizing drivers he called "explorers" while most drivers, content with their existing routes, were "sheep." The explorers were the drivers who realized the expanded route was temporarily faster and switched. The sheep kept to their original routine. Metz's logic suggests that satellite navigation has turned more people optimizing explorers, and that this accelerates convergence. Finally, Ossekina et al is a bit of an outlier; they examine highway widenings between cities in the Netherlands and find that speed gains persist for six years. The resulting gain in welfare, they say, covers about 40 percent of the highways' fiscal cost.

[12] Hanna et al (2017) offer a real world example of this, which occurred when Jakarta ended its HOV restriction policy. The roads previously subject to the policy attracted enough new vehicles to send the total volume handled by the road *down*, and this forced traffic onto other roads, worsening speeds throughout the city.

or more. In every case, two conceptual challenges loom. First, induced VMT cannot be directly observed. Second, new road capacity is likely to induce some VMT on *other* roads. We will introduce these challenges as we examine the different new capacity scenarios.

### 1.1.1 VMT Induced by New Routes

We can first consider a capacity expansion that increases the overall coverage of the road network. If a government builds a new road to a greenfield, the road could, over time, encourage development in that greenfield.[13] If this development occurs, and if it is composed more of homes than jobs (which is usually the case, because jobs disperse more slowly than homes)[14] or if the people who move to the greenfield do not work in or near it, then the road, by virtue of the access it delivers, could induce new vehicle travel. Residents will commute by car to their jobs, and drive to and from their homes to perform other tasks.[15] Much of this new VMT could be described as both induced travel and induced demand: it arises because the coverage of the road network has grown, offering access to more destinations.

Here we arrive at the first conceptual challenge: induced VMT cannot be directly observed. On the road, vehicles producing induced VMT look identical to all other vehicles. There is no way to look at any given car and determine if an expanded road network caused the trip it is making. For this reason, studies of induced travel must be indirect: its presence and magnitude can only be inferred.[16]

This inference, moreover, is confounded by an obvious counterfactual: had the new road capacity not been supplied, *some* driving would still have occurred. Suppose the greenfield hadn't been developed. The people living in it would have lived somewhere else. As long as they would have driven some nonzero amount of miles in this alternative location, we cannot call all the VMT observed in the greenfield induced. The induced VMT would instead be an increment: the difference between the driving that occurs in the greenfield and the driving that *would have occurred* in the counterfactual locations. Induced VMT, by this logic, will be higher if the greenfield development is full of people who would have otherwise lived in job-adjacent urban apartments, and lower in the residents would have otherwise lived in low-density exurban areas.

Unfortunately, it is almost impossible to know where people *would have* lived. Consumer reference data can sometimes show where residents of a new development lived previously, but even this information is of limited use. We cannot assume that in the development's absence its residents would have stayed put. The mere fact that people moved suggests a certain level of dissatisfaction with their previous address, raising the probability that they would have moved somewhere. Without knowing why people moved, it becomes difficult to know where they might have gone.

*

Now consider a case where a government builds a new road between two existing busy roads. This new road, like the road to the greenfield, could encourage new development. The challenge of estimating VMT induced from this road

---

[13] Strictly speaking, this could be an entirely new road or an extension of an existing road.

[14] E.g., Baum-Snow, 2010.

[15] Residents will be particularly likely to drive for non-work tasks if the greenfield development is laid out in a manner unconducive to non-auto modes (most greenfield developments are).

[16] Estimating induced travel is similar to estimating the value that an additional bedroom or bathroom adds to a house. We will never know the "true" answer to these questions, because we cannot observe a market transaction: no one ever just buys a bedroom. We can only infer the answer, by analyzing many house sales that feature different numbers of bedrooms, and that allow us to control for many other aspects of the home that might influence price.

would be similar to the example above, but come with an additional complication. A new road connecting two busy roads is more likely, relative to a road into a greenfield, to concentrate rather than disperse activity.

For example: suppose two roads run parallel to each other, and a cluster of related firms is located along them. The nature of the cluster is immaterial: the firms could be involved in software, media, light or heavy industry, whatever. If most parcels along these roads are already developed and occupied, then additional firms that would benefit from proximity to the cluster, because they share a customer base or labor pool with the firms there, may be unable to do so. If the government builds a road that connects the two existing roads, that new road could open up more parcels for development. The new road and the development it enabled would, as a matter of simple math, add VMT to the area. But it would not be obvious that all the new VMT in the area would represent new VMT overall. Some probably would, but if the new road let additional firms join the cluster, *and* if in the absence of the road many of those firms would have located further away but nevertheless regularly sent buyers or salespeople to the cluster, then some of the VMT on the new road would represent an increase in local driving, but *not* an increase, and possible even a decrease, in *total* driving.[17]

### 1.1.2 VMT from New Space on Existing Routes

The induced travel effects discussed above are primarily “access” or “coverage” effects: an entirely new road makes new destinations available, and those destinations increase the demand for travel. In most of the urban US, however, constructing entirely new roads is now rare. More common are expansions that add capacity to existing roads—e.g., adding lanes in each direction to freeway segments. Such expansions do not increase the coverage of the road network, and are therefore less likely to create VMT through new access to destinations (indeed, adding a lane to an urban freeway often results in some destinations being lost, if expansion requires that homes or buildings are taken by eminent domain). Putting a third lane on a two lane road into the exurbs does not “open up” the exurbs the same way a brand new road does.

Widening an existing road will induce VMT through a different path: it will allow more vehicles to use that road at any given time, and in doing allow some trips to be completed more quickly, more reliably, or with greater real-or-perceived safety.[18] Here again, however, not all vehicles traveling on the expanded capacity will represent a regional increase in VMT. There will be some people drawn to the expanded road who in the road’s absence wouldn’t have traveled at all, or would have traveled by a non-auto mode. Others, however, may still have driven, but done so at other times or via other routes. The traffic on the expanded facility will thus represent a mix of new driving and redistributed driving.

---

[17] Taking this point further: there are situations where new capacity can *reduce* total VMT. This most commonly occurs when the new capacity is a bridge or tunnel that provides a more direct connection between two destinations that would otherwise demand circuitous travel. Imagine a situation where two population centers are on opposite sides of a river, but the nearest river crossing is ten miles away. Constructing a bridge directly connecting those two population centers *could* reduce total VMT. (It would not automatically do so. Anyone traveling between the two towns would now drive fewer miles to do so, but the ease of the trip might also encourage more people to make the drive. So long as the reduction in miles associate with the former outweighed the increase associated with the latter, the bridge would reduce VMT).

[18] Saying that new capacity allows some trips to be completed more quickly is *not* the same as saying that traffic on the road will move faster. As Downs (1962) and others have argued, it is possible for an expanded freeway to have *lower* average speeds at peak hours after the expansion, but nevertheless offer *faster* average trips, if the freeway provides a more direct route between two destinations than a traveler’s next best route. Also see Arnott and Small (1994).

Now we come to the second conceptual challenge in measuring induced VMT. Even some of the redistributed driving could, as a second-order effect, induce some new VMT. Motorists who leave a busy arterial to drive on a newly-expanded freeway open up space on that arterial. That new space can itself attract new VMT, because it will make travel on the arterial faster. Some of the VMT this space attracts will be redistributed, but some of it may also be new, and crucially the portion that is new is caused by the expansion of the freeway. This process, of vehicles backfilling in as other vehicles migrate to faster routes, is unlikely to continue *ad infinitum* (e.g. traffic that leaves a collector street to drive along a newly-faster arterial might be replaced, but traffic that leaves a sleepy residential street to drive along the collector may not be). But the process will occur, and it suggests that some (perhaps most) of the VMT induced by an expanded facility will occur *away* from the expanded facility itself.

A somewhat special case involves going from one lane to two lanes or more. Expansions of this sort are more likely to have access effects because they feature increasing returns to scale: although the road segment's capacity doubles, its performance—in travel time or reliability—could more than double. These increasing returns are possible because a single slow moving or disabled vehicle can bottleneck a one-lane road, triggering substantial delay even in the absence of conventional congestion. A second lane dramatically reduces this potential problem, simply by allowing other vehicles to pull around.[19] In these situations the VMT response to a new lane could be particularly strong, and such expansions could have access effects more typically associated with entirely new roads.[20]

## 1.2 The Empirical Challenge of Measuring Induced Travel

To summarize the discussion thus far: new capacity will attract driving on that capacity. Some of that driving will be "new" (e.g., those miles would not have been driven, but for the road) and some will be redistributed (the miles would have been driven, but the VMT would have occurred in another place or at another time). The redistributed driving, moreover, will create space elsewhere on the network, or on the same road at other times, and this space could in turn attract new driving, if it meaningfully increases speed or convenience along those corridors.

For reasons we have discussed, isolating these different effects is difficult. In the academic literature, the most common way to estimate induced VMT is via a regression analysis at the regional level. Researchers compile large amounts of data on road supply and VMT across many regions, and then try to correlate changes in road supply over an entire metropolitan area with changes in VMT.

Broadly speaking, this methodological approach faces three empirical obstacles: the problems of VMT data, the many factors other than new road capacity that cause driving, and the nonrandom nature of road expansion. We will discuss each in turn.

---

[19] To a limited extent, turnouts and occasional passing lanes—extremely small doses of added capacity—also address this problem.

[20] As an example, consider a mountain town with many desirable natural amenities that is only reachable via a road that is one lane in either direction. The road's dimensions, and their implications for travel time, reliability and safety, could function as a constraint on development in the mountain town. A second lane, by changing the perception and reality of travel to the town, could encourage more vacation homes and hotels, and thus more travel overall.

### 1.2.1 Challenge 1 - VMT Data

There are examples, in the world of applied economics, of statistical analyses that enjoy near-perfect data. Research on the frequency of stock trades, or the timing and outcomes of Congressional votes, or the incidence of home runs in baseball, all benefit from data that is complete and largely devoid of errors. Studies of induced VMT do not fall into this category. Most of the nation's VMT go unobserved. Quantities of driving, as a result, are always estimates. The federal government estimates VMT annually, using a combination of data of gasoline consumption and vehicle counts submitted by each state. These estimates are made available to the public via the Federal Highway Administration's Highway Performance Monitoring System (HPMS).

The HPMS is a remarkable resource, but it has its weaknesses. Approaches to counting traffic have varied over time, and at any given time they also vary by place. While the federal government gives the states some guidelines on how to estimate VMT, it also grants them considerable latitude in how they do so, and states as a result often estimate VMT in different ways (Hymel 2014). Some states, like California, have more, and more sophisticated, counting equipment than others. California's freeway Performance Measuring System (PeMs) monitors vehicle travel on the state's freeways via magnetic loop detectors embedded in the roadway. Other states do not deploy counters of similar quality, or to the same extent. They may also train their staff and validate their equipment in different ways (Fefke et al 2004; Hymel 2014). Even in California, vehicles are missed as loop detectors come on- and off-line (e.g. Steimetz and Brownstone 2005). HPMS estimates of VMT, as a result, are not the product of a uniform counting method. They are the aggregation of many counts made using slightly different methodologies.

The HPMS is not the only source of VMT data. Driving quantities can also be estimated in travel diaries like the National Household Travel Survey (NHTS). Travel diaries ask a representative sample of households about their travel behavior, and may also ask respondents to report the odometer readings on their vehicles. In combination the answers to these questions can be used to estimate total household VMT, which can then be extrapolated to regional or national VMT. Where regression analyses tend to rely on HPMS or similar data, Transportation Demand Models often incorporate travel diary data.

Travel diaries offer a more fine-grained perspective on travel behavior, and allow researchers to tie travel decisions directly to household characteristics (e.g. higher income people drive more miles). This ability to connect travel behavior to other characteristics is one reason travel diaries can be useful components in building travel demand models. But travel diaries have their own limitations, particularly if the goal is to measure induced VMT. Travel diary surveys often have low response rates, and their smaller sample sizes make it hard to draw conclusions about all but the largest geographies. They are also carried out less frequently, and because they focus on household travel they do not capture commercial VMT, which is a substantial share of total driving. (For example, the switch to e-commerce reduced household VMT but increased commercial VMT). In the NHTS, moreover, both the sampling reliability and the approach to estimating VMT have varied over time, in ways that call into question both the survey's representativeness and the accuracy of its VMT estimates. (Manville et al 2017; Alberini et al 2021).[21]

---

[21] In 2001 and 2009, the NHTS was carried out by calling only landline phones, meaning it probably undersampled a nonrandom 10 to 25 percent of US households who did not have a landline (Manville et al 2017). Alberini et al (2021) examine the process by which NHTS staff blend information from the survey to create its two main measures of VMT. Some aspects of this process are opaque and questionable. In the 2009 and 2017 NHTS the VMT estimates were based in part on previous iterations of the survey—suggesting that the VMT measures are not independent of each other over time. Similarly, odometer readings were also multiplied by an artificially generated variable intended to represent a "universal driving schedule."

A final, newer possibility for measuring VMT comes from GPS data, gathered either from mobile phones in vehicles or from onboard GPS devices. Firms like INRIX and Streetlight have pioneered the collection of such data, and now regularly use them to issue reports and/or sell data products to cities and consultants. At present, however, these sources are probably not viable for measuring induced VMT, for two reasons. First is simply that they are relatively new. Estimating a long run induced travel response requires at least ten years of data (or two data points ten years apart), and ideally more. Second is that the quality of the data has become somewhat uneven as time has passed, and the samples have arguably become less representative (e.g. Welch and Widita 2019; Turner et al 2020; Ulrich 2023). If these problems can be corrected, then GPS-based sources may well become important in the future.[22] For now, however, the HPMS, for all its imperfections, probably remains the best source for estimating longer-term the induced travel effects.

### 1.2.2. Challenge 2 - The Many Sources of VMT

Factors other than the supply of roads affect the total amount of VMT in a region. Gas prices, employment levels, and population growth all help determine how much people drive. So too does weather, topography, and the overall pattern of development. When researchers attempt to statistically measure induced VMT, they must control for these factors. Regressions that estimate induced travel normally include variables measuring the economy and population, and metropolitan attributes that might push destinations apart even independent of road construction (sprawl, topography, segregation, etc.)[23] The induced travel effect that emerges from the regression is the association between VMT and new lane mileage *after* controlling for all these confounding factors.

An important point is that many regression analyses do not explicitly control for these factors, but instead rely on place and time "fixed effects." Fixed effects are placeholder variables intended to absorb the influence of unobserved factors, particular to some place or time, that might affect the total amount of driving.

To see the intuition behind fixed effects, suppose we think that Los Angeles just has a different "driving culture" than the rest of the country, and that this culture affects VMT there. Or suppose we think an incident in a particular year made driving idiosyncratically more or less common (this incident could be something generic, like a recession, or something more unusual, like the September 11th attacks in 2001, which reduced air travel and increased long-distance car travel). Or suppose we just think weather and topography matter. These factors might be hard to measure by conventional means, raising the concern that we cannot control for them. So long as these factors are constant within a time or place, however (e.g. L.A.'s driving culture is unique to L.A., the fallout from the September 11th attacks occurred in 2001, San Francisco always has steeper hills than Minneapolis) then place and year fixed effects can control

---

[22] Prior to 2021, mobile phone users who did not wish to be tracked had to explicitly opt out, by going into the settings for individual applications. In late 2021, this changed; Apple began prompting iPhone users with alerts that apps were tracking their locations, and asking users if they wished to opt *in* to this tracking. Most chose not to. Apple's competitors soon followed, and the quantity of mobile phone data plummeted. In response, firms like Inrix pivoted to a greater reliance on connected vehicle data. Not all automobile manufacturers make such data commercially available however, and as a result the GPS sample is not now heavily drawn from GM and a few other brands. For obvious reasons this sample is unrepresentative: it skews toward newer vehicles and higher-end vehicles. (In addition to the sources listed in the text, I thank Sam Speroni and Fariba Siddiq for much of this insight).

[23] The regression models will sometimes include controls for socioeconomic status (SES), such as the area's poverty rates, median income, age distribution and so on. One argument for not including these controls is that new road capacity may attracts people and firms. If it does, then SES variables can be both a cause *and* effect of VMT, and including them could bias the regression (Duranton and Turner 2011).

for them even if we do not directly measure them. Many regressions that appear to have relatively few controls are actually controlling for a wide variety of factors via fixed effects.[24]

### 1.2.3 Challenge 3 – The Endogeneity of New Road Supply

The largest empirical obstacle to measuring induced VMT is that transportation agencies do not randomly assign road capacity, either within or across metropolitan areas. If governments assign new capacity to growing regions because traffic counts in these places are rising or expected to rise, then some VMT that seems *induced* might actually be *anticipated*. In this case the VMT would cause the road, not the other way around. A research design that failed to account for such reverse causality would bias estimates of induced VMT upward.

On the other hand, there is also reason to think that governments assign infrastructure spending, including road expansions, to *declining* regions, in an effort to shore up their economies. Alaska's infamous Gravina Island Bridge, also known as the "Bridge to Nowhere"[25] is a prominent example of such a capacity expansion, but evidence abounds that governments routinely direct road spending and new road capacity to places that have lost jobs or population (Duranton and Turner 2008; Glaeser 2016; Glaeser and Ponzetto 2018), and/or to places where elected officials are better positioned to secure funding (Knight 2005). To the extent such spending represents new supply in the relative absence of demand, VMT growth after a capacity expansion might be small.[26] A research design that failed to account for these political motivations would bias estimates of induced VMT downward.

