Technology Policy Challenges of Electric and Automated Vehicles

Submitted in partial fulfillment of the requirements for

the degree of

Doctor of Philosophy

in

Engineering and Public Policy

Aniruddh Mohan

BEng. Hons. Mechanical Engineering, University of Manchester
MPhil. Nuclear Energy, University of Cambridge

Carnegie Mellon University
Pittsburgh, PA

August 2022

© Aniruddh Mohan, 2022

All rights reserved.

To my parents

Abstract

This body of work covers some important technology policy challenges that will be posed by the deployment of electric and automated vehicles in personal mobility, ridesourcing, and long haul trucking. Substantial electrification of road transport is critical to deep decarbonization. Self-driving technology continues to inch towards wider commercial deployment, despite well-publicized setbacks. Policymaking must grapple with a deeply uncertain technology landscape and trade-offs between multiple societal objectives. I attempt to help inform public policy by carefully navigating the uncertainty present in the deployment of these technologies and highlighting how trade-offs between multiple objectives might materialize.

First, I look at the potential energy impact of automation on electric vehicle range. Will electrification and automation of light vehicles develop in sync? I find that automation will incur expected penalties of 10-15% on electric vehicle range, due to the energy consumption of computing and sensors. This suggests that deliberate technology choices towards efficient computing and aerodynamics will be needed to drive electrification of automated vehicles. In long haul trucking, I investigate the potential job impacts from automation of highway driving. I draw on detailed shipment data to show that replacing all highway driving with automated trucking will lead to more than 90% of operator-hours lost to automation. I find limited evidence to support industry claims that increased short haul jobs will compensate for the operator-hours lost on highways. Third, I present a novel agent based model that can answer a series of research questions related to electric vehicle deployment in ridesourcing. Companies such as Uber and Lyft have committed to fully-electrify their fleets by 2030 in the United States. Should policy incentivize larger battery pack sizes or greater investments in dedicated fast-charging infrastructure for ridesourcing? I instantiate the agent based model to study the lifecycle externalities of fully-electric ridesourcing in the city of Chicago, across different levels of battery size and fast-charging infrastructure. I estimate that greenhouse gas externalities are in the range of 10-13¢ per trip, with lower externalities associated with smaller battery packs. Compared to today, fully electrifying ridesourcing can cut greenhouse gas externalities in Chicago by 50-60%. Solely targeting reductions in greenhouse gas emissions, however, comes at the cost of increases in traffic externalities such as congestion.

ACKNOWLEDGEMENTS

Thank you to all the members of my thesis committee for their help and guidance. I'm grateful to Parth Vaishnav and Venkat Viswanathan for having faith in me and extending me the opportunity to work with them for my PhD. Parth has been a brilliant primary advisor, combining intellectual rigor with an indefatigable dedication to his students. Working with him has meant a rewarding and fulfilling PhD experience. Parth always found time, for the small details as well as the big. Venkat has been an ideal secondary advisor. His ability to connect the dots across seemingly disparate fields is unique and something I was lucky to learn from. He always got me thinking about the bigger picture with respect to my research and helped me to identify important questions. More than anyone, Venkat taught me the importance of perseverance and belief. Jeremy Michalek brought a level of precision to my research and methodological approach that was immensely rewarding and will stand me in good stead throughout my career. Jeremy is the most precise person I've ever met, this may be the only imprecise statement he will let me get away with. I'm grateful to Nick Muller for lending his economics expertise to my research. Working with Nick has challenged me outside of my own comfort zone, making me a better researcher in the process. Some of the most exciting research I did at Carnegie Mellon was with Nick and involved among other things solving differential equations on the whiteboard in his office, that I once jokingly referred to as our Good Will Hunting moment. Last but not least, Jessika Trancik helped me conceptualize my own work through a different lens with her questions and comments. The work she has led at MIT has been a continual inspiration throughout my PhD and I am fortunate to have her participation on my committee.

I would like to thank the organizations who supported my research. In particular, the U.S. Department of Transportation's University Transportation Center, and the Block

Center for Technology and Society and Wilson E. Scott Institute for Energy Innovation, both at Carnegie Mellon. I am also grateful for funding support from the Carnegie Mellon Presidential Fellowship, the Department of Engineering and Public Policy, and a discretionary grant for the Viswanathan lab.

I would like to acknowledge the staff in the Department of Engineering and Public Policy including Vicki Finney, Adam Loucks, and Debbie Kuntz for their kindness and service. They were always on top of things and pulled me through the past four years.

It takes a village as they say. None of this would have been possible without the love, patience, and support of my parents, partner, and friends. I owe them a debt of gratitude.

Finally, to Carnegie Mellon. Thank you. These were the most intellectually rewarding years of my life. My heart is in the work, and always will be.

Aniruddh Mohan
Pittsburgh, PA
August 2022

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

Introduction

Light and heavy vehicles are responsible for about 60% and 25% of U.S. greenhouse gas (GHG) emissions from transportation, respectively (1). Substantial electrification of these vehicles is considered essential for decarbonization (2). California's announcement to ban new combustion engine sales by 2035 is the first of its kind in the United States but follows similar announcements by the UK, France, and other major economies. Furthermore, ridesourcing services such as Lyft and Uber which make up a significant share of the total vehicle-miles traveled (VMT) (3) have also recently committed to a fully electric fleet in the United States by 2030. However, regulators focused on driving electrification should also consider another major innovation in transport, automated vehicles (AVs). The widespread deployment of AVs following the current testing on public roads is a cornerstone of the business strategy of ridesourcing companies. Long haul trucking is also likely to see early adoption of automation due to the lower technological complexity of highway driving and the associated cost savings. Whether automation will go hand in hand with electrification is not clear. For example, while some companies including General Motors have committed to their automated vehicles being electric, other automakers have declared a preference for gas-electric hybrids. Further, deployment of automation in long haul trucking has the potential to cause enormous job losses while also disrupting traditional operational modes, which could lead to opportunities for the introduction of new technologies.

The deployment of early stage technologies always involves significant uncertainty, around costs, feasibility, and intended and unintended consequences (4). Policymaking must grapple with a deeply uncertain technology landscape and trade-offs between multiple societal objectives. For instance, AVs have the potential for major benefits in emissions, congestion, and more but on other hand could also lead to undesirable outcomes if VMT and congestion increases, AVs prove slower to electrify than human driven vehicles due to impacts on range, and employment impacts are large. Electrification of ridesourcing vehicles can help reduce GHG emissions, but is it more cost effective to incentivize larger battery packs or more public fast charging infrastructure? Promoters of novel technologies are both quick to claim their technology is revolutionary and to also comfort concerned regulators with assertions of continuity and that 'this is nothing new' (5). Regulatory decision making, in turn, must walk the tightrope between consistently applying long standing rules while also having to adjust those regulations to respond to new developments. Finally, policymakers often only have a brief window of opportunity to steer the deployment of emerging technologies in a way that helps contribute towards societal objectives.

This dissertation covers some of the important technology policy challenges that will be posed by the deployment of electric and automated vehicles in personal mobility, ridesourcing, and long haul trucking. I attempt to help inform public policy by carefully navigating the uncertainty present in the deployment of these technologies and highlighting how trade-offs between multiple objectives might materialize.

The dissertation is structured as follows. First, I look at the potential energy impact of automation on electric vehicle (EV) range. This is an issue that has received widespread media coverage but limited attention in the academic literature. What are the impacts on the range of EVs if they are automated, from the energy use of computers and sensors required for automation? We study this issue in detail using a vehicle dynamics model and show the impact of automation on EV range and battery longevity. We find that automation will present trade-offs with light vehicle electrification with expected range penalties of 10-15%. Investing in energy efficient computing and aerodynamic integration of sensor stacks on vehicles can help reduce this impact, and may even help increase EV range if the benefits of smoother driving compared to humans can be realized.

Next, I investigate potential job impacts from automation in long haul trucking. Here, we focus on the share of operator-hours that would be lost if highway trucking were automated using an engineering analysis incorporating detailed data on trucking shipment flows. We explore alternative scenarios of deployment in the near-medium term and find that eventually, replacing all highway driving with automated trucking will lead to more than 90% of operator-hours lost to automation. Deployment only in the sun-belt where weather conditions are more amenable and where all testing is currently taking place will impact 10% of operator-hours. Overall, we find that automation is likely to cause significant disruption to long haul trucking even if just restricted to highways, but note that such deployment can create opportunities for partial electrification of heavy trucking.

Third, I study the lifecycle externalities of the full electrification of the operation of transportation network companies (TNCs) such as Lyft and Uber. I present a novel agent based model (ABM) in Chapter 4 that can answer a series of research questions related to EV deployment in ridesourcing. The novel ABM presented in Chapter 4 draws on high geo-spatial detail, labor economics-based literature on the labor supply decisions of TNC drivers, and rich public datasets of TNC trips by riders that provide detail on the demand for these services. I outline how my model presents a significant advance over other ABMs in the literature which have previously studied the operation of taxi fleets in urban areas. In Chapter 5, I then instantiate this model on TNC trip data for the city of Chicago, to understand how fully-electric taxis might impact GHG emissions across different battery sizes and fast-charging infrastructure investments. We find that GHG externalities from fully-electric ridesourcing are in the range of 10-13¢ per trip, with lowest emissions across the different scenarios considered associated with a 40kWh battery size and 1 dedicated fast-charger for TNCs per square mile. We show that compared to today, fully electrifying ridesourcing can cut TNC GHG externalities in Chicago by up to 60%.

Finally, in Chapter 6, I present conclusions from this body of work and discuss relevant implications for public policy. I also briefly highlight the methodological contributions of my work and identify questions for future research in this area.

This work has benefited from collaborations with a number of colleagues and where possible I have attempted to list them. All errors and omissions are mine alone.

CHAPTER II

Trade-offs between automation and light vehicle electrification

This chapter was developed with coauthors Shashank Sripad, Parth Vaishnav, and Venkat Viswanathan. It is based on work published in:

> Mohan, A., Sripad, S., Vaishnav, P. & Viswanathan, V. Trade-offs between automation and light vehicle electrification. *Nature Energy* **5**, 543–549 (2020).

An abridged version is provided here, and the full manuscript is included as an appendix.

2.1 Introduction

EVs form an increasing share of new vehicle sales around the world. Several countries are seeking to phase out internal combustion engine vehicles. Simultaneously, AVs are being tested on public roads. Automation could reduce vehicle energy use through smoother driving, platooning, shared mobility, and optimal routing (7; 8; 9). There is limited literature on the tradeoffs between automation and electrification (10). Early automated electric vehicles (AEVs) may be heavier, need extra computing and sensor power, and (due to the possible need for protruding sensors) be less aerodynamic than EVs. If this reduces driving range substantially, it has been suggested that the first AVs will be gas-electric hybrids,(11) which could slow the electrification of the light vehicle fleet.

In this paper we compare the vehicle-level energy use, range, and battery life of a vehicle equipped to attain Society of Automotive Engineers (SAE) Level 4-5 automation to human driven EVs, by undertaking a careful consideration of the effect on vehicle level energy use of the different components needed for automated driving, as well as the potential increase in drag from LiDAR. We find that automation is likely to reduce EV range; and that the aerodynamic impact of LiDAR is a significant driver of this reduction in suburban driving, while computing loads have a greater effect for city driving. Across different EVs, we find a median reduction in the range of 3-5% for a suburban drive profile and no drag impacts from LiDAR. Including LiDAR increases the median range reduction to between 8-12%. A reduction in range will lead to more frequent charging speeding up battery degradation but we find that this impact on battery longevity is negligible.

2.2 AEV energy consumption

Sripad and Viswanathan (2017) (12) developed a physics-based vehicle dynamics model to estimate the energy demands of an EV given a realistic driving profile. They determined the battery size needed for a given vehicle range or equivalently, the range of an EV given battery size. Using a realistic velocity profile with a one second temporal resolution, the model calculates the instantaneous power needed each second to overcome vehicle inertia, aerodynamic drag, and road friction. We extend this model for AEVs by adding the weight of the different components to the mass of the vehicle and battery pack, increasing the drag coefficient for automated solutions with a roof-based spinning LiDAR. If no LiDAR is used, or if solid-state LiDAR that is incorporated into the aerodynamic profile of the vehicle is used, the increase in drag is zero. We also modify the velocity profile to account for potentially smoother driving and add the computing and sensor loads at each second. Keeping track of the total energy used, we repeat the driving profile until the battery is fully depleted. This gives us an estimate of the AEV range for a given battery capacity. We then compare this AEV range to the EV to understand how automation affects vehicle range. We run this simulation for two types of velocity profiles: the California Unified Cycle Driving Schedule which is a composite profile i.e. a city-highway mix, and the Urban Dynamometer Driving Schedule which is a city-only profile. Full details of the physics model are provided in the Methods section of the published manuscript (6).

There are several different combinations of sensor hardware that are currently being tested on vehicles that aim to achieve full automation. For example, some developers are using solutions that include LiDAR while others are relying solely on cameras and radar. There are also differences in the choice of vendor for LiDAR or radar, and in the number of sensors. Given the numerous possible combinations, we assume a uniform distribution for the sensor and connectivity load. We bound this between 30W to 150W. The lower case represents a low powered LiDAR solution such as the 15 W Ouster OS1 (13) system along with 2 Bosch mid-range radars (MRR) of 4.5 W (14) each, 3 Pt. Grey Dragonfly cameras of 1.5 W (15) each, and 1.5 W for connectivity. Some industry developers have also suggested that a sensor package without LiDAR is sufficient for high level automation. As such, the lower case could also represent an optical-only system with no LiDAR and 9 cameras (1.5 W each), 2 MRR of 4.5 W each , and the remaining 7.5 W for communication and connectivity. The upper estimate of 150 W represents a system with 2 Velodyne 64 LiDAR domes of 60 W each (16), along with 2 Bosch MRR (9 W) radars (14) with 6 Pt. Grey Dragonfly cameras (9 W) (15) and expanded communications and connectivity (12 W). We were unable to find reliable estimates of the power draw of communication systems, perhaps due to the nascent nature of the technology. We assume a power draw of 1.5W for our low estimate and 12W for our high estimate. Modern smartphones consume less than 1 W of power for streaming, connectivity, communication and other functions (17). Many EVs already have Global Positioning System (GPS) and other connectivity installed; so it is not clear that the communications load will be additional to existing loads. The Institute of Electrical and Electronics Engineers (IEEE) standard for dedicated short range communication (DSRC) limits the Equivalent Isotropically Radiated Power to less than 2 W (18), which precludes a large power draw for the transmitter.

Estimates of computing load for automated driving in the literature found power requirements from a few hundred watts (19) to several thousand watts (20; 21). The Nvidia Pegasus system has been advertised as capable of level 5 autonomous driving and has a power load of 500W (22) while the recently released Tesla Full Self Driving (FSD) claims to require only 150 W (23). We therefore bound our estimates between 150W to 1000W with the upper bound representing the higher estimates from the literature. Higher values of computing load are possible given that the technology is nascent, regulators might require redundant systems, or cooling requirements for computing may have been underestimated. On the other hand, over the long term, improvements in chip design and computing efficiency should see power loads fall substantially. We also assume a linear relationship between the computing and sensor loads between their respective bounds as increased data flow from the sensors to the computing platform will require a concomitant increase in computing capacity and therefore power draw.

We are aware of no publicly available, empirical estimates of the effect of roof-based LiDAR on the vehicle drag coefficient (7); so we approximate this effect by using data from wind tunnel tests of drag impacts of roof add-ons such as police sirens, signs, and racks.(24; 25) We therefore estimate a lower bound value of 15% increase in drag (which corresponds to the drag increase from a taxi sign), to 40% (which would impose the same aerodynamic penalty as a barrel). In the case of solid state LiDAR or AVs operating with cameras only, there would be no increase in drag. We therefore consider a second separate case: if there is no LiDAR, or a solid state LiDAR system, we model no increase in drag. In the longer term it is also likely that developers will incorporate LiDAR into the vehicle in a way that does not result in additional drag as consumers may not want to purchase vehicles that have prominent external sensors, as is the case with the AVs being tested today.

Finally, to simulate AEV drive cycles that are smoother than human drivers, and the associated energy savings, we apply a smoothing spline function similar to Liu et al(26), to the composite and city drive profiles. The smoothing function can be adjusted to yield different levels of energy savings. We bound this between 5% to 25% in line with values from the literature that have estimated the energy savings from smoother driving of AVs (27; 28; 8; 29). Details of the smoothing spline function are provided in the Methods section of the published manuscript (6). Our method of smoothing makes no assumptions about the vehicle drive train. We smooth the velocity profile and calculate the energy savings from first principles, using our physics-based model. The effect of smoothing for both drives cycles for different levels of energy savings is shown in Supplementary Information of the published manuscript. We also find possible trade-offs between the energy savings from smoother driving and safety, particularly at higher levels of energy savings wherein the vehicle may be at rest in the original profile but has a non-zero velocity in the smoothed profile. However, we find almost no change in our estimates of AEV range when we constrain our smoothed drive cycles so that the velocity is zero at all instances where the velocity in the original profile is zero.

2.3 Effect of automation on EV range

We use vehicle design parameters based on a Tesla Model 3 with 310 miles of range and an 80 kWh battery pack as our base EV (30). We use a combination of scenario analysis and Monte Carlo (MC) simulations to understand how the deep uncertainty in our input parameters affects our estimates of reduction in range. We consider two broad deployment scenarios for automation: with or without LiDAR. We treat these scenarios separately for two reasons. One, there is considerable disagreement among the manufacturers as to whether LiDAR is essential (31), with developers deploying both LiDAR-based and LiDAR-free technologies. Two, LiDAR can impose a considerable energy and aerodynamic penalty: eliminating it would have a large effect on energy use. Within each scenario, we use MC analysis, since there are too many plausible combinations of other parameters to reasonably justify one scenario over another in the current early stages of AV development. MC simulations allow us to explore the entire parameter space (32).

We find that adding automation with LiDAR would result in reduced range or require a larger battery in more than 95% of simulations (Fig. 2.1a). This indicates that while there is uncertainty around technology choices and power loads of automation enabling equipment, LiDAR-based solutions that have a drag penalty will almost certainly reduce range compared to human driven EVs. Even if we assume no power consumption from LiDAR and no change in drag, energy use increases in 76% of simulations for the composite drive profile, relative to a human driven EV. We find that even with computing load and sensor load each at their lower bounds and no increase in drag, AEVs will need energy savings from smoother driving of at least 4-5% over a human driver to have the same range as EVs. If savings from smoother driving are higher, this will lead to an increase in EV range. In 90% of our MC simulations for the composite drive cycle the reduction in range is between 1% to 16% for an automated Tesla Model 3 with LiDAR compared to an EV with the same battery. The median estimate is a 9% loss in range. Next, we apply the same analysis to estimate the range impacts for other EV models such as the Tesla Model S, Nissan Leaf, Chevy Bolt, and Xpeng G3. The results for the composite drive cycle for all EV models considered are summarised in Fig. 2.2a.

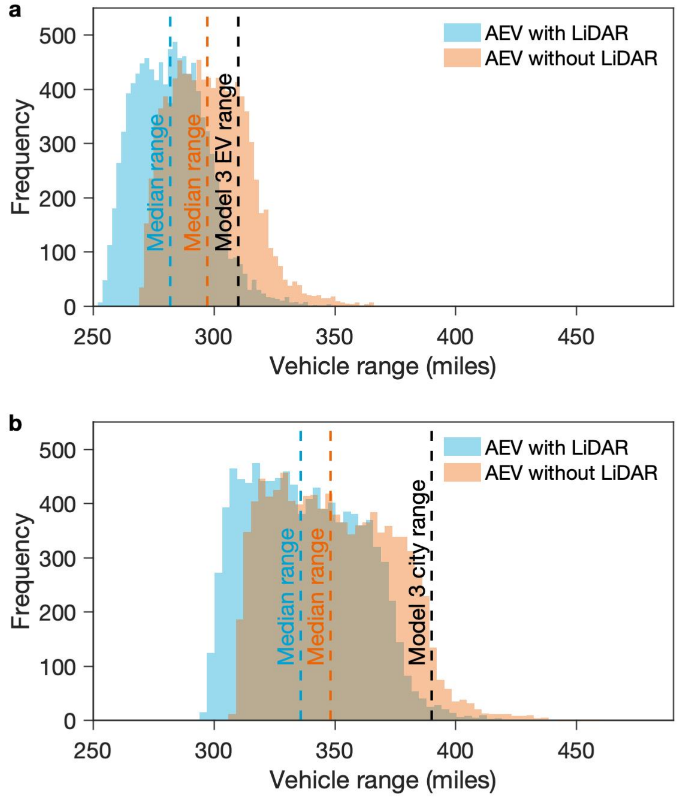

Figure 2.1: **Histogram of range results for composite and city drive profiles for a Tesla Model 3 with an 80 kWh battery pack.** **a**, Results of simulating the vehicle range of an automated Model 3 given the composite driving profile and input parameters for automation with and without LiDAR. The median loss in range is 9% with LiDAR and 4% without. In 90% of our MC simulations the range impact is between [-16%, -1%] and [-11%,+5%] respectively. This is compared to the Tesla Model 3 EV-only range. **b**, Results of simulating the vehicle range of an automated Model 3 given the city driving profile and input parameters, for automation with and without LiDAR. The median loss in range is 14% with LiDAR and 11% without. In 90% of our MC simulations the range impact is between [-22%,-4%] and [-19%, 0%] respectively. This is compared to the Tesla Model 3 EV-only range for city driving.

In 90% of our MC simulations the range penalty for the city drive cycle is between 4% to 22% for a Model 3 AEV with LiDAR (Fig. 2.1b). The median estimate is a 14% loss in range. Results for different EV models are shown in Fig. 2.2b.

Table 2.1 shows the energy demand and range impacts for different autonomous configurations considered. We see that the drag impacts of spinning LiDAR are more keenly felt at higher speeds on the highway. Keeping all other autonomy parameters fixed, the drag

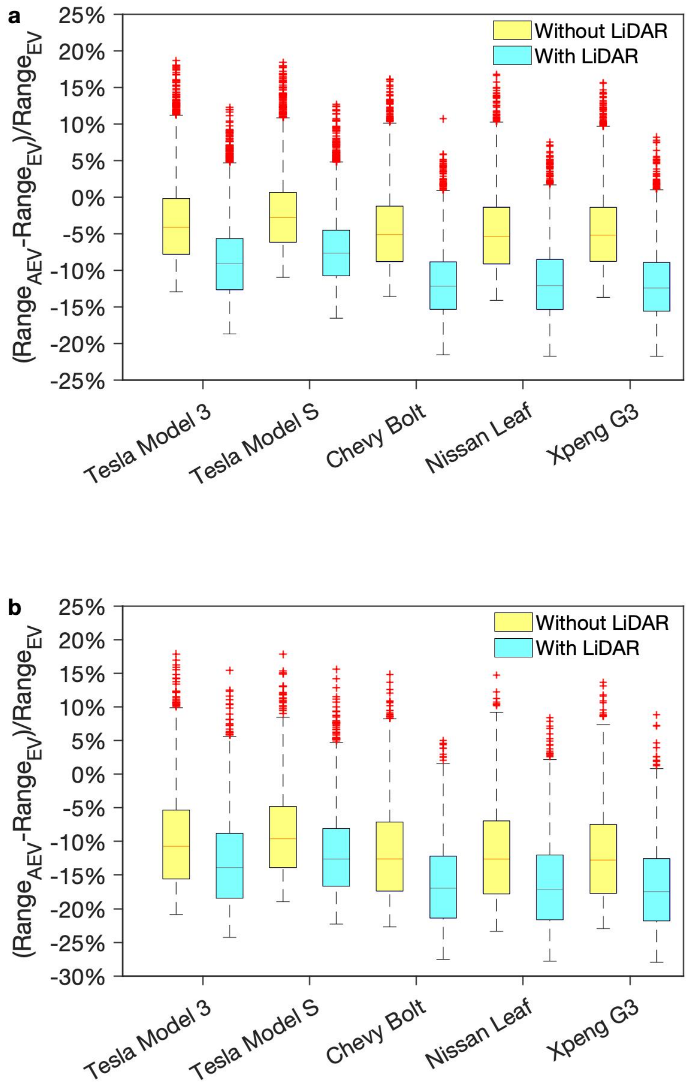

Figure 2.2: **Results of the Monte Carlo simulation of AEV range for different EV models.** The Y axis shows the percentage change in range for the AEV compared to the base EV. The horizontal red line is the median change in range and the red data points are the data that are beyond the whisker where the maximum whisker length is 1.5 times the interquartile range. The X axis lists the five EVs considered in the analysis. **a**, boxplot showing the range impact for the composite drive profile and for automated solutions with and without LiDAR. **b**, boxplot showing the range impacts for the city drive profile and for automated solutions with and without LiDAR.

increase from having roof based spinning LiDAR results in an additional 5% range penalty for the composite profile compared to 3% for the city profile. Our results therefore show that developers may need to consider automated solutions that either eliminate LiDAR or reduce its drag impact. Solid state LiDAR could reduce drag but impose an additional computing load. However, this is unlikely to be as significant as the impact of additional drag, as even an increase of 100 W in the compute load will only decrease range by 1%. The higher range penalty in the city cycle (Fig. 2.2b) compared to the composite profile (Fig. 2.1a) across all EVs we analyze is due to the greater effect of computing loads (which we assume applies at each second) on range as a result of longer trip times in the city for the same distance travelled. As such, we see in Table 2.1 that moving to a more efficient compute stack provides a greater benefit in the city than for suburban driving. Developers who wish to deploy robo-taxis in urban environments therefore have strong incentives to bring down computing power needs in order to maximize vehicle utilization. This could be achieved over time as standardized algorithms are built into efficient specialized circuits or if developers design dedicated chips for neural network computation.

Table 2.1: **Energy demand and range impacts for different autonomous configurations.** All scenarios assume 10% energy savings from smoother driving and a sensor load of 30W. Increase in drag from LiDAR is fixed at 25%.

Drive Cycle	**Technology**	**Range Impact**	**Wh/mile**
Composite	150 W compute + no LiDAR	+6%	244
	500 W compute + no LiDAR	0%	259
	500 W compute + LiDAR	-5%	272
City	150 W compute + no LiDAR	+5%	195
	500 W compute + no LiDAR	-4%	214
	500 W compute + LiDAR	-7%	221

2.4 Negligible impact of automation on battery longevity

A decrease in range could lead to more frequent charging or longer charging times which leads to faster battery degradation. We use the approach developed by Sripad and Viswanathan (2017) (12) to model the longevity of the battery for each type of vehicle; that is, to estimate the number of total miles for which the vehicle can be driven until the battery is unable to charge to more than 80% of its original capacity. To realistically assess battery degradation, we model a series of a 24-hour periods in each of which the vehicle drives for 50 miles for the composite profile (or 30 miles for the city profile), charges until the battery is full, and then rests until it is driven again the following day. For the case of the composite profile and the base AEV with LiDAR, the loss in battery longevity for the median reduction in range is 5% or 5,500 miles and the 5th percentile case with a range reduction of 16% would result in a loss of 9% or 10,000 miles, as shown in Fig. 2.3a. Given the longevity of state of the art batteries is constantly improving this effect is likely to be further negligible in the future (33).

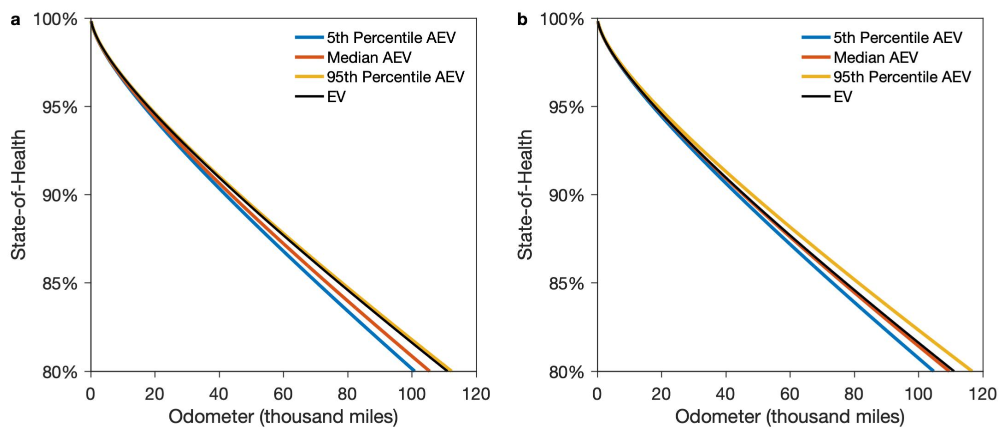

Figure 2.3: **Battery degradation results for our base EV and AEV for the composite drive cycle.** The simulations were performed for a daily drive schedule of 50 miles. A state-of-health or a remaining useful capacity of about 80% is considered the end-of-life of the battery pack. The battery degradation simulations are performed using a battery pack model where the cells are simulated using a pseudo two dimensional electrochemical model which contains a degradation sub-model. A series of 24-hour power profiles which include 50 miles of daily driving is used as an input to the battery pack model. The details of the degradation model are discussed in Methods. **a**, the base AEV with LiDAR. The median loss in battery longevity with LiDAR is 5,500 miles or under four months of driving. The 5th percentile case with LiDAR leads to a loss of 10,000 miles or over half a year of driving. **b**, the base AEV without LiDAR. The median AEV without LiDAR lasts for approximately 110,000 miles and the 5th percentile AEV lasts for approximately 105,000 miles. The median loss in battery longevity compared to the EV is 1,300 miles or a month of driving. The 5th percentile case leads to a loss of 6,000 miles or four months of driving compared to the EV.

The 95th percentile AEV shows a small (1%) increase in battery longevity compared to the EV due to the effect of smoother driving, which lowers the discharge rate of the battery as well as the re-charging rate through regenerative braking. Moreover, due to the higher energy supplied by regenerative braking segments for the EV, during each daily round trip, the average state-of-charge of the battery pack for the EV is about 0.6% higher than that of the 95th percentile AEV. Since a higher state of charge increases the rate of battery degradation (33), the EV battery degrades faster than the 95th percentile AEV, since this state-of-charge effect outweighs the very small increase in the duration of charging time from the 1% loss in range compared to the EV. Battery degradation results for the case without LiDAR are shown in Fig. 2.3b.