In both cases—overestimating and underestimating new capacity's impact—the problem is that lane mileage, which is the independent variable of interest, is *endogenous* to driving. Lane mileage can affect driving, but driving can affect lane mileage. As a result, a simple correlation between lane mileage and VMT could be interpreted multiple ways. It could represent cause and effect (the road causes more driving), or effect and cause (more driving caused the road). For that matter the correlation could be spurious (e.g., roads and driving do not cause each other at all; both result from economic growth). If we are interested in the extent to which lane mileage causes driving, we need to overcome this endogeneity and isolate the one-way relationship by which new road capacity affects VMT.

The only foolproof path around endogeneity, unfortunately, is via controlled experiment. For obvious reasons this option is not available in studies of induced travel. No government will randomly assign new road segments to different metropolitan areas for the sake of scientific curiosity. Researchers must rely instead on a statistical approaches that are designed to reduce or eliminate the endogeneity threat. In induced travel research, the most prominent of these approaches is called instrumental variables (IV) estimation. An IV regression, as its name implies, relies on a variable called an instrument, whose purpose (similar to a fixed effect) is to absorb the influence of the unobserved factors that

---

[24] A related point is that someone looking at the NCST calculator and its output may not understand that the elasticity applied is the elasticity that emerges *after* controlling for all these other factors. Making that clearer could avoid an impression that calculator ignores other determinants of VMT.

[25] See Associated Press, 2007.

[26] Note that in such cases, road *speeds* might rise after the capacity expands, precisely because relatively little demand existed to use the new supply. Poole (2017) in a criticism of the induced travel literature, observed that in 17 areas where road capacity rose faster than VMT from 2000-2010, congestion delay grew more slowly. He interprets this as evidence that the new supply reduced congestion. Examining these 17 cases, however, suggests that at least part of this result may owe to weak demand. All 17 areas were small or economically struggling. The largest areas were Pittsburgh, St. Louis, and Wichita, all of which contended with unstable population and employment bases (the Pittsburgh MSA, for example, lost population steadily from 2000-2010).

might confound interpretation of the regression output.[27] A good instrument for lane mileage mops up any variance that might arise because more driving has caused more road construction, and lets the lane mileage coefficient be interpreted as the causal effect of road construction on driving.

The main problem with IV estimation is that good instruments are hard to find. To accomplish its task, an instrument must be correlated with the endogenous independent variable (this requirement is called the inclusion criterion), but *un*correlated with the dependent variable, *except* through its relationship the independent one (this is the called exclusion criterion). In studies of induced VMT, then, a valid instrument must be associated with new lane mileage but *not* associated with VMT in any way save the association with lane mileage.

While the inclusion criterion is usually easy to satisfy, the exclusion criterion is not. Weak or invalid instruments, as a result, have long been a common problem in empirical economics (e.g. Murray 2006; Davidoff 2016). The literature on induced VMT is no exception. Many older studies of induced VMT used questionable instruments, or otherwise failed to convincingly address endogeneity. Cervero and Hansen (2002), for example, estimate a regression relating VMT to new road capacity in California counties. They control for endogeneity using 19 instrumental variables, most of them county-level measures of weather, air pollution, and political partisanship. All these measures are correlated with lane mileage (satisfying the inclusion restriction) but also correlated with VMT in ways independent of their association with lane mileage (violating the exclusion restriction). Democrats, for example, are more common in dense areas (e.g. Rodden 2019) and dense areas have more overall VMT even controlling for lane mileage. Air pollution, similarly, is often worse in areas with more VMT, even controlling for lane mileage (Borck and Schrouth 2019). And so on. It is for this reason that I do not find the early generation of induced travel studies convincing. They do not adequately control for endogeneity, and their results are probably biased.[28]

## 1.3 Convergence on Unit Elasticity: The New Generation of Induced Travel Studies

A newer generation of induced travel studies relies on more persuasive instruments. The first of these studies was Duranton and Turner (2011), which used three instruments: the routes of major exploring expeditions that occurred between 1835 and 1850, the location or major rail routes in 1898, and the proposed routes of the initial interstate

---

[27] Without going into too much detail: the problem facing an induced travel researcher is that an unobserved variable (the government's nonrandom propensity to increase road supply) is not directly measured in the regression. That variable's influence, as a result, will end up in the regression's error term, and the error term will become correlated with the lane mileage variable. A correlation between a variable and a regression's error term violates one of the key assumptions of unbiased linear regression, and renders the coefficient on that variable suspect. The instrument addresses this problem by recognizing and controlling for the presence of that unobserved variable without directly measuring it. Essentially, the instrument purges the omitted variable from the error term. In doing so reduces or removes the bias from the coefficient on lane mileage. See Angrist and Krueger (2001).

[28] I won't run through every older paper, but the instruments used in Noland and Cowart (2000) and Fulton et al (2000) also do not plausibly satisfy the exclusion restriction. Hansen and Huang (1997) does not use an IV at all, and instead relies solely on fixed effects, which cannot by themselves address endogeneity. A separate problem with Cervero and Hansen (2002) is that it is probably over-instrumented. IV estimation is intended to reduce the biases that arise from a simple OLS regression. If the first stage of the regression (the stage where the instruments are employed) overfits the independent variable, it will bias the coefficients in the second stage of the regression back toward OLS. Intuitively, in a regression with 100 observations and 100 instruments, the R2 of the first stage of the regression will be 1, and the second stage of the regression will simply yield OLS estimates. A regression with 58 counties and 19 instruments will not yield a first-stage R2 of 1, but it is likely to overfit the data on the first stage and shift the second stage results back toward OLS.

highway system in 1947. The logic of these instruments rests on the fact that urban development tends to be path dependent, meaning the location of exploring routes, railroads, and proposed highways all strongly predict the location of road capacity today. Many exploring routes became roads, many routes that were rail lines in 1898 were abandoned and became roads, and the initial 1947 freeway map predicts the size and location of urban freeways today. Put another way, areas more heavily traversed by explorers in 1840 will, all else equal, have more road capacity today than areas that were not. The same is true of areas with rail lines in 1848, and areas where planners proposed freeways in 1947. In this way these variables satisfy the inclusion restriction. But the crucial point is that none of these factors—exploring routes, 1898 railroad locations, or the proposed route of freeways in 1947—should have any correlation with the amount of VMT in the late twentieth century, *except* through their correlation with road capacity itself. In this way they satisfy the exclusion criterion, and are valid instruments.[29]

Duranton and Turner use their instruments in regressions that examine a panel of 228 Metropolitan Statistical Areas (MSAs) observed in 1983, 1993 and 2003. They estimate regressions that measure VMT both across these metropolitan areas (comparing regions to each other) and within metropolitan areas (comparing regions to themselves over time). Their results suggest that on average, the relationship between new road capacity and VMT has an elasticity of 1.[30] A ten percent increase in road capacity will cause a ten percent increase in VMT. The paper contains many regressions, however, and depending on how those models are specified the results vary. The preferred regression, with an elasticity of 1.03, had a confidence interval of +/-0.21.

Duranton and Turner (2011) also include an accounting exercise suggesting that their equation lines up well individually-reported levels of VMT that occur across the entire sample.[31] They also decompose the induced travel. I reproduce that decomposition in Table 1. The decomposition has substantial variance, but suggests that a plurality of induced VMT (28 to 58 percent, depending on the estimate) comes from new household travel and new commercial driving. Some of the remainder (5-21 percent) appears to result from people migrating into the MSA as a result of the new capacity. Only a small amount (up to ten percent of the total) appears to be substitution from other roads. New road supply tends to pull vehicles onto the road. Or, as they put it, induced VMT "mostly reflects traffic creation rather than diversion."[32] A final point, which I will return to in the next section, is that Duranton and Turner also present evidence suggesting that increases in public transportation service also do not reduce VMT.

---

29 These newer instruments only became available because advances in computing allowed for the digitization of historical records, which in turn let researchers express the presence of exploring routes, highway plan mileage and early railroad tracks with quantitative precision. As such, the failure of earlier papers to use such variables is very much not a fault of their authors. Technological improvements made better instruments possible.

30 Estimates below 1 suggest that VMT rises alongside lane mileage but not at a 1-for-1 ratio, while estimates above 1 suggest that for each increment of lane mileage VMT growth will outpace lane mile growth on average.

31 Note however the caution I mentioned in footnote 21, about VMT estimates in the NHTS. In addition, an unpublished undergraduate research paper (Wang et al 2022) plots the predicted VMT from Duranton and Turner for each MSA against actual MSA values and finds that most line up well, although there were a small handful of large outliers.

32 With respect to the nonrandom assignment of road capacity, Duranton and Turner (2011) find that road investment tends to be larger in places that have suffered negative employment shocks.

**Table 1. Estimated Sources of Induced VMT**

| Source | Share of Induced VMT |
|---|---|
| Increased Household Driving | 9-39% |
| Increased Commercial Truck VMT | 19-29% |
| Traffic Diverted from Other Routes | 0-10% |
| Population Increase from Migration | 5-21% |

*Source: Duranton and Turner (2011) also Volker et al 2022*

In a nod to Downs, Duranton and Turner call this elasticity of 1 "The fundamental law of road congestion." They also use that phrase as the title of their paper. The paper has been cited over 1100 times, and is easily the most influential empirical examination of induced travel. But it is not the final word. In the nearly 15 years since its publication, several other papers have tried to measure induced VMT. These papers use different data, different geographies, and/or different approaches to managing endogeneity. In general, however, these papers arrive at the same conclusion: the elasticity of demand for VMT with respect to lane mileage hovers around 1.

Some examples: Hu and Zhang (2014) use map-based instruments similar to those employed by Duranton and Turner, but do so in the context of Japan. Across a sample of 93 Japanese urban areas observed between 1990 and 2005, they find an overall elasticity slightly higher than 1. Hu and Zhang's paper is notable for explicitly differentiating between expansions that increase the coverage of the road network and those that add lanes to existing routes. They find that new coverage accounts for more of the supply-induced VMT growth.

Graham et al (2014) is the rare paper that does not use instrumental variables. The authors instead employ a propensity score matching estimator (a statistical technique that tries to mimic an experiment by creating a control group from other observations in the data). Studying induced VMT across 101 US urban areas from 1985-2010, they too find an average elasticity of 1.

Hymel (2019) examines induced VMT at the state level from 1985 to 2015, using a dynamic panel model. He instruments with variables that measure the differing degree of influence that state Congressional delegations have over transportation funding. (The primary instrument is the cumulative years the state's Congressional representatives have served on Congressional Transportation Committees). This influence is correlated with road expansion, since most expansions use federal funding, but uncorrelated with VMT, since it is a product of political clout rather anything related to the overall demand for travel.[33] In all his estimates, elasticities hover around 1.

Garcia Lopez et al (2022), like Duranton and Turner (2011) use map-based instruments to study induced travel across 545 European cities from 1985-2005. Also like Duranton and Turner, they estimate both cross sectional regressions (comparing cities to each other) and longitudinal regressions (comparing cities to themselves over time). Their

---

[33] An example, again, is Alaska's Bridge to Nowhere. Alaska's congressional delegation secured the earmark for that bridge because Representative Don Young had tremendous seniority (he was the longest-serving member in Congressional history), and he chaired the House Transportation Committee. Neither his rise to that position, however, nor his ability to keep it, hinged on the amount of driving that occurs in Alaska (Alaska is a low VMT state). His authority over road spending owed instead to his political acumen and the overall disposition of the Alaska electorate.

instrumented average elasticity is just above 1. Unlike other papers, Garcia-Lopez et al include a control for the share of a region's roads that are tolled. The coefficient on this variable is negative and statistically significant, implying that more tolling is associated with less VMT. I return to this finding in Section 2.

The most recent paper in the induced VMT literature is Chang et al (2023). Here the authors re-estimate the regressions from Duranton and Turner (2011) but do so using instrumental variables quantile regressions, rather than instrumented ordinary least squares regressions. The benefit of such an approach lies in the ability to estimate the induced VMT conditional on initial levels of VMT—to test the idea, essentially, that places with different levels of VMT to start may see different levels of induced VMT when they add road capacity. Their results suggest that places with more initial VMT (such as Los Angeles) had lower induced travel elasticities. The lowest elasticity they found, associated with the most congested quantile of MSAs, was about 0.8, while the highest was 1.45. A caution here, however, is that quantile regression estimates can be noisier than least square estimates, and indeed Chang et al cannot reject the hypothesis that the results they find for congested areas are not statistically different from the average elasticity of 1.[34]

In sum, the majority of second generation investigations into induced VMT find an elasticity of 1. The result holds across different places, different time periods, and different geographic scales of measurement (e.g., states vs regions). It remains consistent across different methods of addressing endogeneity. Even the lone outlier study finds only a slightly smaller elasticity, and that result is not statistically distinguishable from the conventional findings.

## 1.4 A Skeptical Look at Unit Elasticity

However consistent the literature may be, a natural question, especially for practitioners, is the extent to which these results generalize to *particular* places and times. Is it really the case that every new increment of road capacity will, ceteris paribus, generate new VMT 1-for-1? Can we, based on the output of a regression analysis that covered hundreds of metropolitan areas, truly be confident that widening road *x* in place *y* will yield a VMT increase perfectly proportional to the increase in lane miles?

There are different ways to address this question, but I see four avenues a skeptic might pursue: that the newer literature has serious flaws and gaps; that some older studies finding smaller elasticities have been unjustifiably overlooked; that the time period covered by the induced demand literature is no longer relevant; and that the type of capacity being added is not adequately controlled for. I will discuss each of these in turn, and then also discuss a fifth issue, which is whether Travel Demand Models offer a superior alternative.

### 1.4.1. Do the Existing Studies Omit Important Variables?

No regression is perfect, so every induced travel paper fails to capture some elements that contribute to VMT. That said, most major variables are controlled for—a fact that makes some existing critiques difficult to reconcile. A Fehr and Peers paper from 2019, for example, observes that many regressions do not adequately control for prominent determinants of VMT. The elasticity methods, it says, "are not sensitive to land use context, geographic constraints, (e.g. water or topography barriers) or the level of existing congestion." This critique is repeated in Milam et al (2022).

---

[34] In a separate part of the paper, Chang et al use a travel model to estimate that the elasticity in Los Angeles is actually quite low (0.3) but this exercise is largely separate from the regression. I discuss travel demand models more below.

The authors do not elaborate in either case, so the precise meaning of this criticism is unclear, but if the implication is that the elasticities are derived from regressions that do not control for topography and land use, that is incorrect. The Duranton and Turner paper, for example, includes explicit controls for elevation, the ruggedness of terrain, the dispersion of development, the average climate, the number of heating and cooling days, and so on. Chang et al (2023), as has been discussed, explicitly control for existing congestion levels. Recall too that almost every paper uses place and time fixed effects, which implicitly control for many of these other variables. Readers of these articles might not see topography or climate in the regression tables, but as long as these factors are constant within the category of place or time (e.g. if San Francisco's general topography or climate is relative stable, or if recessions or bouts of bad weather occur in particular years), the fixed effects will absorb them, and prevent them from biasing the estimated elasticity.[35]

## 1.4.2 Does the Unit Elasticity Consensus Needlessly Overlook Other Research?

Skeptics of unit elasticity sometimes mention two papers, both by respected and qualified transportation researchers, which appear to show that new road capacity has much smaller induced travel effects. The first is Cervero (2002). This article uses a path analysis, which is an early form of structural equation modeling, to capture the simultaneous relationships between highway building, travel speed, VMT, and land use changes/development. Cervero examines 24 highway widening projects that took place in 9 California counties from 1980-1994 (his total sample size is thus 360: 24 projects each observed 15 times). He finds an overall elasticity of 0.8—i.e., a 100 percent increase in capacity would see an 80 percent increase in vehicle travel—but contends that only half of that new travel was caused by the new capacity. The remainder would have arrived anyway, as a result of economic growth. Hence the actual elasticity, at least in these California counties, is closer to 0.4—much less than estimates made before or since. In both this article and subsequent publications, Cervero uses this result to gently chide opponents of freeway expansion for their lack of rigor. "Critics of any and all highway investments," he writes, "even those backed by credible benefit-cost analyses, should more carefully choose their battles."[36]

While the paper's conclusions are provocative, it isn't obvious that they provide useful guidance for the question at hand: when new roads are built, how much total VMT will they induce? We can count the reasons why. First, unlike the studies described above, Cervero (2002) does not account for network-wide effects. He restricts his travel analysis to the 24 expanded freeway corridors, and his land use change analysis to a two-mile buffer on either side of them. Any travel or development induced outside these corridors goes unobserved. The corridors, moreover, are all located in "small- to medium-sized municipalities" in the suburbs. None of the studied expansions occur in larger cities (he analyzes no expansions, for instance, in or near Los Angeles, San Diego or San Francisco). These factors alone make the resulting elasticity difficult to compare with other estimates: where the more recent papers study network-wide effects on entire regions, Cervero (2002) measures corridor-level effects in small-to-medium sized suburbs.