2.5 Conclusions

While there is considerable uncertainty with respect to the technologies that will enable fully automated driving, our model provides a way to grapple with this uncertainty and derive useful insights about the vehicle-level energy use of AVs. We find that high level automation will likely reduce EV range by about 5-10% in the case of the suburban drive profile. Given that the specific energy of lithium-ion batteries is increasing at roughly 5% each year without a corresponding increase in cost or a step change in technology, automation will have the same effect on range as a 1-2 year time-lag on improvements in battery specific energy. Will consumers value full automation more than the modest loss in range we estimate? Future work must assess the benefits of automation (e.g., using travel time productively, increased safety) against its cost (e.g., loss in range, increased capital cost, public perceptions of safety) (34). Our study provides a basis for more realistic assessments of consumer preferences, since we estimate the change in EV range associated with automation. In conclusion, to minimize the effect of automation on EV range, developers should focus on bringing down power requirements for computing and on incorporating sensors in aerodynamic designs.

2.6 Data availability

All underlying data are publicly available at https://github.com/battmodels/Automation-EV-Range.

2.7 Code availability

The custom code for the Monte Carlo simulation of the physics model based range impacts of automation on EV range that has been presented in this paper is publicly available on GitHub at https://github.com/battmodels/Automation-EV-Range.

In order to allow readers to engage with our research we have also created a web applet (available at https://tinyurl.com/avrange) that allows users to select different combinations of radar, computing and LiDAR load, cameras, energy savings, and electric vehicles, to assess the effect of different assumptions about automation on vehicle range. The input files for the battery degradation simulations on AutoLion-ST v6.3, Build 2 will be provided on request.

Funding

This work was supported by the CMU College of Engineering, Dept. of Engineering & Public Policy; Scott Institute for Energy Innovation; Center for Climate and Energy Decision Making (SES-1463492; through a cooperative agreement between the National Science Foundation and CMU); the Block Center for Technology and Society at CMU; Technologies for Safe and Efficient Transportation University Transportation Center; Mobility21, A United State Department of Transportation National University Transportation Center.

CHAPTER III

Impact of automation on long haul trucking jobs in the United States

This chapter was developed with coauthor Parth Vaishnav. It is based on work published in:

> Mohan, A. & Vaishnav, P. Impact of automation on long haul trucking operator-hours in the united states. *Humanities and Social Sciences Communications* **9**, 1–10 (2022).

An abridged version is provided here, and the full manuscript is included as an appendix.

3.1 Introduction

Automated driving technology is currently being tested on public roads in the United States in both the light vehicle and heavy duty segments. Given the likely reduced operational complexity involved in highway driving, several companies are currently working on developing automation for long haul trucking which is designed to work as per a 'transfer hub' model (36; 37). This would involve an automated truck (AT) completing the highway leg of the route and human drivers undertaking the more complex suburban-urban segments at both the starting and end points of the journey. Truck ports near highways would be used to switch out the trailer from the prime mover and enable this switch at both ends. For a schematic of the possible transfer hub model see Figure 3.1.

The promise of ATs has led to widespread concern about job losses in long haul trucking, which is a common profession in the United States, particularly for men with high school educations (38). On the other hand, it has also been noted, often by the companies developing

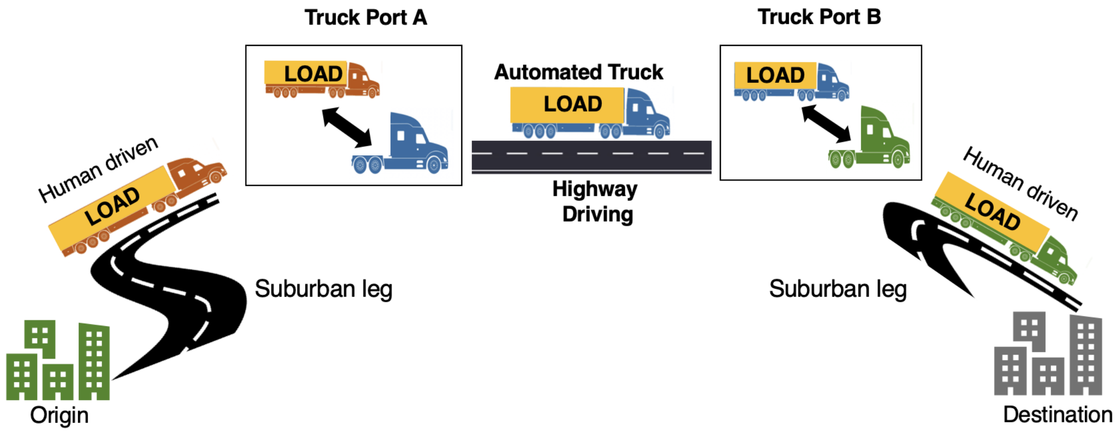

Figure 3.1: **Schematic showing the possible operation of a transfer hub model where a human driven truck drives the load from the origin to the first truck port where the load is switched to an automated prime mover shown in black.** This completes the highway leg of the journey to another truck port where a human driven truck takes the load from the port to the final destination.

this technology, that long haul trucking currently faces a labor shortage and automation will create new short haul jobs which will more than make up for the long haul jobs lost. As a result of these conflicting claims, as well as the uncertainty over the technology itself and its limitations, there is little clarity on how automated trucking will be deployed and its economic and political ramifications, such as the impact on the long haul trucking labor market.

We use data from the Commodity Flow Survey (CFS) 2017 (39) which is a dataset jointly produced by the U.S. Census Bureau, U.S. Bureau of Transportation Statistics, and the U.S. Department of Commerce. CFS produces a sample of shipments in the United States including data on the type, origin and destination, value, weight, modes of transportation, and distance shipped. We estimate the operator-hours required for different routes using origin and destination information for trucking shipments from the CFS dataset, which enables us to estimate highway and (sub)urban splits for each shipment. We consider the technological constraints of automated trucking to make more refined estimates of the possible near term impacts of automation on long haul trucking operator hours in the United States by assessing different scenarios of deployment. We find that contrary to strong claims by companies developing this technology (40), the loss in long haul operator-hours is unlikely to be compensated for by an increase in demand for short haul drivers. We find large labor impacts from the deployment of the transfer hub model, contingent on the pace of progress of automation technology's ability to function in different operating conditions. We compare the capabilities of automation with the tasks truck drivers are required to perform through discussions with industry stakeholders, including long haul operators. This allows us to have a realistic assessment of the tasks the technology will be required to perform over the highway legs and explore the feasibility of such deployment scenarios. In this framework, where jobs are considered as a bundle of tasks, our work builds on an extensive literature in the social sciences (41; 42; 43). Through our limited sample of interviews, we discover reluctance to shift to new modes of operation such as short haul driving that will result from any application of the transfer-hub mode of deployment. The rest of the paper is structured as follows. Section 3.2 reviews the existing literature on job impacts of AT. Section 3.3 describes our methods for both the analysis of operator-hour impacts and semi-structured interviews. Section 3.4 presents the results and 3.5 concludes.

3.2 Literature Review

Given the relatively recent advances in automated driving technology, there are a limited number of studies that have been conducted on the job impacts of automation on truck driver jobs. Viscelli (2018) (44) has analysed the impact of automation based on a similar 'transfer hub' operational model assumption shown in Figure 3.1, where highway driving is automated while human drivers perform the urban segments of the route. Using revenue data from major trucking carriers and estimates of average per driver revenue, Viscelli (2018) (44) estimates nearly 300,000 long haul jobs to be at risk. As per the analysis, these at-risk jobs are primarily concentrated in dry van and refrigerated trucking, which are characterised by high turnover and low wages.

Similarly, Gittleman and Monaco (2020) (45) have undertaken a study analyzing the potential job losses from automation. They also find that contrary to several estimates in the media, the upper bound on job losses from automation are likely to be 400,000 jobs, far less than popular estimates of 1-2 million jobs at risk. Gittleman and Monaco (2020) (45) perform this analysis using 2002 Vehicle Inventory and Use Survey (VIUS) data to gather estimates of the operational use of heavy trucks. They then apportion the number of long haul truck drivers using Occupational Employment Statistics data from the Bureau of Labor Statistics. Their study also notes that only a specific aspect of trucking - long haul highway driving - is ripe for automation and that operators who perform other tasks or are involved in customer facing roles such as in package delivery services are unlikely to face job losses.

Groshen et al (2019) (46) consider different scenarios of adoption of automation and analyze the job impacts in different sectors of the economy using simulations and consultations with industry experts. They estimate 60-65% of heavy truck and tractor-trailer truck driving jobs to be eliminated with full implementation of automation (46).

Waschik et al (2021) (47) leverage a dynamic model of the U.S. economy to simulate the macroeconomic impacts of automated trucking in the economy. They consider three different speeds of adoption: slow, medium and fast and find that employment in for-hire trucking falls by 20-25 percent and in private trucking by 4-5 percent across the scenarios considered (47). Broader benefits in terms of productivity and employment in the economy are found to outweigh the employment impacts on long haul operators and the study assumes that long haul operators will switch to short haul jobs, minimizing the overall effect on the workforce.

Finally, future of work studies around the deployment of automated vehicles have also delved into the issue of job losses in trucking from automation. For instance, Leonard et al (2020) (48) suggest that automation will create new roles such as remote management of trucks, dispatching, and field support while disrupting day driving jobs. However, they do not estimate the number or share of jobs at risk.

We found no examples in the literature of analyses that used trucking routes and shipment data or examined the specific capabilities of the technology (e.g. - only operating in favorable weather conditions) to estimate the labor impacts of automated trucking. We also found no studies which undertook interviews with stakeholders including long haul operators to understand their perspectives on the tasks required over the highway legs.

Outside of the specific case of automated trucking, we also find a broad literature in economics on the impacts of automation on employment. Frey and Osborne (2013) (49) model the probability of computerization for different occupations and find that workers in transportation and logistics are among those at high risk of automation, partly due to their low wages and low levels of educational qualifications. However, their methods, including considering that entire occupations rather than single tasks within jobs can be automated have faced criticism. Arntz et al (2016) (42) have shown that jobs are a bundle of tasks and while tasks may be easy to automate, jobs are often not. As such, forecasts of a large number of job-losses from automation, such as Frey and Osborne (2013) (49) can overestimate the impact of automation. One of the questions that therefore emerges here is the composition of tasks involved in long haul trucking. We attempt to unpack this through our semi-structured interviews with truck operators.

Borland and Coelli (2017) (50) study the impact of automation on employment in Australia and find that total work available did not decrease following the introduction of computerization and that job turnover in the labor market has also not increased due to computer-based technologies. Mokyr et al (2015) (51) have analyzed the history of automation and job losses, explaining how there has always been anxiety on how technological progress will cause substitution of machines for human workers, leading to unemployment. However, such scenarios do not come to pass because the long-run effects of technology are beneficial in terms of net job-creation (52) and technology only impacts the type of jobs available and what they might pay. Technology can also complement labor resulting in increased productivity, earnings, and demand for labor (52; 53).

Nevertheless, perceptions of large job losses from automation are present in the social and political discourse. In terms of the perception of automation in trucking, Dodel and Mesch (2020) (54) have shown how workers in occupations involving a greater number of manual or physical tasks, such as in the case of long haul trucking, can have more negative perceptions regarding the impact of automation on their livelihood. Orii et al (2021) (55) analyzed discussions related to automation among the members of the r/Truckers subreddit and found that less than 1% of comments had positive views on automation.

3.3 Methods

We contribute to the literature in the field by drawing on the CFS data (39) to get a reliable estimate of the density of long haul trucking in different regions and the operator-hours required for different routes. Our primary analysis is centered in the use of freight data along with routing and operator-hour algorithms to estimate the share of operator-hours that may be lost to automation. We complement this quantitative analysis with a limited number of interviews with long haul trucking stakeholders to understand the feasibility of a transfer-hub mode of deployment. Overall, our mixture of quantitative and qualitative methods is based on the triangulation method (56), useful for analyzing socio-technical transitions and emerging technologies. We further elaborate on our methods below.

3.3.1 Data

The Commodity Flow Survey is a well known dataset for transportation planning and research, produced every five years by the U.S. Bureau of Transportation Statistics, U.S. Census Bureau, and the U.S. Department of Commerce. The latest iteration, CFS 2017, is a sample of 5.9 million shipments from approximately 60,000 responding establishments (39). We disregard inter-modal shipments and focus on shipments delivered through for-hire trucks and private trucks. We only consider shipments routed over more than 150 miles as those are commonly classified as long haul (57; 44). This subset contains nearly 1.5 million trucking shipments detailing origin and destination states, shipment distance, weight, and financial quarter. The data also contain a weighting factor, which can be used to estimate the total number of shipments of that type in the population.

3.3.2 Routing

We draw on the Google Maps API and the GGMAP package (58) in R to estimate highway and (sub)urban splits for each shipment. For the purposes of routing we categorize two types of shipments in the dataset: intrastate and inter-state. Intra-state shipments are those where the shipment does not cross state borders. Inter-state shipments are those where the state of origin is different to the state of destination.

We apply a differentiated methodology to calculate the highway and (sub)urban splits for each shipment depending on whether the shipment is within a state or across it and depending on whether the shipment has listed origin and destination Metropolitan Statistical Areas (MSAs). MSAs are listed for many but not all shipments in the dataset. If MSAs are not provided, we use the closest approximation of origin or destination location.

The different types of shipments and the methods used to calculate the highway and (sub)urban splits and highway and urban average speeds are shown in Table 3.1.

For inter-state shipments we proceed as follows. Where possible we use the origin and destination MSAs from the CFS dataset in Google Maps and estimate the highway and urban distance ratios (details provided in the next subsection) for those routes which we then apply to the actual shipment distance from the dataset. To do this, we assume that the precise origin or destination is the centroid of the MSA. For shipments that specify either the origin or destination MSA (but not both), or specify neither origin nor destination metropolitan areas, we use the rest of the state centroid which is defined as the centroid of all other areas of the state that are not listed MSAs. Note that this is an approximation which affects the estimate of the highway and sub(urban) split but *does not* affect the distance of the shipment, which is provided in the CFS dataset.

For intrastate journeys we apply the same method for shipments which have specified origin and destination MSAs. For those that do not, we apply the highway and sub(urban) split derived from the Freight Analysis Framework dataset (59) by splitting the roads into those have average speeds below and above 50mph.

Let the place of origin be designated as $\mathrm{p_{o,i}}$ and place of destination be designated as $\mathrm{p_{d,i}}$ where i is a shipment. Then, consider a shipment from $\mathrm{p_{o,i}}$ to $\mathrm{p_{d,i}}$ where $\mathrm{p_{o,i}}$ and $\mathrm{p_{d,i}}$ are set as per the cases listed in Table 3.1. The Google Maps API where applicable then provides us with detailed route directions which list the amount of time driven for any stretch of road before the next turn and so on. This allows us to calculate speeds for each section and then split the drive into segments which are greater than or equal to 50mph (classified as highway) and below 50mph (classified as urban or suburban). Note that the route suggested by Google Maps may be different depending on the time of day that the API request is sent. We therefore ran several iterations of the routing algorithm at different times of day and found no discernible difference to our results.