---

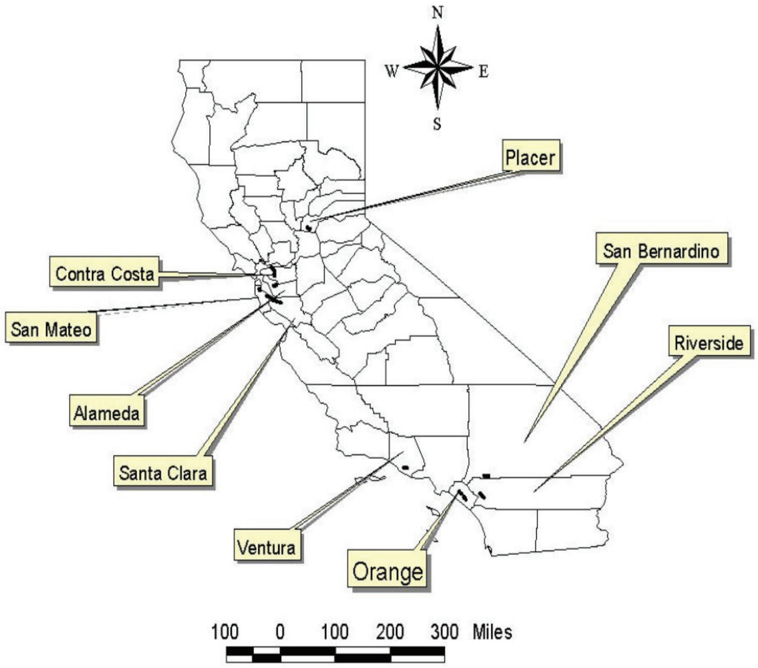
[35] In fairness, the exact meaning of the Fehr and Peers criticism really *isn't* clear, and another way to interpret it is in the context of a different type of impact analysis — one that estimates the local or sublocal effects of adding capacity. In that case the NCST elasticity method could in fact be insensitive to context, because (for example) two very different neighborhoods in the same MSA would be assigned the same baseline level of VMT. But this criticism does not really imply that the NCST calculator is wrong; it is more an observation that the calculator can be used incorrectly. A calculator based on regional estimates should not be used to make neighborhood estimates.

[36] He did add, entirely correctly, that most problems could be solved by pricing any freeways that were built, and existing freeways as well.

This problem, of the elasticities being incommensurate, is exacerbated by Cervero's decision to study a different dependent variable. Cervero never actually measures VMT. The closest he comes is a variable measuring the facility's *proportion* of total county VMT. The paper, as a result, does not directly examine the question of whether new road supply increases VMT. It examines the question of whether new road supply leads to a situation where an expanded freeway segment carries a greater share of the region's total driving than it did before. Those aren't the same thing.

Perhaps the biggest problem, however, is that the paper's already-small sample is actually smaller than it seems. The paper describes its sample as 24 separate projects in nine counties. Both the table and map that show these projects, however (I reproduce the map below, as Figure 2) make plain that this isn't quite right. Nineteen of the projects are actually in only 4 counties, and these 19 are clearly not distinct, separate expansions—they are individual segments of larger widening projects. The regression analysis treats them, however, as independent observations, never accounting for their adjacency, their sequential nature, or the very strong likelihood that the government approved and funded them in clusters. Failing to control for these factors inflates the sample's variance, neglects important information, and almost certainly bias the results.[37]

**Figure 2. Map of Location of 24 Freeway Projects Across Nine California Counties**



---

[37] Cervero employs a fixed effect for each project, but does not use additional county or location fixed effects to capture the hierarchical nature of the data. To see why this is a problem, imagine reading about a new educational intervention given to 400 kids across a large school district. Researchers studying the intervention find it has little impact. Then imagine learning that the students in the study were not evenly spread across the district's schools; that in fact, 80 percent of the kids in the study were clustered in just four classrooms. The study, moreover, didn't account for this clustering. You could be forgiven for wondering if the study is capturing some unobserved aspects of those classrooms, rather than isolating the effects of the reading program, and for wondering if a study where the 400 kids were more evenly distributed might yield different results. Clustered observations cannot be analyzed as though they are independent.

The second paper that skeptics sometimes mention is also from 2002: Mokhtarian et al, published in the journal *Transportation*. This article used a matched pairs analysis, and is the rare study that found no induced vehicle travel at all. The researchers compared highway segments that were widened to similar segments that were not, and found that, over 20 years, traffic volumes did not rise more on the widened segments than on the control segments. Milam et al (2022:4), writing about this study, see it as reason to think that elasticity estimates closer to 1 may not be appropriate for California. The finding of zero induced travel, they note, "may be an example of the limitation ... where the elasticity-method is not sensitive to a unique local context."

Before examining this paper in more detail, a few words about situations like this. A finding of zero induced travel is an outlier finding, and in normal science outlier findings warrant extra scrutiny. The mantra is to trust literatures, not individual papers. The point of this guidance is not to gatekeep or imply that outlier papers are automatically wrong, but to acknowledge that as a matter of probability, a study that finds the opposite of everything preceding it is more likely to represent an error than an intellectual revolution. Mistakes are more common than paradigm shifts.

Sometimes, of course, a single paper *does* overturn the conventional wisdom. When that happens, however, the outlying empirical result is usually accompanied by a powerful theoretical explanation, and the single paper—because of that theoretical explanation—is followed in short order by subsequent articles that replicate and extend its findings, building toward a new understanding of the phenomenon being studied.[38] None of that occurred with Mokhtarian et al (2002). The authors themselves are ambivalent about their findings of zero induced travel and cannot quite explain them, and no one—neither the authors or other researchers—have returned to the paper's method, or argument, in the 20-plus years since it was published. This fact alone should make us cautious about the result.

If we now subject Mokhtarian et al (2002) to a bit more scrutiny, we see that overall its findings are not very persuasive. The article is a replication of an earlier report: Hansen et al (1993). Hansen and his team studied induced VMT by examining 18 freeway expansion projects in urban California counties that occurred between 1972 and 1977. Using regression analyses (but without controls for endogeneity) they examined both corridor-specific and regional induced travel, and found relatively low elasticities. Mokhtarian et al took the same 18 freeway segments but—rather than estimate a regression in the manner of Hansen et al—instead matched the segments with other, similar freeway segments that had not been widened. They then compared the growth in Annual Daily Traffic (ADT) on these paired segments from 1976-1996. Doing so yielded no evidence of induced travel at all; ADT growth was the same on both the treated and controlled segments.

The Mokhtarian et al result is thus different from the result of Hansen et al (1993) which it tried to replicate, and very different from the finding of unit elasticity that is conventional in the literature today. Why? A first potential reason is the nature of the sample. Hansen et al purposefully chose to examine freeway segments where they believed new capacity would not increase speeds. Their study, they wrote, "focuses on capacity additions to existing facilities whose impacts on travel speeds are likely to be slight or nonexistent."[39] Because a substantial share of induced VMT is understood to arise from faster travel, these selection criteria may have biased the study toward finding a smaller

---

[38] As an example: in 1994 two economists published a careful empirical study showing that raising the minimum wage did not decrease employment, as would normally be predicted, but actually increased it. The finding was striking, but what made it influential was the elegant explanation for it. A higher minimum wage should depress employment when markets are competitive; the authors observed that the labor market in minimum wage establishments might not be competitive, and instead be monopsonic. The monopsony could create a situation where workers are paid below their marginal product and employment levels are as a result inefficiently low. Enforcing a wage floor could push both wages and employment up. As a result of that monopsony theory, the paper has been cited over 6,000 times and is credited with launching created "the new economics of the minimum wage." See Card and Krueger (1994).

[39] This quote is from page 3-4 of the report.

effect. A broader study that captured places where road expansions *would* increase real-or-perceived short run speeds would likely find more induced travel.

Second, the Mokhtarian et al replication was not a regional study. Like Cervero (2002) this paper studied corridor effects rather than network effects. And as was the case with Cervero (2002) this restriction to the corridor probably pushed estimates of induced travel downward. A third issue is Mokhtarian et al's method. The authors took Hansen et al's sample and, for each expanded segment, tried to find a suitable and matching control segment. This is no easy task. On the one hand, a control that is geographically close to the expanded segment is advantageous—it provides access to similar destinations, and its flows are likely affected by similar levels of population and economic activity. On the other hand, such a proximate segment may also be affected by the expansion itself. The new capacity might, for example, pull traffic from the control to the expanded segment, thereby contaminating the control. Avoiding this outcome requires choosing control segments that are further away. But moving further away makes the control less similar to the treated segment, in terms of immediate surroundings and other context. A good control, then, is close enough to be comparable but not *so* close as to no longer be a control.

It is not clear that Mokhtarian et al (2002) struck this balance successfully (in fairness, it is not clear *anyone* could). The authors tried to keep controls and matches relatively close to each other, and chose control segments that had roughly the same amount of daily traffic in 1976 as the expanded segments. Eleven of the 18 controls are in different counties than the expanded segments. Some of these are near each other (for instance, on opposite sides of a county line) but some are not. And many of the matches, both those across and within counties, raise questions.

An expanded segment of the 91 freeway in LA County, for example, is matched to a control segment of the 60 freeway, also in LA County. But LA County is large, and the control segment is almost 40 miles from the expanded one. The two segments may have had similar ADT levels in 1976, but their surrounding contexts are strikingly different. The expanded segment of the 91 lies at its terminus, where it intersects with the 110 freeway, and is only a mile away from where the 110 has an interchange with the 405. It runs through a busy, heavily-populated part of west Los Angeles, and during most of the study period it offered access to a dense aerospace industrial corridor. The treated segment is seven miles from LAX airport, and overall, traffic volumes near it are large.

The control segment on the 60, in contrast, lies at the junction of the 57 freeway. The 57, at this intersection, carries less than half the traffic of the 110 at the 91 interchange. Furthermore, where the expanded segment of the 91 is surrounded by population centers, the control segment of the 60 is in an outlying area. It is bounded on one side by the City of Industry, a municipality that (as its name implies) has numerous factories and warehouses and almost no people (the 2019 population was just over 200). It is bounded on other sides by large areas of open space: it is just south of an 1,800 acre regional park, just west of a 14,000 acre state park, and nine miles south of a mountainous 655,000 acre national forest. It is hard to look at these two segments and call them comparable.

A further problem is that the paper's control segments do not always appear to be independent of each other. The same segment of the 60 just discussed, which the paper uses as a control for the 91, lies only few miles west of another segment of the 60 that Mokhtarian et al use as a control for a different expansion. Two of the expanded segments, as a result, will be measured against what is essentially the same control. And while Mokhtarian et al ensure that the control segments were not themselves widened, they do not appear to address the possibility that roads *near* those segments may have been expanded. But doing so is important, and is one reason why regional studies that capture entire networks may be more reliable than corridor specific improvements. Indeed, a quick look at the segments used in Mokhtarian et al shows that the two control segments of the 61 mentioned above are parallel to, at one point within six miles of, a different freeway segment (on the I-10) that is included on the list of *expansions*.

Recall that Mokhtarian et al collect ADT data for both the expanded and control segments from 1976-1996. In doing so they introduce yet another problem. Only five of the 18 expansions were finished in 1976 or after; nine were finished in 1973 or before. For half the expansions, then, at least three years of post-expansion traffic growth went unobserved, and there is no true "before" period for any of the sample. Matched-pair comparisons work best when the treated and controlled observations have parallel trends before the treatment occurs. This paper never establishes that such trends were present.

In summary, this is a replication of a study whose sample was chosen based on the low probability that the chosen expansions would increase vehicle speed. The replication, in turn, uses questionable segments as controls, analyzes no data from before the expansions, and indeed for half the observations doesn't collect any data from immediately *after* the expansions. On top of that, it also never measures VMT (only ADT). Given all these limitations, it seems unwise to rely heavily on this paper's findings, or use them as a basis for discounting the rest of the literature.

### 1.4.3 Should the Current Literature be Discounted Because of the Time Period it Covers?

The Duranton and Turner data set ends in 2003. Much has changed since then. Most notably, in 2004, driving—which had grown steadily for a century—began to fall. From 2004 to 2013, per capita VMT declined. From 2005 to 2011, *absolute* VMT declined, marking the first time since World War II that total driving fell while the economy grew. (The Great Recession occurred during this decline, and contributed to it, but absolute VMT was falling before the recession and continued falling after it). Since 2013, VMT has begun growing again at the national level, but in some parts of the country, including Los Angeles, it remains below its previous levels.

To put this VMT slowdown in perspective: from 1990 to 2003, US VMT rose from roughly 2 trillion to 2.9 trillion—a 45 percent increase in 13 years. From 2004 to 2019, it rose from 2.9 trillion to 3.3 trillion—a 14 percent increase in 15 years. So driving really did slow noticeably. And over the same period, ridesharing services and especially social media exploded.

One could look at this falling VMT and conclude that existing elasticity estimates are wrong, simply because they are the product of an earlier time when driving was growing faster.

A first observation about this argument is that embracing it requires rejecting the argument we just discussed in Section 1.4.2, about how the current conversation unjustly neglects some older literature. One cannot simultaneously protest the omission of older studies *and* argue that newer studies should be distrusted *because they are too old*. If Duranton and Turner (2011) is obsolete because its last year is 2003, then both Cervero (2002) and Mokhtarian et al (2002) must be equally or more suspect, since the expansions they studied date back to the 1970s. Age is either disqualifying or it isn't.

A second observation is that the criticism, even on its own terms, isn't entirely correct. Some induced travel studies *do* include more recent time periods. Both Hymel (2019) and Graham et al (2014) include the years through 2010 in their analyses, and thus capture a number of years when both VMT and per capita VMT were falling. These analyses still find unit elasticity.

Those points aside, the basic question still merits consideration: perhaps something has changed since 2010 that would change the elasticity. In an era of ridesharing and social media, for example, VMT estimates derived from earlier periods might no longer hold.

Evaluating this argument requires that we be mindful of an important distinction: between how much people drive, and how much *more* they drive *because* road capacity has expanded. This is the difference between the total quantity of VMT (which has many determinants) and the elasticity of demand for VMT with respect to lane mileage (which is narrowly about how much a change in road capacity changes demand). While falling overall VMT *could* imply a changing elasticity, it does not automatically do so. Total VMT is (again) a product of many factors, and as such can fall even if drivers' responsiveness to road supply doesn't change.

Falling demand, in short, is not the same as a changing elasticity of demand. When the demand for VMT falls, total VMT declines. Were we to plot this on a standard supply-and-demand graph, we would see the demand curve move to the left, suggesting that at any given price, consumers drive less. Such a shift could occur for many reasons. Perhaps a cheaper substitute for driving arose: to the extent people once drove to socialize with friends, the proliferation of smartphones and social media made it possible to do so from home. When the price of a good's substitute falls, so too does demand for that good.

Or perhaps a complement to driving became more expensive; when the price of a complement rises, the demand curve for the primary good shifts leftward. Gasoline is a complement to VMT (more driving generally requires more gas). Fuel prices, for this reason, figure prominently in explanations for the driving downturn of the 2000s. From 1998 to 2012, real gas prices rose from their lowest point in history to their highest since 1918. The important point here is simple: total VMT can fall for many reasons unrelated to the relationship between induced VMT and lane mileage. When it does we see it as a leftward shift in the demand curve.

A falling elasticity of VMT with respect to lane miles, in contrast, implies that the demand curve has become *steeper* when quantity demanded is plotted against road capacity, regardless of where the curve is. Drivers have become less responsive to that portion of driving's total price that is represented by an increase in lane miles (which is, again, some combination of access benefits and travel time/reliability/safety benefits). Because this portion represents only one component of demand, however, it is entirely possible for the elasticity to remain constant even as demand overall declines. Falling VMT, on its own, does not mean people are less responsive to added lane mileage.

This reasoning can admittedly be confusing, particularly when an elasticity's prediction of induced VMT exceeds a region's *total* VMT (e.g. Systems Metric Group 2023). Such a result, however, is not as nonsensical as it might initially seem. Many factors influence total VMT, some positively (like new roads) and some negatively (higher gas prices). Thus new roads could induce a large amount of VMT, but rising gas prices could swamp that effect, leading to less VMT overall.

This situation, where a real effect in one direction is overwhelmed by a larger effect in the other, is commonplace. One example: by most estimates, the American Recovery and Reinvestment Act (ARRA) created hundreds of thousands of jobs in the aftermath of the Great Recession. Almost every place that received ARRA funds, however, saw its overall employment fall, simple because the recession destroyed more jobs than ARRA created. But it would not be correct to look at this negative employment growth and conclude that ARRA did nothing: its positive effects were simply overwhelmed by the recession's negative effects. A second example: imagine a person who cuts all sugar and alcohol from their diet and begins exercising, but continues smoking three cigarettes a day. This person may well live longer than an identical person who doesn't make those changes, but it would be odd to look at the longer life of the first person and conclude that previous research about *cigarettes* is wrong, and that smoking has no impacts on longevity.

Is it possible that the elasticity could change, and for the same reasons that demand changes? Yes. If the public loses some of its taste for driving, they may well drive less at any given price *and* be less responsive specifically to increases in lane mileage. The point here is only that if we want to learn about elasticities, we need to measure elasticities. We cannot infer elasticities through observed changes in total demand. To the extent arguments about time periods rest on the fact that VMT growth has slowed in recent years (or fallen in per capita terms) they are merely suggestive.

### 1.4.5 Could the Changing Nature of New Capacity Change the Elasticity?

The discussion above addressed the argument that changes in travel behavior since 2003 also implied changed elasticities. A slightly different idea is that we should revisit elasticities because the quantity and nature of new *road supply* has changed. Over the duration of Duranton and Turner's data, lane miles in the LA region grew by about 16 percent. A substantial share of this growth involved opening an entire interstate segment: 150 lane miles of the I-105). In the years since, both the amount and type of new lane mileage has changed. Total lane miles in Los Angeles from 2001 to 2019 were basically unchanged (Luo et al 2023).[40] The little mileage that was added, furthermore, tended to come from widening existing routes, rather than opening entirely new roads. (Many of the expansions were HOT lanes).

Perhaps elasticities calculated when road building was more common are no longer relevant, in when new lane mileage is less common? The simplest answer to this question is that elasticities are proportional relationships, and as such the estimate holds: if a ten percent increase in lane mileage increases VMT by 10 percent, then a 1 percent increase should increase it by 1 percent. A smaller increment of new road will yield a smaller increment of induced driving, but there is no reason to think the proportion is wrong.

A reasonable counterargument, however, is that elasticities are more reliable when they reflect changes actually observed in the data, rather than changes extrapolated from it. By this logic, if we do not observe small increments of new lane mileage empirically, we should be less confident about applying empirically-derived elasticities to projects that add only small amounts of new capacity.

That's a data argument, but we can deploy theory to reinforce it. There is some reason to think that newer road capacity will have less impact than earlier generations of road investment. As Guiliano (1995) argued in an influential article, the relationship between transportation and land use is strongest when new roads and highways open up vast swathes of land to new development, dramatically increasing automobile access. Later additions might expand speed and capacity, but do not open up new destinations in the same way and are thus less likely to alter as travel behavior. One result in Hu and Zhang (2014) in the Japanese context, reinforces this idea. They found evidence suggesting that road expansions that increased coverage accounted for more induced VMT than those that simply widened existing routes, although the difference was not large.

I consider this a potentially fruitful area for future research. It is possible that smaller increments of new road supply do not trigger the same response as larger additions, and the current literature says little to say about this question.

Lest anyone think we have found a large hole in the unit elasticity's armor, however, let me add two cautions. First, even if we find that the elasticity is lower for small additions of new capacity, that finding, as a matter of math, cannot hold *in general*, simply because small changes to a road network will, over time, sum to large changes. Suppose a government proposes a small road expansion, and we determine that because it is small it will induce less travel on the network. Evaluating that expansion based on that determination implicitly assumes that little *further* expansion will occur in years to come. The unit elasticity estimate, remember, is a long-run regional elasticity. A small project may well be different on the improved corridor in the short term, but over 10 or 20 years there may be little difference between ten projects that add one-half of a lane mile and one project that adds 20 lane miles.

---

[40] Luo et al show new lane mileage as falling very slightly over time (-0.28). The negative number probably reflects a combination of changes in MSA boundaries and measurement error, rather than an absolute loss in lane mileage.

The second caution is that while the US does build fewer roads today than we did in the1980s and 1990s, we shouldn't overstate the differences between then and now. It is not the case that the 1980s saw an avalanche of new supply. To the contrary, in 1983, when the Duranton and Turner panel begins, the federal highway program was well past its high water mark. Guiliano wrote her essay about the weakening transportation-land use connection in 1995, in part to capture the new era of declining road building. Similarly, the preeminent empirical analysis of highways shaping development patterns (Baum-Snow 2007) examines the period from 1950-1990, to capture the period (the 1960s) when freeway construction reached its zenith. To give this point some specific numbers: California added between zero and 20 new centerline miles of freeway annually from 1981 to 1993. In most years from 1963 to 1973, in contrast, it added over 200 centerline miles annually (Brown et al 2023).

The upshot is that road building—and especially road building that increased the network's coverage—was already past its peak at the beginning of the Duranton and Turner study period. So while it is possible that the current literature captures substantially larger access effects than we would see today, it may not be likely.

### 1.4.6 Should we Use Travel Demand Models Instead?

The most plausible alternative for estimating induced VMT would be Travel Demand Models. Where regressions are widely used by academics, TDMs are common in professional practice. TDMs vary in their specifics, but most use survey data to simulate a region's current population, travel and employment, then project that simulated population forward to predict travel behavior in different scenarios.

The classic TDM is a Four-Step Model (sometimes also called a trip-based model) and as its name implies it estimates travel demand in four steps. First is *trip generation*, where the model uses the simulated population and employment to determine both the amount and purpose of trips that will originate in a given area (e.g., how many work trips will start there, how many school trips, and so on). Second is *trip distribution*, which matches each trip origin with a destination (where those work, school and other trips, will end). Third is *mode choice*, wherein the model assigns each trip a mode, and fourth is *route assignment*, where the model predicts the path each traveler will take to complete their chosen trip. The route assignment step rests on a form of Nash equilibrium assumption: the model predicts that every traveler will choose the shortest route between their origin and destination, subject to the understanding that all other travelers will be doing the same. What emerges as the model's output, then, is a picture of future regional travel: its quantity and its distribution across modes, places and routes.

Compared to a regression, the advantage a TDM offers is the possibility of an explicit counterfactual. A researcher can use a Travel Demand Model to forecast what will happen in 20 or 30 years if a highway segment is or isn't built. Regressions, in contrast, offer only implicit counterfactuals: by examining hundreds of different places over many years and controlling for many characteristics of those places and times, induced travel regressions let researchers estimate the average effect of an increase in road supply, as well as the confidence interval surrounding that average. The Duranton and Turner regressions, for example, essentially say that over three decades and across and within 228 regions—some large, some small, some growing, some not, etc.—the average elasticity of VMT with respect to road supply is 1.0, that if a similar analysis were run 100 times the average would, in 95 percent of those analyses, fall between 0.8 and 1.2.[41] This offers us reason to think that in most situations, the elasticity is 1. But the regression is, by definition, backward looking: its data cannot include the exact characteristics surrounding a proposed project, for the simple reason that a proposed project does not yet exist.

---

[41] Assuming the assumptions underlying the analysis are correct.

The TDM, in contrast, can make such a place-and-time specific prediction. Using a TDM called OCTAM,[42] for example, analysts in 2024 predicted that adding 16 new lane miles of Express Lanes to I-5 in Orange County would create an additional 7.4 million VMT per year. This calculation could be made because OCTAM is specifically calibrated to represent the present and predicted future of Orange County.

The downside to a TDM is that almost all the data are simulated. Although the model begins with empirical data, it uses those data to predict the size and location of a future population and workforce, predict its travel, and then predict what how that travel will change in the presence of a hypothetical increase in road supply. Where every implicit counterfactual derived from a regression represents behavior observed in the real world, the TDM's explicit counterfactual relies on a forecast of future conditions, as well as a forecast of how future people will react to future interventions that change those conditions. Such forecasts inevitably rest on a strong assumptions, not all of them realistic. The model may need to assume that the quantity of trips stays fixed over time, or that land uses do not change, or the region's economic trajectory is simply a linear extrapolation of the base year.

For this reason TDMs have long courted controversy, and over the years have been the subject of ample hand-wringing and critique (e.g., McNally 2004; Hartgen 2013; Walker et al 2019; National Academy of Sciences 2020; Shewmaker 2012). TDMs have been criticized for inaccuracy and opacity. Skeptics have said that the many assumptions that underlie the models, and the many opportunities the modeler has to influence the output, at best make TDMs unrealistic, and worst breed impropriety—that is too easy to start with a desired outcome and work backward to a model that validates it (Naess 2012; Hartgen 2013; Flyvberg et al 2006).

Such criticism isn't always fair. Most TDMs are in fact impressive programs that can handle vast amounts of data. Even at their best, however, TDMs will not solve the problem of estimating induced VMT. It is more accurate to say that replacing a regression with a TDM will exchange one form of uncertainty for another. Regression-derived elasticities are real-world observations, but they occurred in the past, and not always in places new capacity is being proposed. TDM estimates are specific to the project, but in being forecasts they depart from observations of real world data.

With respect to induced travel, TDMs also suffer from a more relevant and specific limitation: they tend to underestimate it. This problem is bigger in trip-based 4-Step Models than Activity Based Models, and arguably more severe with models that use static rather than dynamic trip assignment (e.g. Marshall 2018), but no TDM is immune to it. At issue is the model's ability to incorporate feedback—the process by which travelers react to changes in the system.

Recall that a TDM starts with a given level of employment and population, then estimates how many trips will occur, to what destinations, and by which modes and routes. If the model is run to measure the impact of new road capacity, one of the first results it will yield is that the expanded capacity will make travel times faster. This initial result has often been used to justify road expansions (Naess et al 2012).

In the real world, however, the process doesn't stop there. The faster travel times will change people's travel decisions, since the decisions were based in part on how long the trips were likely to take. To capture these behavioral changes, the model needs to feed the new travel times back into its initial steps. Many models now partly incorporate such feedback: by feeding the faster travel time into the model's mode split stage. Doing so captures the possibility that because driving is faster, people will make some trips by car rather than transit or active modes.

It is less common, however, for models to feed the faster travel times back into trip distribution or trip generation. As a result, the models are less likely to account for the possibility that faster driving will increase the region's total number

---

[42] Orange County Transportation Authority Model

of trips, and/or make some of those trips longer in distance. And it remains quite rare for models to feed faster travel times back into initial assumptions about the region's growth and distribution of population and employment. This last omission is particularly problematic; the model ignores the possibility that faster speeds can, over time, change population dynamics and land use patterns. These changes, over, can account for substantial shares of longer run induced VMT (e.g. Metz 2021, Volker et al 2020, 2022; Luo et al 2023; Naess et al 2012; Duranton and Turner 2011).

The upshot is that while TDMs can reasonably capture changes in route and mode that arise from faster travel times, they do not account for changes in the length, frequency, or (over the longer term) origins, destinations and total numbers of trips. The TDM will, as a result, underestimate induced travel.

Modelers are not unaware of this problem, and have ways to overcome it. Perhaps the most common approach is an "off-model adjustment"—a situation where the modeler manually changes the model output after the fact to bring it into closer alignment with conventional estimates of induced VMT. The modeler might base this adjustment on the output of a different model, the judgment of an expert panel, or the judgment of the modeler. Caltrans, for instance, in its *Transportation Impact Analysis Framework*, suggests that analysts:

> *Apply off-model approaches using the best available information or tools to compensate for TDM's deficiencies, making approximations as needed where more precise data or information are not available ... Where a quantitative assessment cannot be reasonably undertaken, a qualitative assessment may be undertaken.*[43]

There are different ways to view off-model adjustments. One uncharitable reaction is that they are simply unacceptable, and their existence is an indictment of the models themselves. It is not permissible, for instance, for an academic researcher to estimate a regression and then multiply its output by some factor, and justify doing so because it brings the result closer to the prior consensus of the literature. (Indeed, were that the goal, the simplest approach would be to abandon the regression and just rely on the literature). Given these norms, it isn't clear why we might tolerate after-the-fact modification of TDMs. Models are supposed to impose discipline on researchers; when researchers can discipline models, the research activity starts to look less like science.

A more benign view of adjustments, however, is that large predictive exercises are impossible without them. TDMs are hardly alone among models in needing to be modified after the fact. Large-scale climate models, for instance, also require large post-hoc adjustments—a process that climate modelers call "tuning" (Hourdin et al 2017). If we accept that adjustments are often necessary, what matters is how we carry them out.[44]

Adjustments will be easier to defend if they are predictable, transparent, and rigorous. Predictability and transparency mean that the adjustments are made not made in an ad hoc fashion, and that the modeler makes both the baseline output and the adjustments used clear to anyone looking at the results. In some cases this might involve pre-registering the adjustments: the modeler commits beforehand to adjusting by some specific factor regardless of the model's initial run. Doing so is more defensible than deciding on the adjustment after the fact (i.e., the modeler looks at the result and decides it needs to be adjusted, then goes about testing various adjustments). That latter scenario,

---

[43] Caltrans (2020). Other guidance is similar: Fehr and Peers (2019): "If a trip generation module is not sensitive to travel time and cost, the analyst can manually adjust the vehicle trip generation rates or use off-model processing to increase the VMT forecasts. An important part of the adjustment process is to verify that it is warranted."

[44] Note that climate modelers, too, often evince discomfort about the extent to which their models must be tuned. One 2017 review observed that "The importance of tuning is probably not advertised as it should be," and then explained that this reluctance to discuss it arose "because tuning is often seen as an unavoidable but dirty part of climate modeling" (Hourdin et al 2017:590).

called data-dependent decision making, is generally not considered sound statistical practice. The replication crisis that swept through the social and medical sciences in the 2010s arose in part from the prevalence of data-dependent modeling (e.g. Gelman 2018).

Whether an adjustment process is preregistered or not, the transparency criterion suggests that modelers should document and show the impact of all adjustments *tried* (not just used), to demonstrate that the model's results are not overly sensitive to how it is calibrated.

The biggest issue, though, is rigor: where does the adjustment come from? As mentioned above, industry and government guidance suggest several sources. One possibility is to run a separate land use forecasting model, then use the results of that model to adjust the TDM output. A second option is to employ an expert panel that could qualitatively forecast land use changes. A third option is to simply rely on the judgment of the modelers themselves.

All these options create problems, by increasing what are known as "researcher degrees of freedom." The process of determining an adjustment factor involves many small decisions, some of which have no obviously correct answer. If we are going to feed the results of a land use model into the transportation demand model, what variables and assumptions should go into the land use model? If we want to consult an expert panel, who should serve on it? How should we explain the panel's charge to its members, and how should we weigh their responses, particularly if they disagree? Even modest changes in any of these decisions, individually or summed together, could meaningfully alter the model's results, which suggests that the same dataset could, if analyzed multiple times using slightly different but reasonable assumptions, yield very different answers to the same question.[45] Gelman and Loken (2014) call this dilemma the "garden of forking paths"—decisions without obvious correct answers take analysts to wildly divergent destinations.

In fairness, this problem also afflicts conventional regression analyses. Here, however, transparency comes back into play. In econometrics, both the nature of the analysis, as well as professional standards, can help mitigate the problem. Both econometric research and its outputs are heavily documented and easily checked. At better academic journals, the peer review process demands lengthy descriptions of data and methods, a demonstration that the main results are not sensitive to small changes, and an open data policy that lets other researchers check the findings. Consider Duranton and Turner (2011), the article that is primary source for the NCST calculator. The article, available free online, is 36 pages long, and includes a lengthy description of its data and methods. To demonstrate that their main result is not sensitive to how the regression is run, the authors shows over 60 additional regressions. A supplemental appendix, also free online, shows over 100 more. Finally, both the data and statistical code are available free online, and can be easily executed in a free statistical package like *R*. For anyone with some statistical acumen, then, the work in this article—and thus the work underlying the NCST calculator—is easy to check.[46]

In contrast, now consider the primary VMT analysis for the I-5 Express Lanes in Orange County, mentioned above, which was carried out using OCTAM. The analysis is one and a half pages long. Readers are told that the TDM was run, and then shown a single small table with the model results. The authors mention that more detail can be found in a Technical Appendix, but anyone clicking to that Appendix will find a single page with a message saying the technical

---

[45] Note that this problem does not involve intentional manipulation of results. The same conditions, however, make dishonest use of models easier.

[46] As an example: the replication problems of the social and medical sciences were discovered—admittedly belatedly—precisely because it was relatively easy for other researchers to peer inside the flawed articles, and use their data and documentation to demonstrate their flaws.

memo is only available upon request (Jacobs 2023). The model itself, meanwhile, is inaccessible to the public. A third party cannot easily and independently replicate the analysis and its results.

These facts do not automatically make the TDM analysis wrong, and certainly not every documentation of a TDM analysis will be documented in such a curt and unhelpful way. For that matter the public's inability to access the TDM itself is hardly nefarious: TDMs are gigantic, and require tremendous computing power, and it isn't uncommon for large agencies to restrict access to them. The point remains, however, that for a variety of reasons it is harder for outside observers to check the work of TDMs. So long as such opacity remains professionally acceptable, it will be difficult for TDM proponents to argue that regression-derived elasticities are insufficiently rigorous. Whatever flaws might accompany the NCST calculator, its pedigree is transparent and its methods replicable. In many instances the same cannot be said for TDMs.

A simple step to strengthen the case for TDMs might involve adopting a stronger ethos of shared data and research transparency—as suggested by Walker et al (2019) in their review of modeling practice. If nothing else, such a step would make it easier to accurately compare TDMs and regressions.

# II. The VMT Effects of Managed Lanes

Little is known about the VMT effects of managed lanes, because managed lanes are quite rare. The vast majority of American roads are free general purpose lanes. In 2012, the United States had about 2500 total lane miles of HOT lanes, compared to over 200,000 miles of interstate and over 8 million total lane miles of road. The relatively few managed lanes that exist, moreover, often take different forms. Some are HOV lanes, some are toll roads, and some are HOT lanes.

Even within these categories of managed lanes, moreover, the purpose and quality of the management can vary. HOV lanes can be stringent or lenient in their vehicle occupancy criteria, and strictly or loosely enforced. Priced lanes could be priced statically or dynamically, and governments might design the pricing to reduce driving, increase throughput, raise revenue, or maximize overall welfare. Most cordon charges in European cities were designed explicitly to reduce vehicle trips.[47] Many turnpikes in the United States are intended only to raise (but not necessarily maximize) revenue. These different goals imply different pricing structures, different transportation outcomes, different impacts on overall welfare (e.g. Light 2012). Small and Yan (2001), for example, show that tolls designed to maximize revenue will be substantially higher than those designed to maximize welfare, and that in some circumstances a revenue-maximizing priced lane will be sufficiently underused that it will deliver less welfare than a lane that remains free.