Let the highway segment of this journey be $\mathrm{h_i}$ and the urban sections $\mathrm{u_i}$. Let the origin-destination distance be $\mathrm{d_i}$. The highway to total ratio $\mathrm{r_i}$ is then defined as:

$$r_{\mathrm{i}} = \frac{h_{\mathrm{i}}}{d_{\mathrm{i}}} \tag{3.1}$$

Table 3.1: **Truck routing calculation methods**

Shipment Type	**Case**	**Highway/Urban Ratio Calculation**	**Highway, Urban Average Speeds Calculation**	**CFS Ton-miles share**
Inter-state	Both origin & destination have MSA specified	Origin MSA area to destination MSA area in Google Maps	Google Maps	34%
	One of origin/destination has no MSA specified	MSA area to Rest of State centroid in google maps	Google Maps	42%
	Both origin & destination have no MSA specified	Rest of State centroid to Rest of State centroid in Google Maps	Google Maps	15%
Intrastate	Both origin & destination have MSA specified	Origin MSA area to destination MSA area in Google Maps	Google Maps	3%
	One of origin/destination has no MSA specified	FAF (2012)	Fixed to 55,25 mph	4%
	Both origin & destination have no MSA specified	FAF (2012)	Fixed to 55,25 mph	2%

We then use these calculated ratios for each origin-destination combination and apply them to the actual shipment distance from the CFS dataset. This allows us to calculate the highway and urban leg lengths $\mathrm{D_{S,H,i}}$ and $\mathrm{D_{S,U,i}}$ for the shipments in the dataset.

Let the shipment distance be $\mathrm{D_{S,i}}$. Then,

$$D_{\mathrm{S,H,i}} = D_{\mathrm{S,i}} * r_{\mathrm{i}} \tag{3.2}$$

and then,

$$D_{\mathrm{S,U,i}} = D_{\mathrm{S,i}} * (1 - r_{\mathrm{i}}) \tag{3.3}$$

3.3.3 Operator-hours calculation

The final step involves the calculation of urban and highway operator-hours. We assume the urban legs are equally split at the two ends of the journey with the highway leg in between. We apply a constraint of 11 hours of daily driving as per hour of service (HOS) regulations

(57) and then calculate the operator-hours required for the highway and urban legs of the journey. Using this information and the aforementioned weighting factor we are then able to calculate the total operator-hours as well as the share of highway and urban operator-hours.

Let day_1 hours be the number of hours remaining that can be driven on day 1 of the trip after completing the initial urban leg. Let O be operator-hours described for both highway leg O_H and urban leg O_U. Let highway and urban driving time be T_H and T_U respectively which can be calculated from the average velocities V_H and V_U for the respective segments also derived from the Google Maps API where applicable. Then for shipment i:

$$T_{\mathrm{U,i}} = \frac{D_{\mathrm{S,U,\ i}}}{V_{\mathrm{U,i}}} \tag{3.4}$$

and similarly

$$T_{\mathrm{H,i}} = \frac{D_{\mathrm{S,H,\ i}}}{V_{\mathrm{H,i}}} \tag{3.5}$$

Then, our algorithm to estimate the operator-hours is described below. Note that $\lceil(x)\rceil$ denotes the ceiling of x and x%y denotes the remainder of x when divided by y.

i from $1 : I$

$day_{1,\mathrm{i}} = 11 - \frac{T_{\mathrm{U,i}}}{2}$

if $day_{1,\mathrm{i}} \geq T_{\mathrm{H,i}}$ **then**

  $O_{\mathrm{H,i}} = T_{\mathrm{H,i}}$

else

  if $day_{1,\mathrm{i}} < T_{\mathrm{H,i}}$ **then**

    $O_{\mathrm{H,i}} = T_{\mathrm{H,i}} + \lceil(\frac{T_{\mathrm{H,i}} - day_{1,\mathrm{i}}}{11})\rceil * 10$

  end if

  if $(\frac{T_{\mathrm{H,i}} - day_{1,\mathrm{i}}}{11})\%11 + \frac{T_{\mathrm{U,i}}}{2} > 11$ **then**

    $O_{\mathrm{U,i}} = T_{\mathrm{U,i}} + 10$

  else

    $O_{\mathrm{U,i}} = T_{\mathrm{U,i}}$

  end if

end if

The algorithm can be explained as follows. If the highway driving time is less than the number of driving hours remaining on day 1, then the shipment is simply completed on the day and the highway operator-hours are equal to the highway driving time. However if the highway driving time exceeds this then the driver undertakes the journey over the following days with 10 hours of rest following 11 hours of driving as mandated by law. The urban driving time is simply the time taken to drive the urban leg if the second and final urban segment can be completed staying within the HOS requirements, else it is completed with a day of rest.

With the calculated urban and highway operator-hours for each trip we can then estimate the total operator-hours across both highways and urban areas using the trip weighting factor provided by the CFS dataset. The weighting factor is the estimate of the true number of trips of such type in the actual population and is available for each shipment in the CFS dataset. Let the weighting factor be Π. Further let shipment weight be W. Then for the total operator-hours $\mathrm{O}_{\mathrm{Total}}$ we have:

$$O_{\mathrm{Total}} = \sum_{i=1}^{I}(O_{\mathrm{H,i}} + O_{\mathrm{U,i}}) * \Pi_{\mathrm{i}} * \frac{W_{\mathrm{i}}}{TL} \tag{3.6}$$

where TL is truckload or the total weight that can be carried on one fully loaded semi truck.

Then the urban and highway share of the total operator-hours, US and HS, is simply:

$$HS = \frac{\sum_{i=1}^{I} O_{\mathrm{H,i}} * \Pi_{\mathrm{i}} * \frac{W_{\mathrm{i}}}{TL}}{O_{\mathrm{Total}}} \tag{3.7}$$

$$US = \frac{\sum_{i=1}^{I} O_{\mathrm{U,i}} * \Pi_{\mathrm{i}} * \frac{W_{\mathrm{i}}}{TL}}{O_{\mathrm{Total}}} \tag{3.8}$$

Note that the highway share (HS), across both inter-state and intrastate trucking, is the share of operator-hours at risk from automated highway trucking. US represents the share of hours that must still be driven by a human driver.

Notice that if the truckload is a constant, such as for e.g. fully loaded class 8 semi trucks, then it cancels in both the numerator and denominator of equations (3.7) and (3.8) and is therefore irrelevant to our results. More information on methods including the limitations of our approach are provided in Supplemental Information (SI) of the published manuscript (35).

3.3.4 Interviews

In order to obtain some assurance about the validity of the assumptions underlying the transfer hub model, we undertook semi-structured interviews with stakeholders in the trucking industry using a purposeful sampling methodology (60) which was formulated through authors' prior work in the automated vehicle domain (6) as well as prior informal conversations with companies and researchers in this area which helped us identify relevant questions and stakeholders. For our conversations with drivers in particular, we used snowball sampling (61): we identified an initial set of drivers who have a public profile (e.g., have podcasts or YouTube Channels about trucking) and asked them to introduce us to their colleagues. We stopped when we achieved data saturation (62): that is, when conversations with new drivers did not introduce us to new concepts or phenomena (63). We spoke with stakeholders across automated trucking startups (2), truck drivers (5), trucking logistics operators (1), and labor union representatives (1). In terms of our selection of different interviewees, we deliberately sought to elevate the voices of truck drivers in our sample, relative to other actors such as automated trucking startup CEOs or logistics operators. This is because of two reasons. Firstly, operators were best placed to provide us with the operational challenges and opportunities for the transfer hub model of automation, given they are currently in charge of the major task that automation may replace (driving). Second, much of the narrative and coverage around automated trucking in the popular media has focused on the claims made by private operators, without much consideration of whether long haul drivers themselves believe that a switch to automation is feasible. Note that our sample size was not designed to enable generalization to all the stakeholders in long haul trucking. Our sample size and qualitative method (semi-structured interviews) were instead selected with an idiographic approach (60), focused on gathering detailed insights into the tasks truck drivers performed on a journey. The interviews also highlight interesting areas for future research. Most importantly, as part of the triangulation method (56) we use to analyze automated trucking, the interviews complement our quantitative analysis of the CFS data and the routing and operator-hour algorithms we present by providing a feasibility check on the deployment modes assumed in this paper and which have been promoted by technology companies. The full list of interviewees is provided in Table 3.2.

3.4 Results

Our analysis finds that up to 94% of operator-hours for truck drivers are impacted if the technology is deployed across the continental U.S. in all conditions. However, if restricted to the states where testing is currently taking place, only 10% of operator hours are impacted. The capabilities of the technology and decisions around where and how AT should be deployed will therefore determine the extent of impacts on the long haul operator labor market. Below we first discuss the findings from our exploration of possible scenarios of deployment, and their associated impacts on operator-hours, if realized. We analyze the extent of the possible increase in short haul jobs if AT delivers cost and time savings in freight delivery and show that this is unlikely to outweigh the hours lost to automation. Then, we present the takeaways from our semi-structured interviews with stakeholders in long haul trucking. Finally, we end with a brief discussion on the labor impacts in sectors associated with long haul trucking.

3.4.1 Scenarios

We consider different scenarios of deployment which correspond to constraints the technology may face in the near-medium term. They are as follows:

- Deployment in southern, sunny states only - To our knowledge, significant AT testing currently takes place in Florida, Texas, and Arizona. We hypothesize that initial highway deployment will be restricted to states in the southern sun-belt (see Figure 3.2) to minimize risk of exposure to snow or hail which may be outside the limits of safe operation of the technology and therefore only consider routes between or within these states.

- Deployment across all states over the financial quarters Q2 and Q3 which encompass

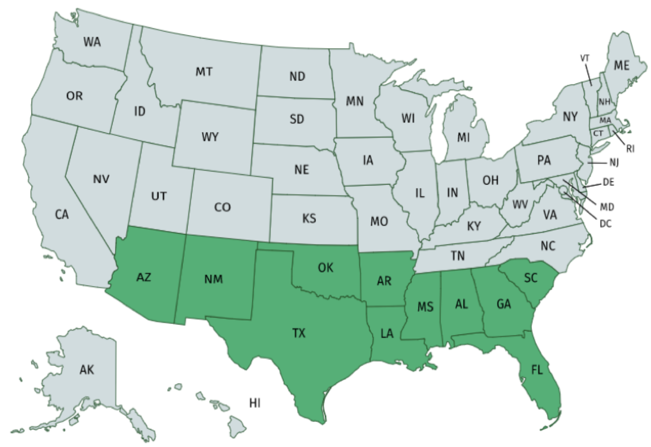

Figure 3.2: **Highlighted southern sun-belt states where a transfer hub model for automated trucking may first be deployed**

the spring and summer months from April 1 to September 30 - The CFS dataset lists the financial quarter of shipment allowing us to assess a scenario where all journeys that would be performed in more favorable weather are automated.

- Deployment for journeys above 500 miles only - While AT may be financially optimal at even short distances of around 100 miles (see SI Section 2 of the published manuscript), the time required for switching trailers and separating urban and highway legs could mean that journeys which currently cannot be performed in a single day given HOS requirements will be automated first. Long haul operators can comfortably cover 500 miles without stopping so we set this as the threshold.

- Widespread deployment - deployment of transfer hub based AT across the United States and automation of all highway driving. While this is likely to take several years, it can be thought of as the most extreme scenario in terms of job losses. Note that most studies in the literature (45; 44) only consider this scenario .

3.4.2 Operator-hours:

Figure 3.3 shows the impact on operator-hours across the different scenarios considered, with each scenario considered as additive (cumulative) to the previous scenario.

In scenario 4, with widespread deployment across the continental United States, 94% of current operator-hours may be automated. On the other hand, we find that deployment

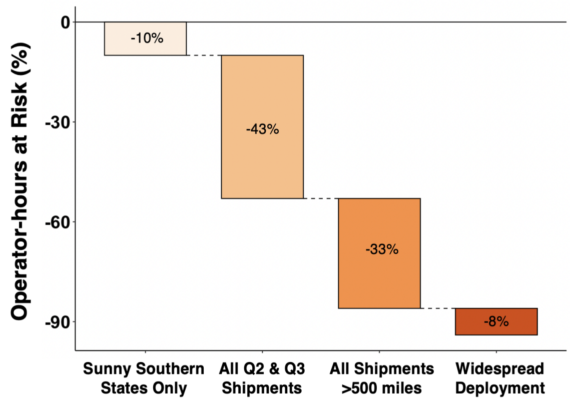

Figure 3.3: **The impact on operator-hours as each deployment scenario becomes feasible, cumulatively building on the previous deployment.** Starting with only a 10% impact on operator-hours, automation can eventually put 94% of current long haul operator-hours at risk

of automation only in places where companies are currently testing the technology, i.e. the southern states, limits the impact to just 10% of operator-hours (scenario 1). These estimates could represent a potential trajectory, over time, for impacts on employment, if the technology improves over time. It is likely that near term deployment, if any, will be in favorable weather conditions and in states with favorable regulation, precisely those states where testing is currently taking place. If this is then further expanded to all parts of the country but only in the favorable weather months, i.e. financial Quarter 2 and 3, then we find that over half of operator-hours could be impacted. If automation is then extended to all shipments over 500 miles, we find that a further 33% of operator-hours are impacted. Other possible deployment scenarios and their impacts are shown in Table S1.

3.4.3 Changes in demand and price elasticity of freight:

Automation is likely to dramatically reduce the need for labor, which constitutes roughly 40% of the cost of trucking (64). It is also likely to eliminate HOS requirements, which currently mean that—unless two drivers work in tandem—a truck remains idle for more than 60% of the time in order to allow the driver to rest (57). As a consequence, automation will reduce the amortized capital cost of the truck, which constitutes another 16% of the cost of trucking (64). Finally, the elimination of HOS requirements will make trucking faster than it is today, since the trucks will no longer have to stop for driver breaks. Cost and delivery time reductions in the delivery of freight through automated trucking may result in increased demand for trucking shipments. This may happen in two ways.

First, freight that was previously routed through air, train, or inter-modal services may now be shifted to long haul trucking. The modal choice of freight could therefore tilt in favor of trucking. There is considerable variation in modal elasticity estimates in the literature (65). For example, Abdelwahab (1998) (66) has estimated an elasticity of -1.44 and -0.99 for truck modal choice given an increase in trucking shipment time and cost, respectively. More recently, Christidis et al (2009) (67) reviewed a number of studies and found cross-elasticities of rail and road that range from 0.3 to 2. Dahl (2012) (68) has found limited elasticity for trucking demand to fuel costs.

Second, cost and time reductions in freight delivered through trucking may increase demand for freight services in the economy overall. This increase in demand will increase the overall operator-hours required to fulfill long haul trucking shipments.

While elasticities vary widely in the literature for these factors as described above, we find that any resultant increase in demand for freight from some or all of these factors is likely to have a small impact on operator-hours. Further, this increase is unlikely to offset the overall operator-hours at risk due to automation. For example, even a 50% increase in demand for trucking services, which translates to an overall elasticity of 5, will only offset 5% of the at-risk hours due to automation, dropping the overall share from 94% to 89% of operator-hours at risk. This is simply because even though such an increase will lead to great demand for trucking services, the large majority of all operator-hours will still be needed on the highway (Figure 3.3).