To my knowledge, there is literally no empirical evidence that examines the question of how a new HOT lane will affect regional VMT.[48] A very limited literature examines impacts along the managed corridors, but for all the reasons already mentioned, studies of regions and studies of corridors do not allow apples-to-apples comparisons. Nor is it possible to derive conclusions about HOT lanes from the larger regional studies. The small fraction of road capacity—even new capacity—represented by HOT lanes would be undetectable in a Duranton and Turner-style regression. Given these limitations, the argument in this section is drawn more from conjecture than a large body of empirical research.

That said, there is no compelling reason to think that the *regional* induced VMT effects of a newly-built HOT lane are different from the regional effects of new free lanes. This is not to say that HOT lanes overall are no different from free lanes. Along multiple dimensions, they are strikingly different. Most of those dimensions, however, are *corridor* specific. For travelers who use them, HOT lanes can improve speed and reliability. Because they offer that increased speed and reliability to travelers who are likely making high-value trips, they can also improve welfare. HOT lanes also, compared to building new free lanes, probably reduce air pollution. With respect to the specific question of how their construction induces regional VMT, however, they are probably little different than new general purpose lanes.

---

[47] Milan's cordon charge was designed to reduce vehicle trips and air pollution and increase bus speeds. Vehicle entrances into Milan's charging zone fell 30 percent after it was introduced (Percoco 2013). Stockholm's cordon charge similarly reduced traffic volumes; in 2012, traffic into the cordon area was about 20 percent lower than it had been in 2005 (Norgesson et al 2012). And London's congestion charge partly reduced vehicle travel, and partly changed its composition—away from private automobiles and toward exempt motorized vehicles like taxicabs and motorbikes (Santos 2008; Leape 2006).

[48] Anderson et al (2021), in a report for the California Air Resources Board, examine four lane expansions, one of which is an Express Lane, and find that it increased speed and flow on the expanded corridor, but could not say anything about region-wide induced VMT. Garcia-Lopez et al (2022) do control for roads being tolled in their region-wide regressions, and find that tolling is associated with less induced VMT, but the context is substantially different: the mean share of highways tolled in their sample is 25 percent, and in 77 of their 545 cities the entire highway network is tolled. Most of the tolls are not dynamic, and most also cover every lane—they have few situations like HOT lanes, where priced lanes sit adjacent to free lanes.

That conclusion might seem surprising, since use of these lanes is restricted. The key point is that restricting use may not mean reducing driving. A HOT lane is a combination of a carpool lane and a toll lane. A carpool lane offers a benefit (higher speed) to people willing to carpool. Can that speed benefit translate to less VMT? The answer depends in part on whether the carpool lane induces new carpooling, or simply offers a new benefit to people who would have carpooled anyway. Shifting existing carpools onto an HOV lane lowers the price of carpooling for those travelers, but is unlikely to reduce VMT—the same number of people and vehicles have formed carpools (Shewmaker 2012).

If the HOV lane induces new carpool formation, VMT reduction becomes more likely, since a vehicle leaves the road. Inducing new carpool formation, however, is a steep hill for an HOV lane to climb. An HOV lane can induce new carpools if the time benefit it offers exceeds the cost of carpool formation. Forming a carpool has upfront fixed costs (finding someone to carpool with) but more importantly ongoing costs in time and flexibility. The time costs are the daily effort of carpool assembly: the driver must circulate and pick up passengers, or the passengers must travel to meet the driver. The flexibility costs arise because the speed benefit is conditional on the carpool members staying together. As a result of that condition, individual members cannot easily change their plans. If one member of a two-person carpool leaves work early for a doctor's appointment, the other member forfeits the carpool lane benefit. The carpool can protect against this outcome by having more members, but more members increase assembly costs.

If the carpool's costs rival or exceed the benefits of faster travel in the HOV lane, carpools of the "ideal" type—composed of multiple people who would otherwise have driven alone—may not be formed. As an example: if the carpool lane subtracts 12 minutes from the freeway portion of a commute, but adds 12 minutes of driving around and picking up the carpool members, the carpool may not be worth it.[49] One result, in this situation, would be a low total number of carpools, meaning that both congestion and VMT in the carpool lane would be low, and speed could be quite high—the carpool lanes could be underused (e.g. Hanna et al 2017).

A second related possibility, when time saving are low relative to assembly costs, is that some new carpools will form, but form *within* households (e.g., a parent and children, or two roommates or spouses). Assuming people in the household are going to destinations near each other (they may well not be) the daily assembly costs of intra-household carpooling are relatively low (everyone is under one roof). Such carpools, however, are also unlikely to reduce VMT, particularly if the passengers are children or people who could not have otherwise driven.[50] Data from the 2017 NHTS suggest that across the US, about two-thirds of peak-hour shared vehicle trips are composed only of people from the same household, and slightly more than half of those are only an adult with a child.[51]

A final point is that even when the time savings associated with carpooling are large enough to induce carpool assembly, the process has diminishing returns. As more carpools are drawn into the lane by faster travel times, they congest the lane and reduce the speed that attracted them in the first place. Carpools lanes, in short, offer larger benefits when they are lightly used, and as their use grows their benefits fall. A successful carpool lane—one that induces a lot of carpooling—risks being self-undermining.

---

[49] The 12 minute figure is probably accurate for many lanes. Small et al (2006) report that time savings on the 91 Express Lanes at about ten minutes; Bento et al (2014) find similar peak time savings on the I-10 West, and data from Los Angeles Metro (2016) also show peak savings of 10-12 minutes across the agency's HOT lanes.

[50] It's possible for a same-house carpool to carry two people who would, absent the HOV lane, have traveled separately, but there is also a reasonable chance that the two people would have traveled together in any circumstance.

[51] I thank Julene Paul for these data. Note that they are not exclusive to HOV lanes. Relatedly, some fraction of carpool trips involves a passenger who does not personally own a car. In these cases the carpool benefit is access to an automobile, but such trips might increase VMT relative to the counterfactual (if the driver picks the passenger up, rather than the passenger walking to a transit stop).

HOT lanes became more prevalent, initially, as a way to address some of the problems affecting carpool lanes. Many HOV lanes were underused, and the HOT lane essentially sold excess HOV space to solo drivers willing to pay a toll. Introducing a toll to carpool lanes offers two advantages. First, it allows drivers to access the lane and its time savings without incurring the time and flexibility costs associated with assembling and maintaining a carpool. Second, the toll can prevent the lane from becoming a victim of its own success. If enough of the lane's traffic is paying a toll, dynamic pricing can prevent the lane from being congested. Where with a carpool lane time savings attract vehicles that reduce the time savings, with a toll time savings attract more vehicles that drive up the price, leaving the time savings intact.[52]

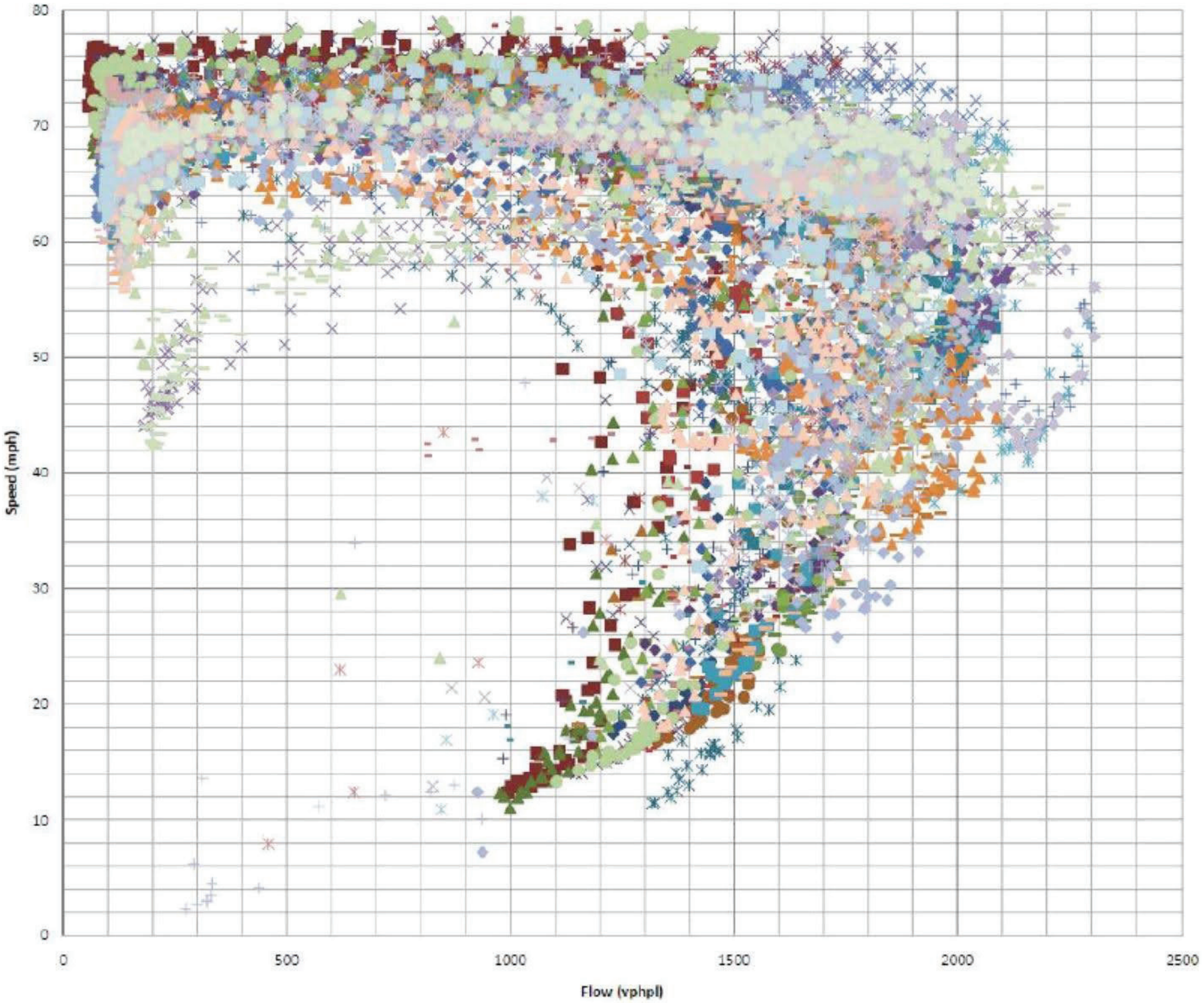
This advantage of pricing, however, introduces another implication of HOT lanes. Assuming they are priced to maintain performance and reasonably well-enforced, they can *increase* VMT along the priced corridor. This is so because prices can prevent a facility from tipping over into hypercongestion (Small and Chu 2003).

Hypercongestion occurs when the traffic density grows to a point where it reduces both speed *and* flow. Graphically, hypercongestion is captured in the backward bend of a standard speed/flow curve, as illustrated in Figure 3. The figure plots real speed and flow data from the 405 freeway in Los Angeles. The curve's upper left portion shows an intuitive negative correlation between speed and flow: very high speeds are only possible when the road holds relatively few vehicles. Cars need space between each other to safely travel fast. The curve's upper right portion, similarly, shows the same relationship in a different way: as more vehicles enter the road, average speed falls.

---

[52] The share of vehicles in a HOT lane that need to be priced to maintain savings probably depends on a number of factors, including the total quantity of vehicles using the facility. LA Metro data show that in most years between 45-55 percent of the Metro Express Lane trips are carpools, and the rest are priced (Metro 2018). These lanes are able to maintain free flow speeds, and note that they also have well-known enforcement problems. As of 2017, the agency believed up to 25 percent of carpools were single-occupant vehicles who had illegally toggled their transponders over to HOV (Nelson 2017). An additional number of vehicles, moreover, are SOVs that make illegal lane incursions.

Figure 3. Speed vs Flow on the 405 Freeway. From Marshall (2018)



In the bottom of the curve, however, this relationship flips. Now when speed is low, *flow* is low as well. Essentially, congestion has become so severe that the many vehicles crowded on to the road can barely move. That's hypercongestion.[53]

A road system won't spend hours in hypercongestion. Hypercongestion is a transient phenomenon that will occur sporadically over peak hours. It is important nevertheless, because it is the costliest form of congestion, in time and reliability. For our purposes, moreover, hypercongestion matters because it functionally reduces the capacity of the road. As we see above, when freeways become extremely congested their flow rates drop: congested lanes can move as few as 1,000 vehicles per hour, compared to almost twice that in off-peak times. Severe congestion, then, *reduces* VMT on the congested corridor. If dynamic pricing can prevent that severe congestion, and stop flow from "going around the bend" it will prevent not just congestion but capacity loss (Fosgerua and Small 2012, 2015; Fosgerua 2015; Kim 2019). VMT on a priced facility will thus be higher, compared to VMT on a similar lane that is unpriced.

The improved flow on HOT lanes has its own implications. First, HOT lanes will likely have lower emissions than unpriced lanes. Booribonsomin and Barth (2008) demonstrate that carbon emissions per mile fall sharply when the worst congestion can be averted, and Currie and Walker (2011) show that the same is true for both criteria pollutants and the health costs those pollutants impose on people living near busy roads.

HOT lanes should also, by virtue of the reliability they sell, serve different trips than newly-built free lanes. *A priori*, solo trips taken on Express Lanes should be higher value. Litman (2023) contends that travel induced by new unpriced capacity is likely to relatively low value: it would not have occurred if speeds were even modestly slower. This assertion is consistent with Duranton and Turner's (2011) finding that large shares of induced VMT are household travel that would not otherwise have occurred.

HOT Lane trips, in contrast, per Small et al (2006) and Bento et al (2020), are more likely to have been redistributed from other lanes: they are among the more important trips drivers are making. In part this redistribution may result because HOT lane drivers have higher average values of time than others (i.e., for any given trip, richer people are more likely to use a HOT lane). But theory and evidence also both suggest a large role for *intra*personal variance in the value of time—everyone, at some point, is in a hurry, and when they are in a hurry they are more likely to pay for reliability. No one goes to the airport simply because HOT Lanes exist. But a middle-income person who is late for a flight might go *to* the airport in a HOT Lane, then economize and stay in a free lane when traveling from the airport back home. Bento et al (2020), in their examination of the I-10 Express lanes, present some evidence for this idea. They find that "urgency"—a trip-specific value—accounts for up to 87 percent of the willingness to pay for LA Express lanes. This combination of increased reliability for more important trips suggests that HOT lanes can be welfare-enhancing; the lanes deliver performance when travelers need it most.

---

[53] The extent of hypercongestion remains a source of some debate. In part this debate arises because Vickrey's (1969) influential bottleneck model of congestion leaves no room for hypercongestion. The bottleneck in Vickrey's formulation had a constant capacity, and congestion costs were purely a function of diminished speed (and congestion costs, were, in turn, half the total costs of congestion, the other half being scheduling costs). Anderson and Davis (2018) examined actual bottlenecks in California and empirically confirmed that they do not become hypercongested: when vehicles enter the bottleneck speed falls but capacity remains unchanged. Vickrey himself, however, thought hypercongestion was real (e.g., Vickrey 1969, Vickrey 1973). The bottleneck model was tractable because it allowed for endogenous departure times, but Vickrey, in the same article where he introduced the bottleneck model, went on to note that empirically a lot of congestion did not occur at bottlenecks. In congested urban areas, so-called "triggernecks"—situations where a queue on one route interferes with traffic taking another route—are common (Southern Californians can picture the 101/110 interchange; anyone can picture traffic backed up across a downtown intersection). Triggernecks result in lost capacity as well as lost speed. (Anderson and Davis explicitly note that their findings may not hold for triggernecks). For discussion see Small (2015), Fosgerua (2015), Kim (2019), as well as various "bathtub" models of urban congestion (Arnott 2013).

Crucially, people allowed to use HOT lanes *without* paying may not be making urgent trips. Bento et al (2015) show that when single-occupant low-emission vehicles are allowed in HOT lanes without paying, they rapidly congest those lanes. These lanes, to drivers of clean air vehicles, are just additional general purpose space. The HOT lanes may therefore deliver more benefits as the share of paying vehicles rises.

For all these reasons, a newly-built HOT lane differs strikingly from a newly-built general purpose lane. Assuming it is managed and enforced well, the HOT lane is likely to perform better, contribute less pollution, and carry trips that are subjectively more important to the people making them. And while the HOT lane may carry more VMT than a similar general purpose lane, much of that VMT may be redistributed. It could pre-existing carpools that switch to the lane, or people with high trip- or person-specific time values who switch to the lane.

None of these facts, however, have much bearing on the policy question at hand: how much VMT will a HOT lane induce? Here we return to the crux of this section's argument. All the impacts described above are occurring on the HOT lane itself. If the vehicles in the HOT lane, be they carpools or toll-paying autos, would have otherwise been in general purpose lanes or on nearby arterials, the obvious question is what happens to that road space they leave behind. Do other drivers backfill into that space? A long literature, going back to Downs (1962) suggests they will. That road space is unpriced capacity, now moving faster because fewer vehicles are on it. This should induce new VMT in the same manner as any other segment of unpriced capacity does.

In short, as long as we think most unpriced capacity has an elasticity of 1 (see Section I), the HOT lanes should have a similar elasticity across the network. One could even argue that at peak hours they might have a higher elasticity. The HOT lanes will move more vehicles than a free lane, meaning they will pull more vehicles *off* free lanes, and allow more backfilling to occur. But we needn't push that point, since it probably involves some additional assumptions that could be quarreled with. For now it is enough to say that a new HOT lane will probably have similar network-level effects to a new free lane.