A large increase in demand of the magnitude of 50% or more for trucking services is in fact unlikely. Trucking already dominates the freight market with 70% of tonnage in the United States shipped through trucks (69). Freight that is today delivered through competing modes such as rail is often because of reasons other than cost or shipment time, for example, gross weight requirements, or because a close coupling between the location of the rail tracks and siting of various industrial facilities. Overall, we find limited evidence in support of the claim that increases in demand for trucking due to the economic and productivity gains from automation will create short haul jobs that will offset the highway operator-hours lost due to automation.

3.4.4 Jobs

What do these impacts on operator-hours imply for jobs? It is important to note that we do not directly characterize job losses from automation in trucking, focusing instead on the share of operator-hours that will be impacted. This is because data for the number of long haul truck drivers in the United States is not available to a high level of accuracy, owing to the large number of owner operators. Estimates have a wide range, from a few hundred thousand to millions of jobs. Studies which attempt to put a number on the total jobs lost therefore run the risk of a large error and are often not comparable due to the differing assumptions about the baseline number of jobs.

Nevertheless, translating the share of operator-hours impacted to different estimates of the number of long haul trucking operators in the United States can provide some insight. Previous analyses has estimated the number of long haul operators to be between 300,000 - 400,000 (44; 45). Waschik et al (2021) (47) estimate that there are roughly 550,000 long haul operators, in their modelling of the macroeconomic impacts of automated trucking. Our results on the share of operator-hours at risk from automation would therefore mean that anywhere from 30,000 to more than 500,000 jobs may be impacted, depending on the scenarios presented above.

3.4.5 Discussion with Stakeholders

Scholars note that typically, only some of the tasks that constitute a job, are amenable to automation (52; 42). As such, tasks are easy to automate; jobs are often not. Existing literature has noted how truck driving jobs have increasingly been reduced to solely the task of driving (44; 70). While this supports automation, one of the major challenges we envisaged to AT deployment on highways is maintenance and repairs. However, upon speaking with actors in the trucking industry (interviews 3-9) we found that employed truck drivers do not perform any significant maintenance and repairs on their truck. Instead, drivers simply call for assistance and trucking companies send out repair teams or arrange for a repair appointment at the nearest service station. ATs will therefore need to be able to send out distress signals and get assistance when needed. While some drivers indicated that they do perform some maintenance on their trucks (interview 7,9) this was restricted to minor fixes, for e.g. a broken headlight. Often drivers indicated that their companies did not want them to try to repair problems with the truck but instead rely on expert help. Some trucking jobs are union jobs and in those cases the drivers often have clauses in their contracts that restrict them to only driving and not having to perform work such as maintenance and unloading (interview 2).

The benefits of automation in terms of shorter trip times and lower costs could quite easily and quickly be incorporated into trucking operations and logistics, as per one of our interviews with a Senior Manager at a logistics firm (interview 4). AT startup executives acknowledged that both weather and lighter regulation motivated their decision to test in states such as Florida and Texas (interviews 1,3). This raises questions regarding the widespread application of this technology to all parts of the continental United States.

Interestingly, every operator (interviews 5-10) we spoke to said that they could not see any major barriers to ATs performing highway journeys. Several of them (interviews 5,9,10) highlighted the difficulties posed by inclement weather conditions, places where lane markings are absent, and routing if Global Positioning System (GPS) signal is lost, which could mean that some routes continue to be driven by human drivers. All drivers we spoke to try to use their maximum allowed driving time of 11 hours in a day. They all also indicated that there are often several weeks where they do not have any direct contact with their employers and simply perform their tasks as assigned. All drivers also indicated their trucks are tracked with GPS and therefore companies and customers are continuously aware of their location. Our discussions indicate that the job of long haul trucking has indeed effectively been reduced to a single task, which—given conducive external conditions—makes it amenable to automation.

We further offer some **propositions** for future research to explore. Due to the limitations of our sample size, the lack of segmentation between different types of drivers, and no consideration of potential confounders, these are not intended as conclusions.

Proposition 1: *A larger volume of shorter trips may not compensate for the loss of work associated with automated long haul trips.* A transfer hub model will require drivers to shift to short haul jobs. The most common pay structure currently for "truck-load" drivers, who haul full truckloads worth of generic containerized freight for a trucking company (and not for a shipper like Walmart or Target), is payment per mile of haulage. They are often not paid for the waiting and paperwork that occurs at the beginning and end of each trip. Truck load drivers seek to maximize the time they keep the cargo moving. A shift to shorter trips would increase the ratio of stationary (unpaid) to driving (paid) time, reducing their wages per hour worked (70). In our limited sample there was near consensus among our interviewees that the shorter trips and lower pay that may come from urban driving jobs will be unattractive (interviews 5-9). Therefore one of the interesting questions for future qualitative research based on a large sample size would be to understand operators' views on shifting to short haul jobs and whether these might hinder or accelerate a shift to the transfer hub model.

Proposition 2: *Transfer hub deployment could create short haul jobs in locations that are different from where long haul truckers currently live.* A shift to only urban driving will likely require operators to live in sub(urban) areas. At least one operator we spoke with expressed reluctance to shift to short haul trucking for this reason alone (interview 9). Around 40% of older truckers come from rural areas so it's possible that the geographical shift will prove a barrier to transitioning current operators to short haul jobs, which speaks to the cultural significance long haul trucking jobs have carried in the U.S. (71). However, these new jobs may prove attractive to new truckers joining the workforce, who are increasingly from urban areas (72). Again, this is an important question for future work and will require careful study. Shifting of employment from rural to urban areas for trucking will naturally have political implications, particularly given the existing rural-urban divide in the American political landscape (73).

Proposition 3: *Partial automation is viewed negatively by heavy truck operators, as previous studies have suggested* (74). Partial automation systems have also been criticized in the broader safety literature as they may lead to disengaged and distracted operators who are too quick to trust the technology and will be unable to react in a timely and safe manner should something go wrong (75). Many companies involved in self-driving technology are in fact skipping partial automation systems and focusing solely on full automation (76; 77; 78).

Economically, if companies still have to pay drivers for their labor and also pay for the system, it is difficult to see how this would be attractive compared to the current system of only paying drivers, unless the cost savings from increased safety were significant. All drivers (interviews 5, 7-10) in our limited sample expressed dislike for partial automation systems that they have used or experienced such as lane assist and emergency braking. How might views on such technologies differ depending on the age and experience of truck drivers? We believe that future work on the transition from human driven to automated trucks and the potential role of partial automation systems as a bridge technology must take into account the views of drivers, as they will be the primary users of such technologies.

3.5 Conclusion and policy implications

Automation of the major part of the job - in this case highway driving - will naturally put downward pressure on wages in the long haul trucking industry. It is unclear that the labor supply will easily adjust to the new level of prevailing wage and operating requirements (short haul jobs) in the market. Our limited number of interviews certainly highlight the challenges employers may face in transitioning long haul operators to different jobs such as short haul driving. It also suggests that the deployment of AT is being driven by techno-economic considerations alone with limited understanding of the social consequences, consistent with the broader narrative around automated mobility technologies (79). Although companies have claimed that such technologies will benefit truck drivers, our evidence does not suggest that the motivations of truck drivers are part of the designed operation of this technology. Moving away from industry led visions of AT futures will require a greater understanding of the motivations and interests of long haul operators, and a participatory approach to shaping AT deployment (80).

In the currently envisioned transfer hub model, short term adjustment costs are likely and potentially notable. As we show, a significant share of operator-hours will be affected if the technology is deployed in all conditions and locations. Further, we argue that this result is robust to increase in demand for freight delivered through trucking if as assumed, the cost of long haul trucking falls due to automation.

We do caution however that the potential loss of a significant share of operator-hours to automation need not be necessarily viewed as permanent unemployment or as a permanent welfare loss. Long haul trucking has been characterized by turnover rates of nearly 100% in recent years (81). The profession is increasingly unattractive to potential new entrants with most new operators lasting less than a year in the job. This has occurred in substantial part through a concerted effort to make trucking cheaper by paying drivers less; for example, by encouraging many drivers to operate as independent contractors (70). Wages may need to increase as these arms-length employment arrangements are challenged in court and as it becomes increasingly difficult to find new drivers. This dynamic may strengthen the economic case for automation.

Historically, technological change has resulted in short term employment shocks but realignment in the labor market means that these shocks have limited impact on the broader economy in the long run, as new industries grow and workers transition to new jobs with new skill requirements (53). Long haul operators may therefore move across to different sectors after a period of unemployment, some may transition to lower paying short haul jobs that will be created by the transfer-hub model, others may retire prematurely (47).

The sharp reduction in labor cost makes the economics of ATs compelling but will disrupt livelihoods and, by potentially shifting demand from rail to trucks, likely also increase emissions of greenhouse gasses and other air pollutants (82). The threat of jobs lost due to automation in trucking may also have profound political impacts, the existing literature has found increased support for radical right wing parties as the risk of automation increases (83). Policymakers could demand that—in exchange for permission to deploy ATs on public roads—truck operators re-invest some of the monetary benefits of reduced labor costs to ameliorate the disruption to employment and in reducing the environmental footprint of the trucking industry (84). Ultimately, societal and political choices can determine the mode of deployment of AT capabilities, and accordingly, the winners and losers of any shift to automation of long haul trucking.

Table 3.2: **List of interviewees**

Number	Job Title	Affiliation	Date
1	Chief Technology Officer	Automated Trucking Startup A	24th April 2019
2	Legislative Representative	Teamsters	14th August 2019
3	Chief Operating Officer	Automated Trucking Startup B	22nd November 2019
4	Senior Manager	Large Trucking Logistics Firm	10th February 2020
5	Trucker	JB Hunt	22nd September 2020
6	CDL Coordinator	CCAC	22nd September 2020
7	Trucker	Holland	15th October 2020
8	Trucker	Holland	20th October 2020
9	Trucker	Poly Trucking	21st October 2020
10	Trucker	American Service Line	12th November 2020

3.6 Data Availability

The Commodity Flow Survey (2017) dataset used in this paper is publicly available from the U.S. Census Bureau.

Acknowledgements

The authors would like to thank Venkatasubramanian Viswanathan, Jeremy Michalek, Nicholas Muller, Jessika Trancik, and Jack Stilgoe for their helpful comments on earlier drafts of this paper.

Funding

This work was supported by the CMU College of Engineering, Department of Engineering & Public Policy; and the Block Center for Technology and Society at CMU.

Informed consent

Informed consent was obtained from all interviewees.

CHAPTER IV

Agent based modelling of fully-electric taxis in ridesourcing

4.1 Introduction

Major TNCs such as Uber and Lyft have committed to a fully electric vehicle fleet by 2030 (85). Several cities, states, and countries in which major TNCs operate have also announced goals for full electrification of the light vehicle fleet. Going forward, urban areas will therefore see greater deployment of electric taxis and their share of the ridesourcing fleet will increase. Automation is also a major objective for TNCs. Lower operating costs and therefore prices due to automation may increase VMT and lead to congestion (34). A combination of these two technologies - automation and electrification - therefore has the potential to disrupt existing patterns of transport emissions and energy use. Governments must carefully manage both these transitions in order to support the eventual decarbonization of road passenger transport. Careful policy design is also needed to ensure that this transition does not exacerbate unequal access to transport (e.g., by cannibalizing public transport services or inequitable distribution of services).

This chapter presents a novel agent based model (ABM) that can answer a series of questions questions around automated and human-driven electric taxi deployment related to GHG emissions, energy use, equity implications such as impacts on transit, local air pollution, local electricity demand impacts, and more. While there exist several ABMs studying the impact of taxi fleets, I improve on the existing literature by building a model that can incorporate real life TNC trip data; draw on the labor economics literature regarding labor supply decisions of TNC drivers; and analyze in high geo-spatial detail the vehicular and infrastructure needs for all electric ridesourcing.

4.2 Literature Review

There has been strong research interest in modelling the operation of taxi fleets using an ABM approach. Broadly, the literature has largely focused on the replacement of the personal vehicle owned fleet with a shared automated vehicle (SAV) fleet. The research questions in these exercises have therefore been related to fleet size, waiting time, trip coverage, VMT, and other aspects of operating a fleet that can meet daily travel requirements. These studies do not consider the replacement of existing ridehailing trips with AETs but rather, the replacement of all personal vehicle travel with travel through a SAV fleet. Most of these studies also do not consider AEVs but focus on conventional engine AVs. An exhaustive literature review of ABMs for automated taxi fleets can be found in Jing et al (2020) (86).

An early example of such modelling is Fagant and Kockelman (2014) (87), who build an ABM to study how a SAV fleet can service simulated trips in a hypothetical symmetric grid. They find that a shared fleet can lead to significant reductions in emissions and the number of cars needing to be self-owned for personal travel(87). Boesch et al (2016) (88) also build an ABM using a well known ABM for transport, MATSim, to study the fleet size required to meet simulated travel trips in Zurich, Switzerland. They also find that fleet sizes can be substantially reduced compared to the original personal vehicle fleet if SAVs are widely adopted. AEVs are not considered. Liu et al (2017)(89) compare the probability of choosing a human driven vehicle to a SAV based on a stochastic process that involves four different fare levels for Austin, Texas. They find that longer distance travelers prefer SAVs to human driven vehicles and SAVs are also preferred to transit options for trips shorter than 5 miles which is the majority of journeys(89). Martinez and Viegas (2017) (90) assess the impacts of a SAV taxi and shared autonomous taxi-bus fleet in Lisbon, Portugal through a custom built ABM that draws on a synthetic population of trips within the city based on a mobility survey. They also find significant reductions in carbon emissions (up to 40%) with increased vehicle occupancy and despite increased VMT, through the deployment of shared automated taxis. This result of emissions reductions also comports with Liu et al (2017)(89).

Chen et al (2016) (91) do consider AEVs in their simulation and also use travel survey data for more realistic trip simulations in their study of a shared AEV fleet in a simulated grid based on Austin, Texas. They also consider the number and placement of charging stations in order to enable recharging and continued operation of the AEV fleet. Fleet size is found to be highly dependent on charging infrastructure and vehicle range, trade-offs between charging investments and EV range is considered in the paper through a financial analysis model (91) which finds a 80 mile AEV to be the optimal choice that minimizes costs while meeting travel demand. Chen et al (2016)(91) however do not constrain the number of chargers or the size of investment, therefore allowing for an unlimited number of stations. Trips are also considered for a broad use for travel purposes, i.e. replacement of personal owned vehicle travel with an SAV, rather than the specific use by TNCs or a taxi fleet replacing existing ridehailing trips. Loeb at al (2017, 2018) (92; 93) also perform a similar analysis for the same region, but with some additions for realism including a more realistic network grid for movement of agents and time dependent shortest routes for trips. The analysis however yields similar conclusions and once again does not draw on ridehailing trip data. Bauer et al (2018) (94) do consider AEVs and draw on taxi trip data for New York City to predict the battery range and charging infrastructure requirements for a fleet of automated electric taxis operating in Manhattan. Routing is achieved through the use of the Google Maps API which calculates distances and travel time for approximately 500 set of points in Manhattan (248004 point pairs), which are then used to interpolate values for trips between other locations.