Before moving on, I want to offer a clarifying point. This report is about induced VMT and managed lanes specifically, and not about the larger issue of how California chooses to "count" VMT impacts of all kinds. Nevertheless, some readers might at this point be wondering if the Fundamental Law of Road Congestion is being selectively applied. My argument is that the state is correct when it says new HOT lanes have an elasticity of 1, because when drivers leave free lanes for the HOT lanes, other drivers will backfill into that free space and add VMT. Scholars generally accept, however, that drivers will do the same if an agency adds *rail* service. A new train might pull some drivers off a highway, but other drivers will replace them, and the level of congestion will return to what it had been. This reasoning goes as far back as Downs (1962), and as I mentioned above, Duranton and Turner (2011) offer some evidence to back it up. In light of theory and evidence, should we worry that the state is unfairly singling out roads? Perhaps light rail will also induce VMT, and should thus require mitigation.

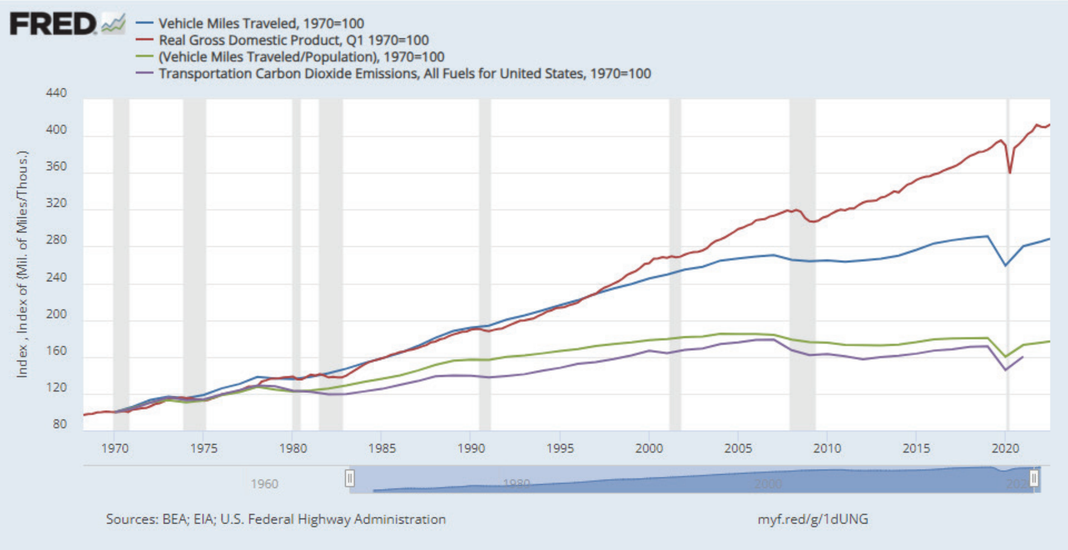
This concern is unfounded. The difference lies in the fact that the new HOT lanes involve constructing new road capacity. We can illustrate thusly. Suppose that instead of building two new priced lanes we build a rail line running parallel to the freeway. The rail line, like the priced lanes, peels some drivers off the road, and frees up road space. That road space attracts more driving. The key is that when drivers move from the road to the train, VMT *falls*—car trips have become transit trips. When new drivers fill in for the cars that left, VMT goes back up. When a driver leaves a general purpose lane for a managed lane, in contrast, VMT does *not* fall—it just shifts. After that shift, *new* VMT is added, via backfilling in the general purpose lane. As a result, the new HOT lane increase total VMT. The train, meanwhile, doesn't *reduce* VMT (which should cast some doubt on its utility as a mitigation), but is also unlikely to *increase* it, because (at least roughly speaking) the train can only induce as much VMT as it reduced in the first place.

# III. What VMT Does and Doesn't Tell Us

At different points in Section II we observed a tension between measures of VMT and measures of system performance. An underused carpool lane carries few vehicles. It has little VMT, and little congestion, but along its most important dimension (is anyone carpooling?), it doesn't work. Similarly, a hypercongested road carries less VMT than a moderately congested one, but hypercongestion is costly in almost every way: time, stress, reliability, pollution and so on. The lower VMT on this facility implies system dysfunction. Both these examples suggest that VMT may not, by itself, be a very meaningful measure.

One could argue, reasonably, that both the examples above are corridor-specific, and that what matters is regional VMT. Fair enough. And one could then note that on a regional scale VMT is correlated with many problems: a strong association exists between VMT and GHG, VMT and criteria air pollutants, and VMT and motor vehicle deaths. True indeed. But VMT is *also*, at the same larger scale, correlated with many *benefits*: household income, employment, and economic growth. Here a skeptic could pounce, and observe that the relationship between VMT and economic growth has been weakening over time. That's correct. But the relationship between VMT and transportation carbon emissions has also been weakening over time (Figure 4). It isn't clear why one of these weakening relationship should matter more than the other.

**Figure 4. VMT. GDP and Transportation Carbon Emissions**

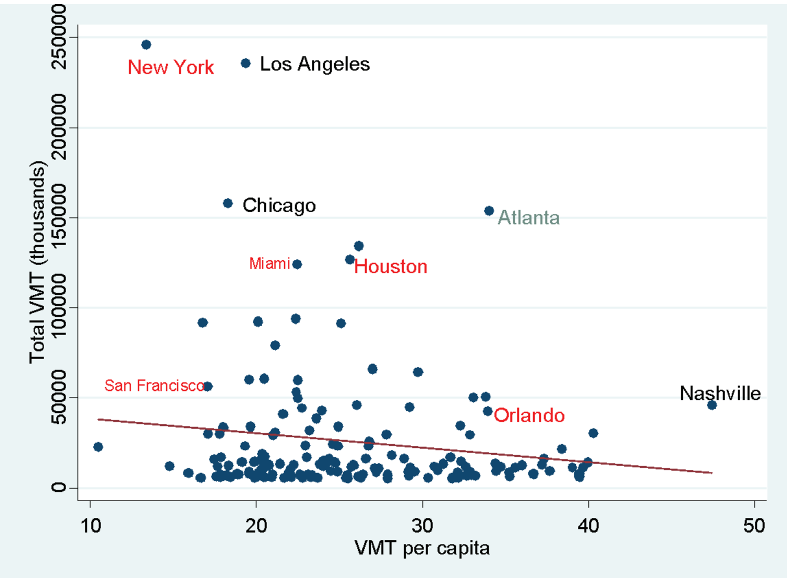


Sources: BEA; EIA; U.S. Federal Highway Administration

myf.red/g/1dUNG

Figure 4 shows that transportation carbon emissions remain closely tied to VMT *per capita*, but VMT per capita and VMT are quite different measures. Their relationship to each other, like the relationship between VMT and carbon (or VMT and economic growth) has weakened over time. It is also the case, as we see in Figure 5, that at any given time VMT and VMT per capita are only weakly correlated by place. Across large US urban areas, the two variables exhibit if anything a mild negative relationship (more VMT per capita is associated with less VMT overall). But really there is no clear relationship, and if anything the graph suggests that places with very different transportation cultures can have quite similar measures of VMT. New York is a place with high VMT and VMT low per capita. But so is Los Angeles.

**Figure 5. VMT and VMT per Capita, 50 Large UZAs, 2021. Source: Federal Highway Administration**



So why focus on VMT, and consider it a pure cost? In my reading and discussions with state officials, I have encountered four arguments for doing so: 1) many other states do it 2) VMT's costs cannot be separated from its benefits, so managing the costs requires reducing VMT outright, 3) VMT cannot have benefits because driving is a derived demand, and 4) VMT causally suppresses the use of other modes. I will discuss each in turn.

## 3.1 VMT is a Common Metric for Environmental Impact

This argument isn't persuasive. It's true that other states have committed to reducing VMT, but it is also wholly irrelevant. In transportation policy as in other arenas, prevalence should not be confused with wisdom. It wasn't that long ago, after all, that many places enthusiastically widened their roads and called doing so a best practice, happily convinced that induced travel was a myth. The fact that other places have a policy doesn't make that policy correct. Believing otherwise is known as the "bandwagon fallacy." Policies must stand or fall on their merits.

## 3.2 VMT's Costs Cannot be Separated from Its Benefits

A foundational idea in the economics of social cost is that if an activity imposes negative externalities, public policy should target those externalities specifically, rather than the activities that create them. The logic behind this idea is that most activities have some benefit (otherwise no one would engage in them), and that as such the least costly way to reduce social harm is to let individual people weigh the costs of that harm against the benefits they receive.

Electricity, for example, has both benefits and costs, with the most notable cost being the air pollution emitted when electricity is generated. Society *could* fight that pollution by simply restricting electricity use. But doing so would be neither efficient nor fair. While some people use electricity wastefully, and could cut back at little cost, for others additional extra increments of electricity are extremely important. An across-the-board reduction of electricity use lands more heavily on a person relying on a ventilator than on a person who tends to overdo it with the AC. For that matter not all kilowatt hours of electricity are equally harmful. Some are generated with highly-polluting coal, others with cleaner solar or hydro power. Categorizing all electricity as pollution would blur this distinction, and reduce the incentive of producers and consumers to seek out cleaner sources.

To avoid this outcome, anti-pollution policy in the electricity sector generally tries to regulate the *specific emissions* associated with electricity, rather than just limit electricity use. Cap and trade programs, for example, place caps on pollutants like sulfur dioxide, not on the production or consumption of electricity per se. In doing so they benefit producers who can generate power with less pollution, and reward innovations (like improved solar technology) that deliver more of the benefits with fewer of the costs. These incentives to pollute less and innovate more are only possible, however, because the policy explicitly separates the productive activity (electricity consumption) from its harmful by products (pollution).

California argues, essentially, that such a separation cannot be accomplished with VMT, and that as a result the best course of action is to treat “traffic as the pollutant”:

> *If we were to list all of the environmental effects of traffic, could we dispense with VMT and address those effects separately? Probably not. Even if we could enumerate every environmental impact of traffic, dealing with them separately would be like trying to address the public health effects of smokestack emissions one by one ... it’s better to reduce the emissions in the first place. Thus, as with other pollutants, our goal is to reduce traffic, not simply try to accommodate it as we used to.*[54]

This short paragraph contains much to unpack. A first observation is that the metaphor is mixed; it confuses the activity for the externality. Caltrans compares VMT to a smokestack. A smokestack, however, is a device that does *nothing* but emit pollution. No one builds a smokestack for its own sake, or values the output a smokestack produces. Smokestacks instead expel the byproducts of goods or services that people *do* value (like electricity). A smokestack in this way is less analogous to VMT and more analogous to a vehicle’s exhaust pipe. And just as it is true that we’d rather not deal with all the pollutants in a smokestack one-by-one (we’d rather just install a scrubber) it’s also true that we’d rather not deal with everything coming out of a car’s exhaust pipe one-by-one (it’s better to install catalytic converters). But none of that tells us anything about *VMT*. VMT doesn’t come out of an exhaust pipe any more than electricity comes out of a smokestack. The sequence here is exactly backward: smokestack emissions are the residual of electricity production, and tailpipe emissions (and other externalities) are the residual of VMT. “Reducing emissions,” as such, *can* mean reducing electricity use or VMT, but doesn’t have to. It’s also possible, and in some cases optimal, to find ways to drive or consume electricity that emit less pollution.

One counterargument here might be to reiterate that driving isn’t like electricity. Smokestack scrubbers might let more electricity get produced with less pollution, but driving and its negative effects cannot be separated—there is just no way to get more of driving’s benefits while controlling its costs. But this argument doesn’t work, theoretically or empirically. We can start with theory: situations where costs and benefits appear inseparable are hardly foreign to the study of externalities. In these situations the theory is clear: it’s still better to write policies that target the externality. This is so for the simple reason that circumstances change, and well-written regulations can in fact accelerate those changes. A rule that draws a distinction between a productive activity and its externality encourages people to *find* ways to separate the two. A rule that caps or restricts the activity itself, in contrast, offers no such

---

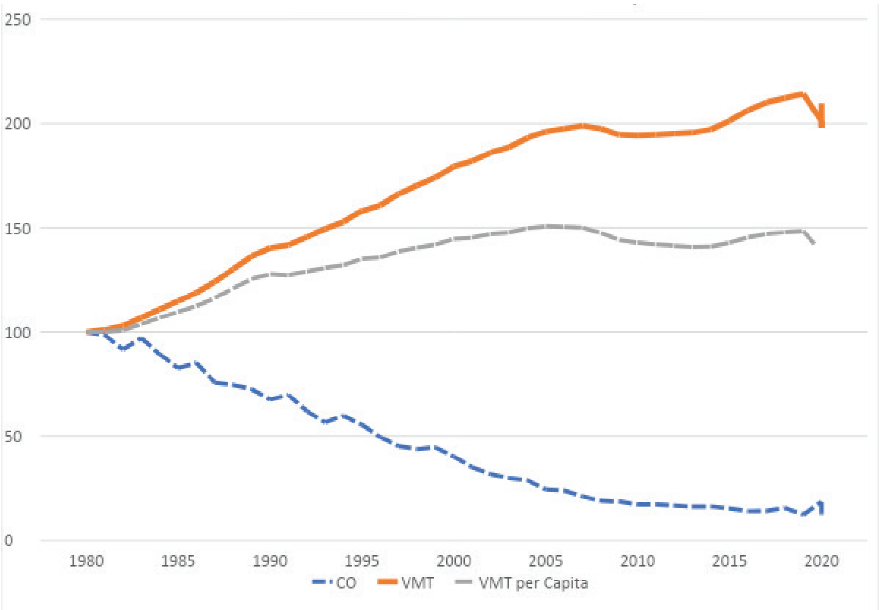
[54] Caltrans, no date.

incentive. (Regulations that target nitrogen dioxide, for example encourage investment in solar electricity production. Regulations that target electricity don't).[55] It is better to target pollution, or congestion, than to target VMT.

But what if separating VMT from its externalities really is impossible? No harm is done. The rule will simply reduce VMT. The beauty of targeting the externality is that it allows the optimal outcome, but doesn't preclude the next-best outcome.

In practice, however, there is little reason to believe that the costs of VMT cannot be separated from its benefits. We have already seen that VMT's correlation with transportation GHGs is falling. A still more dramatic trend is the weakening relationship between VMT and criteria pollutants. From 1970 to 2005, VMT rose over 150 percent, and criteria pollutants fell almost 50 percent, largely due to catalytic conversion and other emission standards (Moore et al 2005). Figure 6 below shows that, from 1980 to 2022, VMT grew by over 100 percent while carbon monoxide emissions fell almost 90 percent:

**Figure 6. Carbon Monoxide Emissions and Vehicle Travel, 1980-2022**



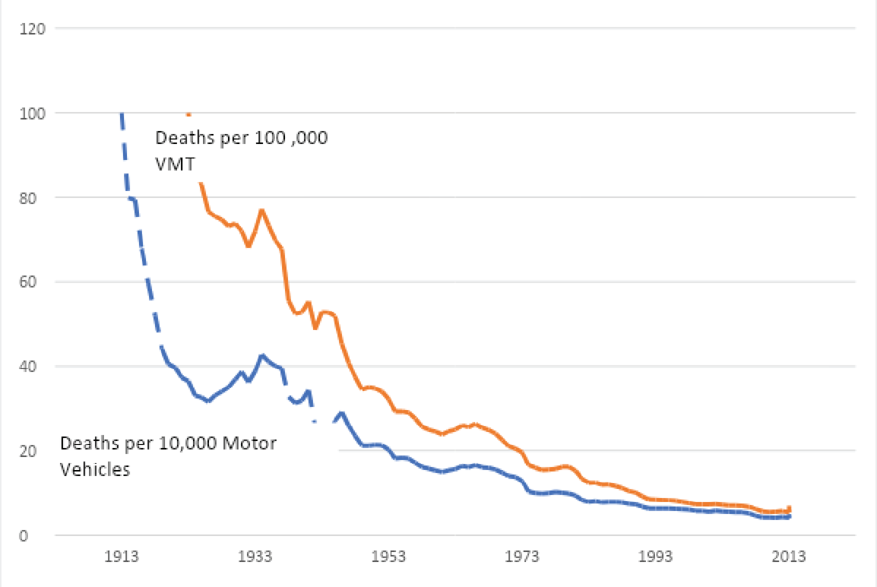
*Sources: US EPA. US BEA, US Federal Highway Administration*

---

[55] Here is a relevant but possibly gross example: livestock flatulence is a large source of greenhouse gas emissions. This may seem like a case where the productive activity and the externality are inseparable. Cows burp and fart. But policies that target methane emissions can reward farmers who feed livestock different foods that reduce bovine gas, and/or who invest in equipment that captures the gas and prevents it from entering the atmosphere. Laws simply limiting the size of cattle herds would not (Kwok 2023).