Further work has included ridesharing capabilities in an ABM framework to allow for further reductions in fleet size and other operational targets in meeting travel demand (95; 96). Specifically, Lokhandwala and Cai (2018) (96) study ridesharing in automated taxis in the city of New York using trip data from the New York Department of Transport. They however do not consider electric AVs and associated charging decisions. Farhan and Chen (2018)(97) expand on Chen et al (2018)(91) and consider ridesharing for a shared AEV fleet for Austin, Texas. They find that ridesharing can decrease fleet size and the number of charging stations required but leads to significant increase in waiting times. Lu et al (2018)(98) apply an ABM built in GAMA, a platform for spacial ABM simulations to Ann Arbor, Michigan to study the environmental implications of an AV taxi fleet. Trip demands are simulated and matched with the US National Household Travel Survey (NHTS)

2009 data. Interestingly, they find that when using conventional internal combustion engine AVs, energy consumption, GHG, and SO_2 emissions are 16%, 25%, and 10% higher than conventional solo commuting due to unoccupied miles between trips, which is commonly referred to as deadheading. Lu et al (2018)(98) suggest that converting to electric AVs offers little improvement in environmental outcomes due to the high emission intensity of the local electricity grid.

ABMs for human-driven taxis are less common in the literature. Bauer et al (2019) (99) build an ABM which draws on data regarding driver shifts to study the infrastructure needs for ridesourcing fleets in New York City and San Francisco to be fully electrified. They use trip data from the NYC Taxi and Limousine Commission as well as the San Francisco County Transportation Authority for Uber and Lyft trips. For New York City, their model clusters origin and destination locations into cells with a total of 6,500 cells and trip times between these cells are pre-processed using the Google Maps API. For the city of San Francisco, they cluster origin and destination into zones and create a distribution of trip times and distances for each zone-zone pair. Any vehicular movement in the model is then assigned by randomly drawing from this distribution. As such, this work does not involve any in-model routing or step by step knowledge of where taxis are at any given moment. Rather, similar to the other studies reviewed above, once a vehicle is assigned a destination, it is provided the estimated time and distance to that destination. The model then stops tracking that vehicle until it has relocated to the destination point. This relocation strategy is a common feature in ABMs looking to minimize computational costs for routing and accordingly gives up on geo-spatial detail.

Finally, Moniot et al (2022) (100) build on the Highly Integrated Vehicle Ecosystem (HIVE) simulation model presented in Fitzgerald et al (2021) (101) to study direct current (DC) fast charging infrastructure needs for the city of New York drawing on driver shifts, home locations, and overnight charging access rates. Routing is achieved through the OpenStreetMap road network, coming closest in approach to the model I propose in the next sections. However, the network is limited to just 47,000 nodes, which is significantly smaller than actual city networks. For example, the city of Chicago, which I focus on in the next chapter, has 293,532 intersections.

From this overview of the literature I note four points of interest which motivate my work. First, a majority of the models reviewed do not use real world ridehailing trip data, relying on simulated data which in some cases is matched to mobility surveys. As such, these studies look to create a SAV fleet that can service all personal travel within the city, assuming that SAVs will completely replace the need for personal owned vehicles. Second, few studies consider battery-electric vehicles. Third, geo-spatial detail is often extremely coarse in these ABMs, with most routing and infrastructure siting done at the zone or block level which can lead to over or underestimations of trip distances, times, and energy use. It also precludes studying ride-pooling given that taxi locations are coarse and not tracked in every step. Fourth and finally, I note that many of the studies in the US have focused on the Austin region owing to the large transport research community located at University of Texas. A few others looked at New York City. I found no ABM studies for the city of Chicago, which I focus on in the next chapter.

Overall, I identify space for a contribution in the literature through using TNC dataset of trips, a focus on battery-electric ridesourcing vehicles, significantly improved detail on vehicle routing and mapping, and focusing on driver decisions rather than consumers as most transport ABMs have done previously given that it is drivers and vehicles which will have infrastructure needs and impacts.

Emissions and energy impacts of these studies were also estimated in the aggregate through a change in VMT multiplied by a constant emissions factor. Instead I will consider marginal, i.e. consequential emissions from vehicles charging during different times of the day as per their operations and the resultant emission impacts using long run marginal emission factors.

4.3 Model Research questions

The ABM is designed to enable research on a number of different questions relating to the deployment of electric vehicles in TNC operations across a number of different cities. With certain modifications to agent rules, it may also be possible to use the model to study future deployment of automated taxis.

I list a few questions of immediate interest here, noting that the first question is the focus of the next chapter.

- **Effects of battery size and charging infrastructure on life cycle externalities of Uber and Lyft**: If TNCs fully electrify, what are the likely lifecycle externalities from GHG pollution and VMT? How do those externalities vary across different battery sizes for vehicles and different levels of DC fast charging infrastructure?

- **Effects of charging infrastructure on equitable access to TNC services:** What are the impacts in terms of equitable access to mobility services depending on charger installations? Since many variables in the TNC dataset for Chicago such as pickup and drop-off locations are reported at the census tract level, I can augment these data with auxiliary datasets at the census tract level such as socioeconomic and demographics variables from the American Community Survey and transit supply variables from the General Transit Feed Specification. This will allow a study of how fully-electric TNC services might interact with riders across different socio-economic levels, as a function of charger placements.

- **How can fleet charging complement grid decarbonization strategies?** TNC vehicles could be responsible for the majority of DC fast charging demand due to their high utilization and requirements for charging during shifts. Existing literature has shown how EVs can play a role in decarbonizing the electricity grid, for e.g. by soaking up excess electricity during certain hours of the day which allows for greater penetration of renewables. How should incentives for DC fast charging during certain hours be designed and how might those impact driver decision making and TNC operations? Combining this ABM with a capacity expansion model could shed light on the value of alternative charging strategies on GHG emissions.

4.4 Methods

I present a novel ABM which is summarized in Figure 4.1 below. The model takes as input a 24 hour period of real world ridesourcing data from operating TNC fleets. Additionally, the model has as an input the Open Street Map based geographical layout of the city, direct current (DC) fast charging infrastructure spread across the city, and labor economics literature based estimates of driver shifts. There are three main types of agents in the model: customers, taxis/drivers, and charging ports (DC fast charging). Customers are dispatched into the model as per the TNC data on trip origin times. They have origin locations and destinations they wish to travel to and accordingly request a ride. Taxis fulfil those ride requests by matching with customers. Taxis enter the model in response to rising customer demand, and each taxi exits when it either hits its own revenue target for the day, or when its fuel drops below a critical threshold. Finally, charging stations earn revenue when taxis charge at the stations. Details on each of these agents and the agent rules that govern their behavior are discussed in the following sub-sections.

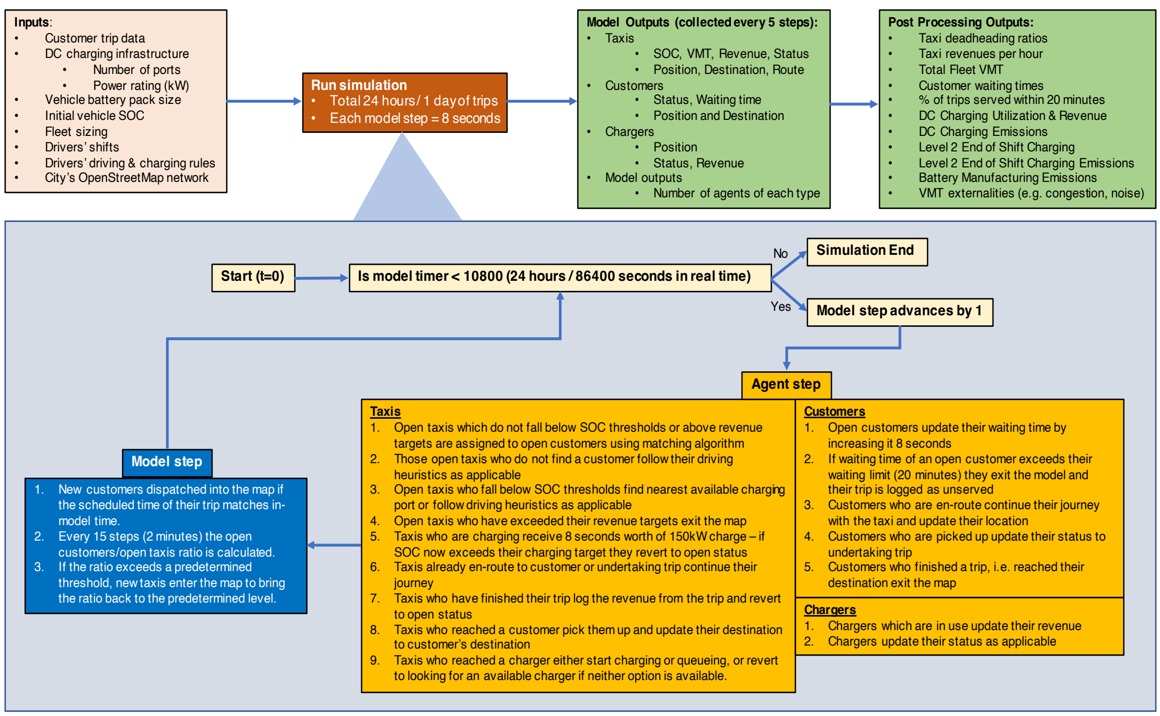

Figure 4.1: **Flow chart for the ABM showing model inputs, model stepping sequence along with agent actions at each step, and model outputs**

4.4.1 Agent Rules: Customers

Data on customer trips is obtained from real world TNC trip data. Each customer agent enters the simulation at the time of their travel. They search for a taxi. First, customers search for a taxi. More details on the matching between customers and taxis is provided in the sub-section on Fleet Manager. Once customers are picked up by a taxi, they move in the taxi to their final destination. After being dropped off, customers are removed from the simulation. All the time that they participate in the simulation their activity in the model is recorded, in particular their ID, location at each step, and how long they had to wait for a ride. I enforce a waiting time limit of 20 minutes for customers, i.e. if customers do not get picked up within 20 minutes they exit the simulation and the model records that the trip was not serviced. Customer waiting thresholds are common in the agent based modelling literature [e.g. - (89; 94; 99)]. Many studies assume a waiting time of 10 minutes while some others consider 15 minutes. I use 20 minutes as a threshold as it is more likely to correspond to real world threshold waiting times for Uber/Lyft as opposed to 10 or 15 minute wait times which are more aggressive metrics. This is particularly true in recent months given the ongoing driver shortage and longer wait times experienced by consumers as reported frequently in the media (102). As TNCs do not publicly reveal any data on their wait times it is not possible to benchmark this assumption to real world data.

4.4.2 Agent Rules: Taxis

Taxis or taxi drivers have attributes such as their ID, position on the map, their current destination and route, their revenue for the day, revenue target, and state of charge (fuel level). Each individual taxi enters the map at a random location, this is a necessary because there is no publicly available data on where Uber or Lyft drivers live. They then search for a customer. Once they match (rules described in the next sub-section on Fleet-Manager) with a customer they go to pick the customer up, then journey to the customer's destination, and start the process again. At any point of time during searching for a customer or searching for a charging station, i.e. if the taxi is not engaged in picking up a customer or doing a ride, if their fuel level drops below a certain threshold (20% SOC during peak hours and 30% SOC during off-peak hours), they look for the nearest charging station that is available. Fuel thresholds to look for refuelling stations is a common approach in the ABM literature [e.g. (87; 92; 99; 100)]. Lower thresholds were set for looking for a charger in peak hours because empirical data on charging by battery-electric vehicles in TNC fleets in California has shown consistent dips in charger utilization during peak hours as drivers try to maximize time on the road and earnings (3). It is assumed that drivers would have access to an app which would inform them of the available charging ports in the city. Once they identify a charger, taxis journey to the charging port and start charging on arrival, unless the port is taken before they arrive, in which case they queue. Queuing is however restricted to a maximum of 1 car in the queue per port, so if there is already a car charging at that port and a car waiting behind it, taxis look again for an open charging station. At any point of time during searching for a customer or searching for a charging station, i.e. if the taxi is not engaged in picking up a customer or doing a ride, if their fuel level drops below a critical threshold (10% SOC), they exit the map.

Taxis at charging stations charge back up to 80% SOC. I choose 80% as the threshold as DC fast chargers are meant to get vehicles up to 80% SOC in 15-20 minutes, but charging speeds are much slower after. Drivers in the middle of their shifts are therefore unlikely to wait long for the additional 20% energy.

Finally, drivers end their shift once they have crossed their revenue target for the day. More details on driver shifts and revenue targets are provided in the next Chapter when I instantiate this model for a particular study for the city of Chicago. Drivers can only end their shift when they are not engaged in picking up or dropping off a customer. At the end of their shift, I assume that drivers go back to their home location and charge back up to their starting SOC using a L2 charger, if their final SOC is less than the SOC they started the day with. L2 chargers are *not* tracked in the model as they are assumed to be widely available at or near each driver's home location. L2 chargers are also only assumed to be used at the end of the driver's shift, while DC fast charging is used in-between shifts and their utilization is tracked through the model (see Subsection 4.3.4).

4.4.3 Agent Rules: Fleet Manager

Matching of customers to taxis and vice versa are undertaken through the use of a virtual fleet manager. This would be analogous to the role that the TNC platform provides for customers and drivers.

Previous ABMs usually assign the nearest taxi to customers (87; 88; 89; 92; 94; 99). Certain metrics, such as taxi SOC can be used as a tie-breaking mechanism (94), for e.g. if multiple taxis are equidistant to a customer.

TNCs use a complex mix of rules and algorithms to match riders to drivers, and have individual-level data on customers' willingness to wait from their ride history, drivers' willingness to take certain rides and more. Given the lack of data and publicly available knowledge on such algorithms, I design a simpler matching sequence that builds on the literature while also staying relevant to what is known about Uber and Lyft's real world ride-matching rules.

Each open customer in the model searches for a taxi first within a 5km radius. All possible open taxis are identified. Then, the customer is assigned the taxi which has the minimum match metric. The match metric is an equally weighted sum of the distance of the taxi from the customer (measured in metres) as well as the revenue of that taxi. Note that distance is measured as per the spherical distance between locations on the earth's surface and not the precise driving distance on the road network in order to speed up computation. The motivation behind the heuristic of using both distance and revenue in matching customers to taxis is that TNCs attempt to equalize driver revenues across the fleet (103). Given a choice between a taxi that is 1 minute away and has earned average revenues for the day, and a taxi that is 2 minutes away but has earned no revenue so far, it is likely that TNCs would assign the lower revenue vehicle and reasonably expect the customer to wait an additional minute. How much weighting should be given to revenues and distance is clearly an open question. Seed runs demonstrated equal weighting to be a reasonable heuristic for meeting the preset performance metric of serving 95% of rides within 20 minutes. Simply assigning the nearest vehicle as most previous ABMs have done lead to certain vehicles having an out-sized share of revenues while others do not earn any money at all. Clearly, this is unlikely to be the operational model of TNCs.

If no open taxi is found, the search radius is doubled and a search is conducted over a 10km radius following the same process. Note that the search radius is limited to 5km in the first search to prevent an unnecessary number of calculations of the distance for taxis that are further away if there are already taxis present in the 5km radius. If no taxi is found again in the 10km search, the customer continues to wait and repeats the search in the next step, until the 20 minute maximum wait time has elapsed.

Each open taxi in the model searches for a customer first within a 5km radius. All possible open customers are identified. Then, the taxi is assigned the customer which has the minimum match metric. The match metric in this case for the taxi is an equally weighted sum of the distance of the taxi from the customer (in metres) as well as how long that customer has been waiting. The motivation behind this heuristic is that TNCs attempt to serve all customer demand in order to ensure customer retention. If taxis were only assigned to the nearest customer, as most previous ABMs have done, certain customers in some parts of the city would never receive a ride or the taxi fleet size would have to grow significantly to serve all demand. Further, given a choice between a customer that is 1 minute away and has waited for zero minutes so far, versus a customer that is 2 minutes away but has waited for 10 minutes already, TNCs might prefer to allocate the taxi to the customer that is further away as the nearby customer is likely to accept waiting for a bit longer for a ride. How much weighting should be given to waiting time and distance is again an open question. Seed runs demonstrated equal weighting to be a reasonable heuristic for the goal of meeting 95% of trips within 20 minutes.

If a taxi is assigned a particular customer, and that same customer had also been assigned that taxi from the customer search described above, which is conducted first, the two are matched together. However, if the driver does not match, either because their assigned customer actually matched with another taxi, or because there were no customers in the 5km radius, the taxi must decide what to do. In fact, one of the key questions from a model design perspective and also for environmental outcomes is what do drivers do when they do not find a ride? Again, I make use of heuristics to guide behavior. The location of a customer that was assigned to the taxi within that 5km radius but with which it did not match is used as a heuristic. The driver moves in the direction of that customer anyway, even though they were not matched. The real world behavior this is meant to imitate is that drivers can anticipate where demand is likely to be from previous driving days, and might drive in that direction while they are waiting to be assigned a ride. If the taxi however does not identify any open customers at all within 5km, they search within 10km. If they match based on the process described above, they are assigned the customer for a pickup. If they do not match, they simply start driving towards the area of the map which has typically highest demand, which for most cities is the downtown area. Again, this is meant to imitate the real world heuristics driver might use, as most TNC demand is downtown in cities, drivers are likely to gravitate towards the downtown area if looking for a ride, particularly if they are in less travel dense areas of the city where there are no customers within a 5km radius.

Previous ABMs have also grappled with the issue of relocating taxis when they do not find a ride. Fagnant and Kockelman (2014) (87) assume relocation by a central fleet manager or waiting in place for any time necessary. Chen et al (2016) (91) also assume variety of relocation strategies, again centrally controlled by a fleet manager. Bauer et al (2018) (94) assume that the fleet manager would have well trained algorithms to relocate taxis to future demand in order to meet trips within the 10 minute maximum waiting time. Since all these studies focused on automated vehicles owned by a central operator, such approaches can be considered reasonable. However, current TNC operations are not centrally owned and managed. Drivers make decisions on what to do and so I attempt to model driver heuristics and decisions.

4.4.4 Agent Rules: Charging Ports

DC fast charging ports are located at fixed points on the map, often with multiple ports at a single location/charging station as is common, and have attributes such as their position, their status (open or not), their queue status (open or not), and revenue. Tracking the utilization of chargers in each step by monitoring their status allows me to infer the level of charging demand at any given point of time in the day. Utilization and revenue both allow for identification of chargers which are better performing in terms of servicing TNC vehicles.

Chargers cannot be reserved by taxis but their status can be viewed by taxis in the simulation. As indicated above, this is meant to represent a real world application where drivers might be able to see which ports are free. Multiple drivers might select the same port to drive to at any given point of time and whoever reaches the port first would be the first to charge.

4.4.5 Model attributes

4.4.5.1 Fleet Sizing

Sizing the vehicle fleet to serve demand through the 24h period is a key question in the design of ABMs for ridesourcing operations.

As the model step is described in Figure 4.1, I increase the fleet size in response to the ratio of open taxis to open customers. Uber and Lyft do not have any way to bring drivers online apart from providing them price signals and incentives when they see consumer demand rising and insufficient taxis on the road. Two key questions from a model design perspective are: 1) What is the appropriate ratio of open taxis to open customers? 2) How quickly should the supply of drivers, i.e. the ratio, adjust as new customers come online?

Several seed runs were performed for the city of Chicago using TNC data to test exactly this, how many available taxis to available customers are required to be able to serve more than 95% of trips within 20 minutes while also not significantly increasing deadheading and reducing driver average revenues. I found that a 15:10 ratio is often sufficient for most days, i.e. a ratio of 15 available taxis to 10 waiting customers. Other days required a 25:10 ratio, i.e. 25 available taxis to 10 waiting customers. This led to more than 95% of demand being satisfied within 20 minutes with median waiting times of 2-3 minutes (more details are provided in Chapter 5).

Next, on the question of how quickly this ratio should re-adjust as new customers come online - the range can be once again reasonably inferred. Bringing more drivers online every few seconds in response to consumers coming online is likely to be a) inefficient - because other drivers may complete their trips in the next minute or so and become available again, leading to an oversized fleet and b) unrealistic - because TNCs do not have any physical way to bring drivers online and can only send them signals of price increases or incentives. Drivers will take time to respond to those and leave their home or current task to start driving. However, adjusting the supply every 10 to 15 minutes is likely to result in a significant amount of trips going un-served. As a base case the model re-adjusts to the ratio every 2 minutes in real-time or 15 steps in model time. From seed runs for the city of Chicago, every 2 minutes was found to be a reasonable interval. Naturally, this leads to periods where there is over-supply or under-supply, and in as much as this happens, it mirrors the real world where there may be periods with lower or over supply of drivers. Modelling analyses can incorporate sensitivity runs with different adjustment times, e.g. every 5 minutes.

Previous ABMs which focused on automated vehicles had centrally controlled fleets did not have drivers ending their shifts and dropping off. Fleet sizes were fixed through seed runs in order to fulfil all demand with set thresholds of time (for e.g. 10 minutes or 15 minutes) (87; 88; 25; 95; 89). Loeb and Kockelman 2018) (92) varied fleet sizes for different customer: taxi ratios. Bauer et al (2018) (94) chose 10 min as the maximum waiting time for customers and if no taxis could fulfil the request within that window from the entire fleet, a new taxi was created to serve that demand. As such, the taxi fleet grew gradually over the day, analogous to my approach described above. Bauer et al's (2018) (94) simulation however accordingly produces the minimum number of taxis required to serve all demand, whereas my approach does not explicitly seek to do so, as TNCs do not have full control over the number of drivers that drive on a particular day. Rather, my approach attempts to smooth an oscillating signal (the open driver to open customer ratio) to a more constant signal, similar to control theory approaches.

4.4.5.2 Map and routing attributes

The spatial aspect of the model works on the basis of Open Street Maps. .OSM files are loaded into the model which specify the geographical layout of the city in question, containing many hundred thousand intersections and the road network. Routing in the model can take place between any two road positions on the map. This model is the first to bring this granular level of geo-spatial detail to the movement of customers and taxis in agent based modelling for transportation. For example, the city of Chicago, which I use as a test case in the next chapter, has 293,532 intersections and 174,408 roadways. In order to speed up computation, customer pick up and drop off locations are approximated to the nearest intersection instead of the middle of a street - this is likely to also happen in the real world as in many cases taxis might only be able to stop at intersections or meet up with customers at those locations for a pick up. Vehicle speeds in the model can be exogenous or calculated in the model on the basis of the particular road the vehicle is driving on, for e.g. 40mph on highways, 20mph on city streets. For the purposes of my simulation, I set the speed to 10m/s or 22mph which is an approximation for the average speed of vehicles in urban areas. To simplify computation and increase speed, I do not compute route alternatives and routing is achieved through the shortest path algorithm, which is executed once for agents that need a route. Agents follow that calculated path until they reach their destination.

4.4.6 Computation

A high level of routing detail along with hundreds of thousands of customer trips and several thousand taxis naturally involves a heavy computational cost. This is no doubt one of the reason why most transport ABMs in the literature do not incorporate such level of detail and prefer facile routing setups such as hypothetical, highly simplified grids (87; 25; 91) or low resolution mapping with a few hundred nodes in a city (94).

To address the issue of computation cost I build the model in Julia. Julia is a relatively new, open source computing language built for high performance computing with several in-built parallel computing capabilities. The Agents.jl package in Julia (104) allows for the creation of ABMS, including spatial ABMs based on OpenStreetMaps. While Agents.jl offers its own stepping functions, these are not specifically designed to take advantage of parallel computing capabilities. In particular, routing the agents is the most computationally heavy part of the model. Given that several thousand agents in the model might need to calculate a route at any given point of time, using the default stepping functions would slow down the model considerably.

As such, I design a custom stepping function to allow routing computations to benefit from multi-threading capabilities that Julia offers. In each step, as agents update their attributes, those agents requiring a route in that step (for e.g. taxis who have just been assigned a customer and must now journey to pick them up) are collected. The routing calculation for those subset of agents is then threaded across multiple cores, taking full advantage of the native Julia multi-threading feature. This parallelization allows me to reduce run-time of the model by 50-80%, compared to a naive implementation which uses the default Agents.jl stepping functions or does not parallel process routing.

The second expedient means used to speed up computation is to avoid calculation of distances on the road network. Distance calculations between agents are required in the model, for e.g. when matching customers and taxis. However, instead of precise driving distances on the road network, I approximate distances using the spherical distance calculation, which calculates the distance between two latitutde and longitude points on a sphere (the earth's surface). This approximation may in some cases under-estimate driving distances between riders and drivers, but precise distances are not particularly crucial, as recall that the matching radius is already set within a certain range, to enable nearby taxis to be located for drivers.

Finally, overall model design is guided by the principle of avoiding, as much as possible, calculations that require full visibility over all agents. An example of this is matching riders to drivers. A top down method of doing this might seek to minimize total driving distance or time. In this case, given a set of vehicles, and a set of customer locations, one could ask - what is the appropriate allocation of vehicles to customers that minimizes distance? Or that minimizes time? Such a formulation is NP hard, as commonly described in the computer sciences, which means that the required solution time increases exorbitantly with size as more and more routes and combinations need to be processed. Instead, while my matching sequence described in Section 4.3.3 may not always result in the minimum driving distance or driving time required to fulfil all rider pick ups, it allows for bottom-up calculations where one does not need to calculate all possible combinations between agents, and instead agents self-select into appropriate matches. To what extent is the matching algorithm of Uber or Lyft likely to represent one or the other approach? Evidence suggests that Uber or Lyft do not assign drivers to riders based on overall fleet metrics because supply is highly individualized given rider and driver histories. Uber and Lyft offer prices and waiting times to riders on the basis of past data on willingness to pay, willingness to wait, current weather (for e.g. is it snowing or raining), and even the battery level on the individual's smartphone.

As a result of these model design steps, a full day (24h worth of trips) simulation involving more than hundred thousand customer trips and several thousand taxis can be completed on a personal laptop (8 cores) in just 3-4 hours. I also run some simulations on the Great Lakes Computing Cluster, which further speeds up computation due to the advantage of greater parallel processing capabilities (36 cores) and can complete certain 24h simulations in less than 1 hour of run-time.

4.4.7 Data Collection from Model Runs

Data is collected from the model every 5 steps (40 seconds in real time). In theory data collection can be performed at every step. However, this offers limited benefits while significantly increasing the size of the data frame in memory, which slows down simulations.

At the end of the simulation a dataframe is written into memory which contains information on all the attributes of all the agents that were active in that step. For a 24h simulation, this results in several million rows of data.

4.5 Model Summary

4.5.1 Future extensions

It is common for models to go through various stages of development. The ABM presented in this chapter is the first iteration. I plan to continue to develop increased capabilities and two important future extensions are listed here.

First, the high level of geo-spatial detail in this model compared to prior ABMs will be particularly useful for adding ridesharing or pooling as a model capability. This is especially relevant because media reporting on TNCs has shown that pooling will be a critical component of their operations going forward and their path to profitability (105). Currently, all trips in the model are considered as solo trips and there is no ability to pool different customers into the same vehicle. Going forward, one of the first steps in future model development will be to add ride-pooling which will naturally require careful design to avoid a significant computational cost.

Second, vehicle speed in the model is currently considered exogenous and constant, which is a limitation given that speeds can vary significantly by time of day due to congestion, the type of road (urban highway vs residential street), and road elevation. Incorporating more realistic speeds for vehicles will be a key consideration for the model going forward as this will impact trip times and energy use.

4.5.2 Limitations

ABMs by design involve a number of different assumptions regarding agent decision making. Capturing millions of complex interactions between humans, technology, and infrastructure that influence TNC operations requires a host of simplifying assumptions. I have detailed the assumptions underlying this model extensively in the Methods section. In particular, the assumptions regarding driver behavior (e.g. thresholds on when to charge and what to do when no ride is found) are homogeneous, i.e. all agents of a particular type behave in the same manner. In the real world, different drivers may have different decisions when faced with the same situation, as some may be full-time drivers while others might be part-time. Further, some drivers may have access to home charging while others may not which may influence charging decisions. Absent extensive real world data on Uber and Lyft drivers in terms of their home locations, habits, and preferences, it is difficult to ascertain how realistic some of these assumptions might be. Similarly, part of the challenge with designing a matching algorithm to match riders with customers is that TNCs do not publicly share their algorithms, which means that we are left approximating what might be happening in the real world, with limited information on how close that approximation might be.

Further, despite the careful computational design, 24h model runs still require several hours to run on a personal laptop, which limits the number of cases and sensitivity analyses that one can run. This means that the capacity to thoroughly interrogate all the underlying assumptions with a large number of sensitivity runs is constrained.

4.5.3 Next steps

The next chapter showcases a particular instantiation of this model, focusing on the research question of lifecycle externalities from fully-electric TNC operations in the city of Chicago, across different levels of charging infrastructure and battery pack sizes. Many of the modelling details and assumptions which might seem abstract in this chapter will come alive in the next, and help the reader understand how the model is operationalized with real world data to study a relevant research question.

CHAPTER V

Effects of battery size and fast-charging infrastructure on lifecycle externalities of Uber and Lyft

This chapter was developed with coauthors Parth Vaishnav and Jeremy Michalek. Manuscript *in preparation.*

5.1 Introduction

We use the TNC data from the city of Chicago as a test case for the model presented in Chapter 4 and study a specific research question related to lifecycle GHG externalities of fully-electric ridesourcing operations, across different decisions regarding level of charging infrastructure and battery pack sizing. The chapter is structured as follows. Section 5.2 provides a brief review of the existing literature on externalities of TNCs and studies on charging infrastructure for fully-electric TNC operations. Section 5.3 describes our methods including the TNC dataset we draw on, data pre-processing steps, model assumptions specific to this research question, and model setup steps. It also describes our process for calculating the external costs of TNC operations and the scenarios we consider. Sections 5.4 and 5.5 present our results and 5.6 discusses the implications of our findings. Finally, Section 5.7 notes some limitations of this work.