The same weakening relationship can be seen between VMT and deaths from motor vehicles (Figure 7):

**Figure 7. Traffic Fatalities and Motor Vehicle Use, 1913-2021**



*Sources: National Center for Health Statistics, National Safety Council, Federal Highway Administration*

The point of these graphs is *not* that driving is devoid of costs, or that it requires no policy intervention. Neither should anything in these graphs suggest that there is no value in creating places where people don't need to drive as much as they do now. Driving remains a substantial and serious source of air pollution, and remains one of the most dangerous things the average American does on a regular basis.[56] The point is only that driving's substantial costs are plainly *not* inseparable from driving itself. One viable way to reduce the harms associated with driving is, again, to make it easier for people to drive less. But there is no basis for the idea that it is the *only* way to do so, and that we must therefore simply treat VMT as pollution. Some of the great policy triumphs of the 20th century involved making driving safer and cleaner than it had been before, letting us capture more of driving's benefits while incurring fewer of its costs. A policy declaring VMT a pure cost, had it been in place decades ago, may well have prevented or impeded these policy victories.

## 3.3 VMT Actually Has No Benefits

This is perhaps the most unusual argument for considering VMT a pure cost. It is rooted in the idea that VMT is a "derived demand." Absent the occasional pleasure drive, people do not travel by automobile because they want to. They do so because driving allows them access to the things they really want: getting to work, visiting family, going

---

[56] In recent years pedestrian deaths have risen, and this fact is obscured in the overall fatality data. But the pedestrian death increase has little to do with the amount of VMT (which has been flat to falling in many places) and more to do with increasing driver distraction, and the increasing front-end height of vehicles. See Tyndall (2024).

to the coffee shop or gym, etc. Driving is not what we want, but an instrument for satisfying actual wants. Hence the demand for it is "derived."

Scholars sometimes argue about the extent to which travel really is a derived demand (e.g. Mokhtarian and Salomon 2001), and for that matter some wonder if the concept of a "derived demand" is even coherent.[57] We needn't concern ourselves with those debates. We can stipulate, for now, that most demand for VMT is derived. Conceding as much does not convert VMT into a pure cost. Nothing about the demand for a good being derived implies that the good is devoid of benefits. It may imply that if a superior substitute arose, people would happily switch (who would drive if they could easily teleport?), but that's just another way of saying that in a different world people might value different things. It does not change the fact that in the world we have, the good has benefits.

Consider that if VMT is derived demand, so too is electricity. Just as people rarely wake up yearning to drive, they also rarely wake up hankering to consume electricity for the pure satisfaction of doing so. Electricity instead is instrumental: it powers many other valuable aspects of people's lives. But it would be unusual to conclude, as a result of that fact, that electricity has no benefits. Indeed, recall that Caltrans, in justifying its classification of VMT as pollution, did *not* analogize VMT to electricity. It analogized it to a smokestack. But a smokestack doesn't produce electricity. It emits some of the negative externalities of electricity. Electricity, a derived demand, has costs and benefits. VMT too.

We can make the same point without leaving the realm of transportation. If VMT is a derived demand, public transit is also a derived demand. Like driving, transit is useful only insofar as it lets us access other goods and services. Is California willing to say that passenger-miles of transit travel, like vehicle miles of travel, have no benefit? It is not. California promotes transit use while discouraging driving, which does not make much sense if instrumental goods are presumed to have zero benefits.

Of course, when California promotes transit use and discourages driving, it does so with an argument that sits on much firmer ground: on average, driving has larger negative externalities than transit. To give just one example, the average transit trip in the United States emits 55 percent fewer GHGs per person mile than the average automobile trip (National Academies 2021).

As long as it makes sense to draw externality-based distinctions *across* modes, however, it should also make sense to draw such distinctions *within* them. If we can say that a train ride has fewer social costs than a car trip, we should also be able to say that some train rides, and some car trips, are more or less costly than others. Indeed, the variance in externalities within modes is often as large, or larger, as the variance between them. The average person mile of transit mile emits 55 percent fewer GHGs than the average person-mile in a vehicle, but that average is heavily biased by New York City, which accounts for over 40 percent of American transit trips and carries most of its passengers in electrified rail cars that have high average occupancy. Many American bus-miles, because they are on low ridership coverage routes, emit more GHGs than the transit average, and some emit more per mile than a standard automobile trip. The emissions gap between low- and high-emission person-miles of transit is much larger, as a result, than the average gap between transit and automobiles.

---

[57] The worry here is that many goods can, with a little effort, look like derived demands. It's easy to say that driving is only a derived demand, and that the "real" demand is to (for example) get to work. But many people don't want to work for its own sake; they feel like they have to work to make money. By that logic getting to work is also a derived demand. And yet work only gets you money, and money, unless you are Scrooge, isn't valuable for its own sake. It is valuable insofar it can be spent on other things—some of which themselves look like derived demands (health care, food basics, etc.). At a certain point, enough demands are derived that the distinction starts to lose meaning.

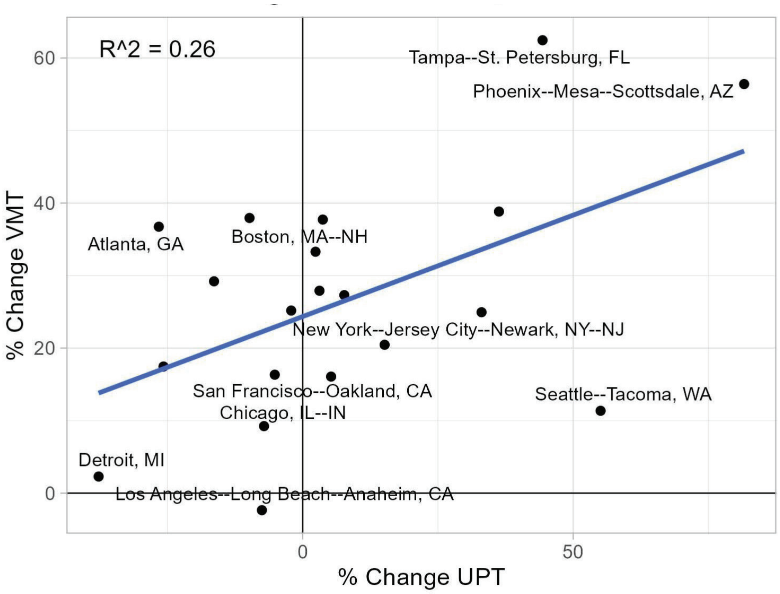
So too is the gap between high and low-emission VMT. The externalities produced by an auto trip, similarly, can vary greatly depending on the time of day, the type of vehicle, the vehicle's occupancy, and the existing level of traffic. Adding VMT at congested times creates more delay, more crash risk, and most of all more pollution (Fosgerua and Small 2012; Currie and Walker 2011; Edlin and Picar-Mandic 2006; Bento et al 2014; Knittel et al 2016; Booriboonsomin and Barth 2008). Bento et al (2014) estimate that adding a single unpriced vehicle to a Southern California HOT lane at 2 AM each day creates zero congestion externalities. Adding the same vehicle 7 AM, in contrast, creates $4,500 in annual externalities. A policy that simply classifies all VMT as pure costs risks eliding these differences.

## 3.4 VMT Suppresses Use of Other Modes

A final argument for classifying VMT as a pure cost is that driving has impacts beyond pollution and crashes: it actually makes walking, cycling and transit use less feasible. "Higher VMT," Caltrans observes, "makes it harder for people to travel by walking, biking, and transit. The traffic itself and the infrastructure needed to accommodate it can make other modes less viable, e.g., adding additional wait times and travel distances for pedestrians, including those using transit, or adding to the 'level of traffic stress,' which discourages active travel."

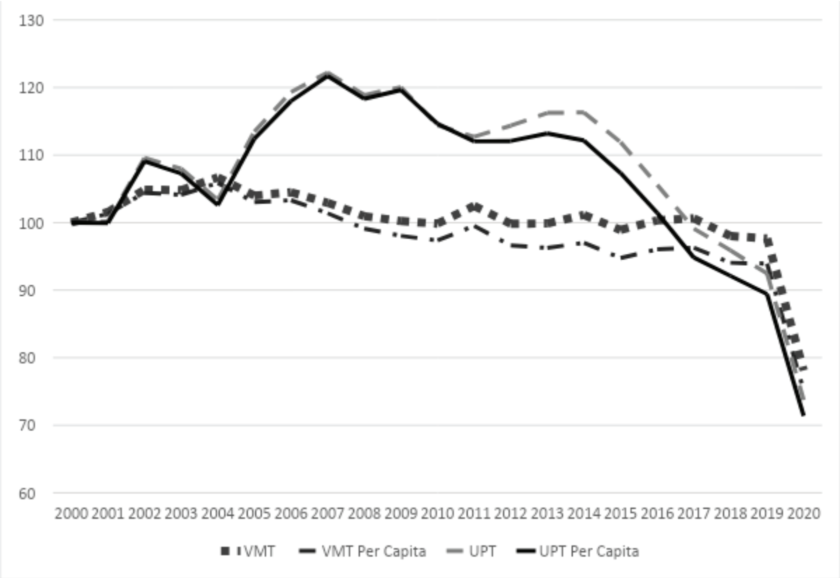
Is this correct? If it is, we should see that, all else equal, when VMT falls transit use rises, and vice-versa. At the regional level, little evidence suggests this is the case. Figure 8 plots the change in VMT against the change in unlinked transit passenger miles for 20 large urban areas between 2000 and 2019. The relationship overall is modestly positive. VMT grew almost everywhere, transit fell in some places but not others.

**Figure 8. Driving and Transit Use, 2000-2019**



The graph admittedly shows only a simple bivariate relationship. But more comprehensive analyses show that American transit ridership was flat or falling during the period when VMT was in decline, even as gas prices and transit investment were on the rise (e.g. Manville et al 2017). Note too that in Figure 8 Los Angeles is the only urban area where VMT declined from 2000-2019. Transit fell alongside it. Figure 9 examines LA in more detail. As context, from 2000 to 2019, LA invested heavily in new transit service. It opened multiple new rail lines even as nominal gas prices soared. If VMT suppressed transit use, in these conditions we should see at least some signs of a transit resurgence. We don't (Figure 9). Transit use did not steadily rise as transit declined. Instead it rose briefly as driving dipped (coinciding with the steepest increase in gas prices), but then if anything moved in sync with VMT (the correlation between VMT and transit trips during this time is a modestly positive 0.3).[58]

**Figure 9. VMT and Transit Ridership, 2000-2020, Los Angeles (2000=100)**



Is a simple graph wholly dispositive? No. But a plausible explanation for why transit use didn't rise even as driving fell is that L.A.'s *built environment* didn't change. Even with less VMT Los Angeles is a region of wide roads, intrusive freeways, and abundant parking, all of which conspire to push destinations apart and undermine non-auto ways of moving around.

Let's return to the Caltrans statement above: "The traffic itself and the infrastructure needed to accommodate it can make other modes less viable." Embedded in this statement are two different assertions. The first is that traffic by itself leads people to not use other modes: that VMT makes people walk, cycle or use transit less. The second is that *auto infrastructure* has these effects. While the first assertion isn't impossible, it probably holds only in very particular situations. The second assertion, in contrast, is eminently plausible. Automobile infrastructure almost certainly causes both higher VMT and lower use of other modes. Automobiles, historically, have exhibited network externalities: the more people use cars, the more useful they become—and the *less* useful other modes become. But the mediating mechanism in that relationship has been the built environment. Cars gain their relative advantage over other modes because cities oriented the landscape around automobility, and doing so penalizes other ways of moving around (King et al 2022).[59]

---

[58] The graph shows both modes plunging in 2020, but correlation doesn't change if 2020 is removed.

[59] To reiterate a point made in the introduction, King et al (2022), in describing the network externality created by auto infrastructure, argue that absent changes to that infrastructure it can be welfare enhancing to have *more* VMT, through targeted programs that carless low-income people drive more.

Imagine an urban neighborhood split by a wide busy road with infrequent intersections. Many people on one side of this road, who need access to destinations on the other, may be reluctant to make those trips by walking or biking. The road is dangerous to ride along, and finding a safe place to cross it requires traveling well out of the way. So most people drive instead. This scenario appears to be a clear-cut case of VMT—the sheer volume of traffic—causing people nearby to walk less and drive more. And it's probably true that if the amount traffic on the road fell to zero, more people would walk or bike nearby.

The chances of VMT falling to zero, however, are probably themselves close to zero. Suppose more realistically that the road suddenly carried 20 percent less VMT. Fewer vehicles would be on the road, but because the vehicles were fewer each would probably be moving faster. Would the road and its environs now be 20 or even 10 percent more appealing for biking and walking? A 20 percent reduction is a large drop in VMT, but it's hard to imagine that non-auto modes would flourish as a result.

Now imagine the city makes a substantial policy intervention. It widens the road but also *buries* it. All motor vehicles are now in a tunnel, and the city converts the land atop the tunnel, which had once been the busy arterial, into a grassy park with bike and pedestrian paths. What happens? VMT will rise, because the road is wider. But walking and cycling will probably rise as well. There are still many fast moving cars, but those cars no longer threaten or interfere with other modes. It was the infrastructure, not the VMT, that was creating problems.

That situation might sound fanciful, but it isn't unheard of. It essentially describes Boston's Big Dig. By replacing the Central Artery/Expressway with a tunnel that was wider than the Expressway had been, and then adding an additional tunnel across Boston Harbor, the Big Dig induced VMT. But there is little doubt it also increased walking. The project removed a double-decked freeway that had sliced whole neighborhoods apart, and installed a manicured greenway that helped stitch them back together. Changing the infrastructure increased both driving and walking.

We can consider other examples as well. Figure 10 shows a neighborhood in West Los Angeles, and illustrates what are known as the "severance and barrier" effects of freeways. There is a park on the top side of the picture, less than 500 feet from the neighborhood on the south side. But the 10 freeway cuts through the area and blocks many surface streets, so anyone wishing to walk from the neighborhood to the park has to travel seven-tenths of a mile—3700 feet rather than 500 feet—and the walk is unpleasant.[60] This is a good example of an auto-oriented, high VMT environment reducing active travel.

---

[60] See Millard-Ball et al (2022) for a more thorough analysis of barrier effects from highways.

**Figure 10. Barrier Freeway Effect**

*Source: Google Maps*

The key, however, is that it is probably the auto-orientation — the presence of the freeway itself — and not the high VMT, that matters. It is far from obvious that more people would walk to the park if the freeway carried fewer vehicles. For instance: Is the pedestrian experience really better at 8 AM, when the hypercongested 10 can move only 1,500 vehicles per lane mile at an average speed of 10 mph, than at 1 PM, when the freeway moves 500 vehicles per lane mile at an average speed of 75 mph? Or is the issue mostly the physical presence of the 10 itself, which blocks the most direct walking routes? To invoke the Big Dig example again, if the 10 were buried, it could be carrying 2000 vehicles per lane mile—maximizing its VMT—and the walk from the neighborhood to the park would be both shorter and more enjoyable.

No one, sadly, is going to bury the 10 anytime soon. But these examples allow us to rethink Caltrans' admonition that we should no longer "accommodate VMT." Is this right? If we are concerned about the impact on other modes, then certainly we should not accommodate VMT *by building new road infrastructure*. But one of the most efficient ways to avoid building new roads is to dynamically price the roads we have. As discussed in the previous section, such pricing functionally increases the capacity of the road (which could be seen as "accommodating VMT") *without* making built environment interventions that undermine transit, cycling and walking. Indeed, pricing a road can increase the VMT on that road *and* increase transit use.

*

To summarize: the case for calling VMT pollution is quite weak. That weakness, however, is probably of little use to anyone who wishes to expand freeways. The state probably errs when it asserts that VMT causally suppresses other modes. It is more accurate to say that auto infrastructure simultaneously increases VMT and decreases walking, biking, and transit use. But while correcting this error might absolve VMT, it obviously does not build a case for more roads. VMT may not be the problem, but road space is.

Perhaps the best a proponent of new lanes can do, using this reasoning, is suggest that two additional lanes on an already existing freeway would only marginally exacerbate severance and barrier problems. The damage of expanding a freeway that has already split a neighborhood, in other words, is much less than the damage imposed when a new freeway carves up a neighborhood that is currently intact.

This logic isn't terribly convincing in its own right; it essentially says that since walking to the park is already miserable, there is little harm in making it more so. But the larger issue is that this reasoning still ignores the larger underlying inefficiency. It remains inescapable that—as discussed—most congested facilities are suffering, as a result of their congestion, from the equivalent of large losses in road capacity. The most efficient way to solve that problem is to restore that capacity with performance pricing—or at least *try* to do so—not to build entirely new lanes.

# IV. Conclusion

My goal in this report has been to take the areas of disagreement that surfaced during the expert panel and examine them in more detail, in the hope that further background and explanation might help later discussions arrive at more consensus. I have offered opinions here that I am sure not everyone will agree with. Specifically: I think that right now the unit elasticity estimate is the best available estimate of induced VMT, for both general purpose and HOT lanes. Conceptually, I do not like the idea of formally classifying VMT as a cost, and I think the state would do well to revisit this practice. In practice, however, the flaws in this approach would not change the wisdom, on performance or welfare grounds, of building new road capacity.

In part I have advanced these opinions because I believe they are correct. But I have also done so because advancing these opinions may give the debate around these issues more specificity. Specificity allows us, rather than simply saying we believe or disbelieve a given model or elasticity, to address particular strengths and weaknesses in these methods and the arguments underlying them. With luck, then, any disagreement I provoke will be productive, and enable those whose views differ to zero in on particular errors and omissions in the case I have made, and move the overall discussion forward.

Let me close with a recommended next step. As I said from the beginning, no empirical research specifically asks how HOT lanes induce travel at a regional level. This absent empirical base arises largely because HOT lanes remain quite rare, and until more are built, it may be difficult to fully resolve that problem. In the meantime, however, researchers could replicate some of the recent induced travel studies, like Hymel (2019) or Duranton and Turner (2011) but separate new unpriced road capacity from new HOT lane capacity. Doing so would create a new independent variable of interest, and we could compare the elasticity associated with HOT lanes to that associated with unpriced lanes. This idea has limitations—in earlier years there were no HOT lanes, so a panel could not go back as far, and even in recent years most places will have very little HOT lane mileage, so estimates about a HOT lane elasticity are likely to be noisy. Nevertheless, this may be the only step available right now. If mobile phone or GPS data improve, they may offer another way to study this question. But we would need to not just resolve the selection questions that currently hang over these data, but also let enough years pass to allow estimation of a long-term elasticity. As such these data, while they hold potential, do not offer an immediate research opportunity.

# References

Alberini, Anna, Lavan Teja Burra, Cinzia Cirillo, and Chang Shen. 2021. Counting vehicle miles traveled: What can we learn from the NHTS? Transportation Research Part D: Transport and Environment, 98.

Alexander, Soctt. 2014. Beware Isolated Demands for Rigor. Slate Star Codex. August 14.

Anderson, Michael and Lucas W. Davis. 2020. An empirical test of hypercongestion in highway bottlenecks. Journal of Public Economics 187, 104197.

Anderson, Michael, Lucas Davis and Leila Safavi. 2021. Estimating Induced Travel from Capacity Expansions on Congested Corridors. Report 18RD022 to the California Air Resources Board.

Angrist, J.D. & Krueger, A.B. 2001. Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments. Journal of Economic Perspectives, 15(4), 69-85.

Arnott, Richard. 2013.A bathtub model of downtown traffic congestion, Journal of Urban Economics, 76:110-121,.

Arnott, Richard and Kenneth Small. 1994 "The Economics of Traffic Congestion," American Scientist. 82: 446-455

Associated Press. 2007. End sought bridges to nowhere. New York Times. September 22

Banister, D. (2008) The big smoke: congestion charging and the environment. In Harry Richardson and Christine Bae (Eds) *Road Congestion Pricing in Europe: Implications for the United States*. Cheltenham: Edward Elgar.

Baum-Snow, Nathaniel. 2007. "Did Highways Cause Suburbanization?" Quarterly Journal of Economics, 122(2): 775–805.

Baum-Snow, Nathaniel. 2010. "Changes in Transportation Infrastructure and Commuting Patterns in US Metropolitan Areas, 1960-2000." *American Economic Review*, 100 (2): 378-82.

Beevers, Sean and David Carslaw. 2005. The impact of congestion charging on vehicle emissions in London. *Atmospheric Environment*. 39:1-5.

Bento, Antonio, Daniel Kaffine, Kevin Roth, and Matthew Zaragoza-Watkins. 2014. "The Effects of Regulation in the Presence of Multiple Unpriced Externalities: Evidence from the Transportation Sector." American Economic Journal: Economic Policy, 6 (3): 1-29.

Bento, Antonio, Kevin Roth and Andrew Waxman. 2020. Avoiding Traffic Externalities: The Value of Urgency. NBER Working Paper 26956.

Booriboonsomin, K and M. Barth. 2008. Impacts of Freeway High-occupancy Vehicle Lanes on Vehicle Emissions. Transportation Research Part D: Transport and Environment. 13:112-25.

Borck, Rainald and Philipp Schrauth.2019. Population Density and Urban Air Quality. Cesifo Working Paper 7629, Munich.

Caltrans, no date. SB 743 at 10: The Environmental Effects of Traffic. https://dot.ca.gov/programs/esta/sb-743/resources/10-years-sb743

Card, David & Krueger, Alan B, 1994. "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania," American Economic Review, 84(4): 772-793.

Cervero, Robert and Mark Hansen. 2002. Induced Travel Demand and Induced Road Investment: A Simultaneous Equation Analysis. *Journal of Transport Economics and Policy*. 36 (3):469-490

Cervero, Robert. (2003), "Road Expansion, Urban Growth, and Induced Travel: A Path Analysis," Journal of the American Planning Association, Vol. 69, No. 2 (www.planning.org), pp. 145-163.

Chang, Huibin, Debarshi Indra, Abhradeep Maiti. 2023. Metropolitan area heterogeneity and the impact of road infrastructure improvements on VMT, Transportation Research Part A: Policy and Practice, Volume 175.

Chatman, Daniel, Ruoying Xu, Janice Park, and Anne Spevack. 2019. Does Transit-Oriented Gentrification Increase Driving? Journal of Planning Education and Research. 39(4)482-295.

Cook, Benjamin and Willard Manning. 2013. Thinking beyond the mean. Shanghai Archives of Pyschiatry. 25(1)55-59.

Currie, Janet and Reed Walker. 2011. Traffic Congestion and Infant Health: Evidence from E-Z Pass. *American Economic Journal: Applied Economics*. 3(January):65-90.

Deakin ,Elizabeth et al. (2020), Calculating and Forecasting Induced Vehicle Miles of Travel Resulting from Highway Projects: Findings and Recommendations from an Expert Panel, California Department of Transportation (https://dot.ca.gov); at https://bit.ly/3nwaf10

Davidoff, Thomas. 2016. Supply Constraints are not Valid Instruments for Home Prices. Critical Finance Studies. 5:177-206.

Duranton, Gilles and Matthew Turner. 2011. The Fundamental Law of Road Congestion.

Duranton, Gilles, and Matthew A. Turner. 2008. "Urban Growth and Transportation." Unpublished.

Downs Anthony 1962. The Fundamental Law of Highway Congestion. Traffic Quarterly". 393-409.

Edlin, Aaron and Pinar Karaca-Mandic. 2006. The Accident Externality from Driving. *Journal of Political Economy*. 114(5): 931-955.

Fehr and Peers. 2019, Induced Vehicle Travel Impact Analysis. Technical Guidance -1.0/

Fekpe, Edward, Deepa Gopalakrishna and Dan Middleton. 2004. Highway Performance Monitoring System Traffic Data: Best Practices and Guidelines. Report for Office of Highway Policy Information

Federal Highway Administration U.S. Department of Transportation Washington, D.C.

Flyvberg, Bent, Mette Skamris Holm and Soren L Buhle. 2006. Inaccuracy in Traffic Forecasts. Transport Reviews. 26, 1: 1–24.

Fosgerua, Mogens. 2015. Congestion in the Bathtub. Economics of Transportation. 4:241-255.

Fosgerua, Mogens and Kenneth Small. 2012. Marginal Congestions Costs on a Dynamic Expressway Network. Journal of Transportation Economics and Policy. 46(3):431-50.

Fosgerua, Mogens and Kenneth Small. 2013. Hypercongestion in Downtown Metropolis. Journal of Urban Economics. 76:122-134.

Fulton, Lewis M., Robert B. Noland, Daniel J. Meszler, and John V. Thomas. 2000. "A Statistical Analysis of Induced Travel Effects in the U.S. Mid-Atlantic Region." Journal of Transportation and Statistics, 3(1): 1–30.

Garcia-López, Miquel-Àngel Ilias Pasidis, and Elisabet Viladecans-Marsal (2020), Congestion in Highways when Tolls and Railroads Matter: Evidence from European Cities. Journal of Economic Geography. 22(5): 931-960.

Gelman, Andrew. 2018. Ethics in statistical practice and communication: Five recommendations. Significance, 15: 40-43.

Gelman, Andrew and Eric Loken. 2014. The Statistical Crisis in Science. American Scientist. November-December: 460-463.

Glaeser, Edward and Giacomo Ponzetto. 2018. The Political Economy of Transportation Investment. Economics of Transportation. 13(2018)4–26.

Goodwin, Philip, and Robert B. Noland (2003), "Building New Roads Really Does Create Extra Traffic: A Response to Prakash et al.," Applied Economics.

Graham, Daniel, Emma J. McCoy & David A. Stephens (2014) Quantifying Causal Effects of Road Network Capacity Expansions on Traffic Volume and Density via a Mixed Model Propensity Score Estimator, Journal of the American Statistical Association, 109:508, 1440-1449.

Hanna, Rema, Gabriel Kreindler, and Benjamin Olken. 2017. Citywide Effects of High Occupancy Vehicle Restrictions. *Science*. 357:89-93.

Hartgen, D.T.2013. Hubris or humility? Accuracy issues for the next 50 years of travel demand modeling. Transportation 40: 1133–1157.

Hourdin, Frederic, and 14 coauthors. 2017. The art and science of climate model tuning. Bulletin of the American Meteorological Society. March: 589-602.

Hsu, Wen-Tai and Hongliang Zhang. 2014. The fundamental law of highway congestion revisited: Evidence from national expressways in Japan J. Urban Econ, 81:65-76

Hymel, Kent. 2014. Factors Influencing Vehicle Miles Traveled in California June 24 Report t

Hymel, Kent. 2019. If you build it they will drive. *Transport Policy*. 76:57-66.

Jacobs Engineering. 2023. I-5 Managed Lanes Project, Traffic Operations Analysis Report. May 22.

Kim, Jinwon. 2022. Does roadwork Improve Speed? Regional Science and Urban Economics. 93

Kim, Jinwon. 2019. Estimating Social Cost of Congestion Using the Bottleneck Model. *Economics of Transportation*. 19.

King, D. A., Smart, M. J., & Manville, M. (2022). The Poverty of the Carless: Toward Universal Auto Access. *Journal of Planning Education and Research,* 42(3), 464-481.

Knight, Brian. 2005. "Estimating the Value of Proposal Power." American Economic Review, 95 (5): 1639-1652.

Knittel, Christopher, Douglas Miller and Nicholas Sanders. 2016. Caution, Children! Drivers Present: Traffic, Pollution and Infant Health. The Review of Economics and Statistics. 98(2): 350–366.

Kwok, Karen. 2023. War on cow gas is stinky but necessary job in climate-change struggle. Reuters. March 24.

LA Metro. 2018. Expresslanes Operations Performance Report. Los Angeles, CA.

Ladd, Brian. 2012. You Can't Build Your Way out of Congestion – Or Can You? The Planning Review. 48(3):16-23.

Light, Thomas. 2012. HOT Lane Performance Under Alternative Pricing Policies. Transportation Research Forum. 51(2):61-82.

Luo, Anna, Chelsea Richer, Ron Milam and Jeremy Klop. VMT Quantification Tools and Preferred Methodology. Fehr and Peers Memorandum to LA Metro, February 21.

Manville, Michael, David King and Michael Smart. 2017. The Driving Downturn: A Preliminary Assessment. *Journal of the American Planning Association*. 83(1):42-55.

Manville, Michael. 2017. Automatic Street Widening: Evidence from a Highway Dedication Law. *Journal of Transport and Land Use*. 10(1): 375–393.

Manville, Michael, Brian Taylor, Evelyn Blumenberg and Andrew Schouten. Vehicle Access and Falling Transit Ridership. *Transportation*.

Marshall, NL. 2018. Forecasting the Impossible. Research in Transportation Business and Management. 29:85-92,

McNally, Michael. 2005. The Four Step Model. In David Hensher and Ken Button, Eds, *Handbook of Transport Modeling*. Amsterdam: Pergamon.

Metz, David. 2023. Digital navigation negates the economic benefits of road widening: The case of the M1 motorway. Transportation Research Part A: Policy and Practice. 174,

Metz, David. 2021. Economic benefits of road widening: Discrepancy between Outcome and Forecast. Transportation Research Part A: Policy and Practice.147:312-319.

Milam. Ronald, Marc Birnbaum, Chris Ganson, Susan Handy, and Jerry Walters. 2017. Closing the Induced Vehicle Travel Gap Between Research and Practice. Transportation Research Record. 2653:10-16.

Milan, Ronald, Jerry Walters and Melanie Gill. 2022. Balancing Congestion Relief and Induced VMT. Western Meetings of the Institute of Transportation Engineers.

Millard-Ball, Adam. 2015. Phantom Trips. *Journal of Transport and Land Use*. 8(1):31-49.

Millard-Ball, Adam, Benjamin Silverstein, and Purva Kapshikar. 2022. Dividing Highways: Barrier Effects and Environmental Justice. UCLA Institute of Transportation Studies.

Mokhtarian, Patricia and Ilan Salomon. 2001. How derived is the demand for travel? Some conceptual and measurement considerations, Transportation Research Part A: Policy and Practice. (35) 8: 695-719.

Murray, Michael. 2006. Avoiding Invalid Instruments and Coping with Weak Instruments. J*ournal of Economic Perspectives*. 20(4)111-132.

Naess, Petter, Morten Skou Nicolaisen, and Arvid Strand. 2012. Traffic Forecasts Ignoring Induced Demand. European Journal of Transportation Research. 12(3): 291-309.

National Academies of Sciences, Engineering, and Medicine. 2021. An Update on Public Transportation's Impacts on Greenhouse Gas Emissions. Washington, DC: The National Academies Press. https://doi.org/10.17226/26103

National Academies of Sciences, Engineering, and Medicine. 2020. Traffic Forecasting Accuracy Assessment Research. Washington, DC: The National Academies Press. https://doi.org/10.17226/25637.

Nelson, Laura. 2017. Bad Traffic in Express Lanes? Blame the 25 Percent of Drivers who Don't Pay. *Los Angeles Times*. October 13.

Noland, Robert B., and William A. Cowart. 2000. "Analysis of Metropolitan Highway Capacity and the Growth in Vehicle Miles of Travel." Transportation, 27(4): 363–90.

Moore, Adrian Samuel Staley and Robert Poole., et al 2010 The role of VMT reduction in Meeting climate change policy goals. TR-A 44:565-574.

Ossokina, Ioulia V Jos van Ommeren, Henk van Mourik. 2023. Do highway widenings reduce congestion?, Journal of Economic Geography, Volume 23 ( 4) 871–900.

Percoco, Marco. 2013. Is road pricing effective in abating pollution? Evidence from Milan, Transportation Research Part D: Transport and Environment.

Poole, Robert. 2017. Rethinking America's Highways. Chicago: University of Chicago Press.

Rodden, Jonathan. 2019. Why Cities Lose. Hachette Press.

Shoup, Donald. 2003.Truth in Transportation Planning. Journal of Transportation and Statistics. 6(1):1-16.

Shoup, Donald, 2011. The High Cost of Free Parking. Chicago: Planner's Press.

Small, Kenneth, & Xuehao Chu, 2003. "Hypercongestion," Journal of Transport Economics and Policy, 37(3): 319-352..

Small, Kenneth, Clifford Winston and Jia Yang. 2006. Differentiated Road Pricing, Express Lanes and Carpools. Brookings-Wharton papers on Urban Affairs. 53-96.

Small, Kenneth. 2015. The Bottleneck model: An assessment and appreciation. Economics of Transportation. 4:110-117.

Shewmake, Sharon. 2012. Can Carpooling Clear the Road and Clean the Air? Journal of Planning Literature. 27(4):363-74.

Steimetz, Seiji and David Brownstone. 2005. Estimating Commuters' Value of Time with Noisy Data. Transportation Research Part B. 39(10):865-889.

Systems Metric Group. 2023. Regional Induced Demand Analysis Summary. Memo to Southern California Association of Governments. September 25.

Turner, S., Tsapakis, I., & Koeneman, P. (2020). Evaluation of StreetLight Data's Traffic Count Estimates From Mobile Device Data (MN 2020-30; Minnesota Department of Transportation, p. 24). Texas A&M Transportation Institute. https://rosap.ntl.bts.gov/view/dot/57948

Tyndall, Justin. 2024. The effect of front-end vehicle height on pedestrian death risk, Economics of Transportation, 37.

Ullrich, A. (2023, October 26). All Vehicles Data Updates—December 2022. StreetLight Data Support. https://support.streetlightdata.com/hc/en-us/articles/10397273102363-All-Vehicles-Data-Updates-December-2022

Varaiya, Pravin. 2005. What We've Learned About Highway Capacity. Access. Fall: 2-7.

Vickrey, William (1969), "Congestion Theory and Transport Investment," American Economic Review, Vol. 59/2, May, pp. 251-260

Volker, J. M. B., Lee, A. E., & Handy, S. (2020). Induced Vehicle Travel in the Environmental Review Process. *Transportation Research Record, 2674*(7), 468-479

Volker, Jamey, Amy Lee and Dilllon Fitch. 2019. Streamlining the Development Approval Process in a Post-Level of Service Los Angeles. Journal of the American Planning Association. 85(2):114-132.

Volker, Jamey and Susan Handy. 2022. Updating the Induced Travel Calculator," National Center for Sustainable Transportation.

Walker, Joan, Daniel Chatman, Ricardo Daziano, and others. 2019. Advancing the Science of Transportation Demand Modeling. UC Berkeley Institute of Transportation Studies.

Wang, Jenny, Gabi Leovan, and Eliana Arroyo. 2022. The Fundamental Law of Road Congestion: Is it Truly Fundamental? Unpublished.

Welch, T. F., & Widita, A. (2019). Big data in public transportation: A review of sources and methods. Transport Reviews, 39(6), 795–818.

Winston, Clifford, and Ashley.Langer. 2006. The Effect of Government Highway Spending on road users' congestion costs. Journal of Urban Economics. 60(3):463-483.

UCLA Institute of Transportation Studies

3320 Public Affairs Building
Los Angeles, CA 90095-1656
uclaits@ucla.edu
its.ucla.edu