5.2 Literature Review

The potential impact of fully-electric ridesourcing services on GHG emissions has started to receive attention in the recent literature. A study focused on California found that converting ridesourcing to EVs will triple emissions reductions (3) compared to converting personal-owned vehicles, due to the larger VMT for vehicles used in ridesourcing. Greenblatt and Saxena (2015) (106) suggest *automated* electric taxi deployment will lead to up to 94% reduced per mile GHG emissions in 2030 due to rightsizing for trips, decreases in electricity sector emissions intensity, and higher efficiency vehicles gaining market share due to the need for more VMT. The analysis is based on city taxi data for New York, Denver, and San Francisco. Smaller size AEVs are considered along with conventional vehicles. Bauer et al (2018) (94) also consider AEVs and find that in Manhattan, a shared AEV fleet would reduce GHG emissions by 73% compared to an automated fleet of conventional internal combustion engine (ICE) vehicles. Bruchon et al (2021) (107) study TNC operations and find that when lifecycle air pollution and GHG emission externalities are internalized, the share of fully-electric vehicles in a least cost taxi fleet increases. They find GHG costs of 6-11 ¢/trip-mile for fleets that produce the lowest total cost (private and social costs).

Should electrification proceed on the basis of more chargers per square mile or bigger batteries? There have been no studies which have specifically looked at this question for TNCs and fully-electric vehicles, although Wenig et al (2019) (108) consider the trade-off between battery capacity and charging infrastructure for plug-in hybrid electric vehicles. They find that only a small share of long-range drivers benefit from a dense charging network. While the study does not explicitly consider vehicles in ridesourcing, its results offer evidence that TNC drivers, given their higher mileage, would be the biggest users of public charging infrastructure.

The required level of fast-charging infrastructure for TNC operations has been extensively studied in recent literature. For instance, Bauer et al (2019) (99) draw on data for ridesourcing trips in New York City and San Francisco and find that a network of three to four 50 kW chargers per square mile will allow the same level of service from fully-electric ridesourcing as from ICE vehicles and at lower cost. Nicholas et al (2020) (109) analyze the charging infrastructure required to support electric ride-hailing in the 2020–2030 period in multiple U.S. cities. They offer a number of conclusions including that dedicated chargers for ridesourcing can serve 4 to 10 times the number of drivers as public fast-chargers, widespread home charging can reduce public charger requirements by 70-90%, and that the timing of public use and TNC driver use of fast-chargers is likely to be complementary, i.e., peak usage hours for TNC drivers and the general public are at different times of day. Moniot et al (2022) (100) study taxi operations in New York City and find that roughly 1 150kW fast-charger is needed for 20 ridesourcing vehicles. Another analysis by some of the same authors studies fully-electric ridesourcing infrastructure needs in multiple U.S. cities and conclude that on average, 1 charger is required per 60 ridesourcing vehicles, albeit with significant variation across different cities due to geographical layout, temperature-related EV energy consumption per mile, home charging availability and more (110). Finally, a burgeoning literature has focused on the placement of chargers to best serve demand [e.g. (92; 111; 112)].

Overall, we found no studies in the literature which have studied the specific issue of externalities from fully-electric ridesourcing operations across different levels of charging infrastructure and battery pack sizing.

5.3 Methods

5.3.1 Chicago TNC trips dataset

The City of Chicago has mandated all active TNCs (Lyft, Uber, and Via), beginning November 2018, to publish their historical TNC trip demand, as part of the licensing for their continued operations. We use this dataset as an exogenous input into our model. The dataset for any given day lists the origin and destination of the trip (to the nearest census centroid), the time at which the trip started (to the nearest 15mins), the trip duration, trip distance, as well as trip fare and other attributes. The TNC Chicago dataset has been extensively used in transportation research (113; 114; 115; 116; 117) since its public release. Chicago is also one of the largest ridesourcing markets in the U.S. with ridesourcing making up about 3% of the total regional VMT (118).

5.3.2 Heterogeneity in customer demand

Demand for TNC services can vary considerably across days (for e.g. weekday vs weekend). The COVID-19 pandemic has also caused significant changes to demand. In the city of Chicago, the number of daily trips served by TNCs averaged roughly 300k prior to March 2020. Subsequently, demand dropped by as much as 80%, prior to vaccinations being readily available. Since summer 2021, the number of daily trips has started to show signs of recovery although they are still approximately 40-50% pre-pandemic levels.

To capture heterogeneity in customer demand and ensure that our modelling accounts for different types of TNC operation days, we simulate a number of different days. The selected days and the characteristics of these days are summarized in Table 5.1 below.

Table 5.1: **List of Days Selected for Modelling**

#	Date	Type	Number of trips
1	7th November 2019	Pre-COVID Fall Weekday	332,111
2	12th January 2021	COVID Winter Weekday	94,687
3	1st May 2021	COVID Spring Weekend	166,425
4	26th July 2021	Post-Vaccinations Summer Weekday	112,015
5	6th September 2021	Post-Vaccinations Labor Day Holiday	106,063
6	15th October 2021	Post-Vaccinations Fall Weekday	212,854
7	20th February 2022	Post-Vaccinations Winter Weekend	145,284

Our selection of days tries to capture different dynamics : a) pre-COVID vs peak COVID vs post-vaccination periods, b) weekdays vs weekends or public holidays; and c) Summer/Winter days vs Spring/Fall as the former would have cooling/heating requirements which would increase vehicle energy-use per mile. The post-vaccinations period was categorized as after more than 50% of the U.S. adult population were fully vaccinated, which was a threshold crossed on July 12th 2021 according to the Centre for Disease Control and Prevention (CDC). We deliberately biased our sample set to lean more heavily towards the post-Vaccinations period as there is increasing evidence that the business model of TNCs has changed in recent years, with more recent higher fares the norm as they pursue profitability (105).

The number of trips by hour across the 24h period on these different days is shown in Figure 5.1. Figure 5.1(a) shows the four weekdays in our sample. Figure 5.1(b) shows the three weekend or public holiday days we consider.

5.3.3 Data pre-processing

Three pre-processing steps were undertaken on the TNC dataset for the seven different days in our sample. These pre-processing steps were necessary due to data quality issues (missing

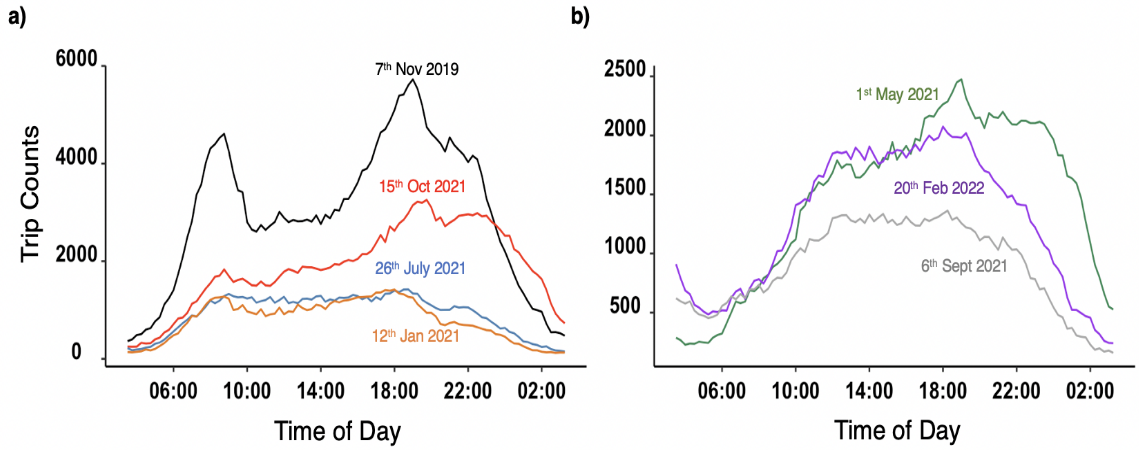

Figure 5.1: **Number of trips starting at different times of the day across a) weekdays and b) weekends or public holidays.** Note the different scale on the Y axis.

locations) as well as data resolution issues with regards to the precise timing of customer trips and start and end locations.

The steps are described below.

5.3.3.1 Data filtering

Roughly 10-15% of trips across the different days in our sample did not contain either a start or end location. These trips were removed from the dataset.

5.3.3.2 Enhancing Time Resolution

The Chicago TNC dataset aggregates trip start times in blocks of 15 minutes. The exact second the trip started is not provided to protect user privacy. Given the modelling setup described in Chapter 4, we need a way to assign a specific second in the day that the customer should enter the simulation and request the ride. To achieve this, we assign all trips within a 15 minute block (900 seconds) to start at random seconds within that block.

As the model steps through the 24h period, each block of trips that are within that model step (8 seconds as described in Chapter 4), are dispatched into the simulation.

Note that the Chicago TNC dataset timestamp describes the start of the trip itself, *not* the time when customers request a ride. That information is not publicly disclosed. Our method therefore approximates the trip start time from the Chicago TNC dataset as the time customers that requested a ride. Given low waiting times, and the fact that the Chicago

TNC dataset only aggregates in blocks of 15 minutes anyway, the timing discrepancy from this approximation is likely to be low.

5.3.3.3 Enhancing Geo-spatial Resolution

The Chicago TNC dataset aggregates trip start and end locations to latitude and longitude points of census tract or community area centroids. These offer coarse data on start and end locations compared to our modelling geo-spatial resolution, which as described in Chapter 4, is down to the street level.

To enhance the detail on start and end locations to exact street intersections, we undertake the following process. First, we draw a rectangular perimeter around the provided latitude and longitude location in the TNC dataset by increasing and decreasing latitude and longitude location by 0.01. This creates a box of approximately 1 square miles around the provided point which is approximately the size of a census tract. We then assign a random latitude and longitude point which corresponds to a street intersection within that box. This location is then assigned in lieu of the coarse centroid location.

Naturally, assigning latitude and longitude points in this manner may lead to underestimating or overestimating trip distances, compared to the actual trip distances which are reported for each trip in the Chicago TNC dataset. To ensure the error induced by this method is not large, we run the following test.

We randomly sample 2000 trips from a selected weekday and selected weekend from our sample of days (in this case 15th October 2021 and 20th February 2022 respectively). We then calculate the difference in driving distance between the our assigned latitude and longitude locations and the actual trip distance as reported by the TNC dataset. We use the Google Maps API to estimate the driving distances between our assigned start and end locations.

Figure 5.2 below shows the distribution of this difference.

In the large majority of cases the error induced by our method of enhancing geo-spatial resolution is zero or near-zero. We do find a relatively longer negative tail, i.e. we are more likely to underestimate the trip distance than overestimate it. Why might this be the case? One explanation could be that TNC trips are longer because customers are making detours

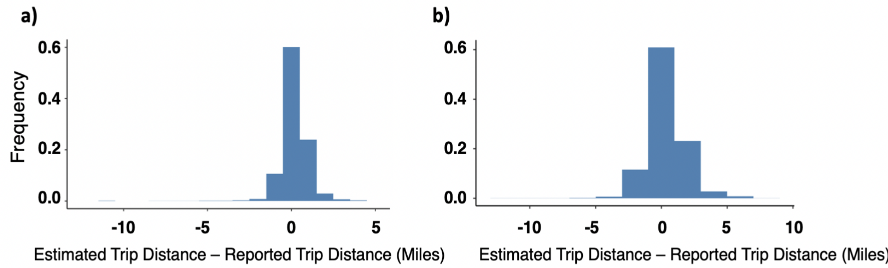

Figure 5.2: **Estimated trip distances versus reported trip distances for a random sample of 2000 trips on a) weekday of 15th October 2021 and b) weekend of 20th February 2022.**

en-route, either stops that they planned beforehand which made the trip longer or because drivers were forced to make detours due to construction or other unforeseen events. It could also be due to drivers taking a different route as opposed to that suggested by Google Maps. Finally, some of the trips in the Chicago TNC dataset are pooled trips which will involve longer routes and we ignore this characteristic and treat each trip as individual rider demand.

Overall, figure 5.2 provides evidence that our method of enhancing geo-spatial resolution is likely to induce a negligible level of error in our modelling.

5.3.4 Charging infrastructure

To site charging infrastructure in Chicago we tried to build on information regarding existing charging infrastructure as much as possible. This is because locating charging stations involves a number of regulatory decisions including zoning laws, electricity infrastructure permits, and more. It is reasonable to assume that future charging infrastructure would build on existing locations where permits have already been obtained and charging stations are already in operation.

The Department of Energy's Alternative Fuels Data Center (119) has a dataset of electric vehicle charging stations by location and type (L2 vs DC fast-charging). We found that there are currently 118 DC fast-charging ports within 25 miles of Chicago. At a base level, the existing 118 ports were assumed to be capable in the future of fast-charging at 150kW and exclusively available to TNC drivers. This roughly corresponds to an average of one port every 2 square miles for the city of Chicago, although these chargers are not uniformly distributed throughout the city. Note that our research is not focused on answering how much charging infrastructure needs to be built for all-electric ridesourcing. Other work in the literature has looked at this specific question as highlighted previously in Section 5.2. Answering the question of infrastructure requirements would naturally require estimates on how much of the DC fast-charging infrastructure would be used by non-TNC vehicles. Instead, we simply consider different levels of charging infrastructure availability, and assume that those stations are exclusively available to TNC drivers when they need them. In reality, this would of course correspond to a certain greater number of stations overall, given appropriate information on non-TNC usage. Furthermore, TNCs such as Uber and Lyft have already begun exploring contracts with companies building charging stations for their drivers to exclusively access such stations.

At a second higher level of charging infrastructure, we double the ports to correspond to a charging infrastructure level of roughly 1 port on average per square mile in Chicago. To site the 100+ additional stations, we use the Level2 charging dataset from DOE. The dataset currently lists 677 L2 ports in Chicago. We run a seed run of the model with these 677 chargers and identify the top 120 ports in terms of utilization. We then assume these ports could be converted to DC fast-chargers from their current L2 speeds. Again, we assume that these 230 stations would be exclusive to TNC drivers, which would imply a higher number of charging stations overall given non-TNC usage.

As described in Chapter 4, drivers are assumed to only use DC fast-charging infrastructure during shifts. At the end of a shift drivers are assumed to drive back to their home location and use L2 charging capacity near their home, which is assumed to be widely available, i.e. there are no fixed number of ports and therefore queuing or waiting assumed for these chargers. All vehicles charge back up to their initial SOC by the end of the day if they finish with less than what they started with, which ensures that the total energy needed for the 24h of trip demand is balanced.

5.3.5 Battery pack sizing and vehicle energy use

We consider two different levels of battery pack size. The first small battery pack of 40kWh corresponds to the base model of the 2022 Nissan Leaf with an EPA rated energy consumption of 0.3kWh/mile and 149 miles of range (120). We then consider a second, larger battery pack size of roughly 60 kWh which corresponds to the Nissan Leaf PLUS with an EPA rated energy consumption of 0.31 kWh/mile and a range of 226 miles (120).

We use the EPA rated energy consumption as the default energy use per mile for spring and fall days. However, for simulations that represent summer or winter days, we increase the energy use by 30% to reflect cooling or heating needs (121).

5.3.6 Driver Shifts

To assign shift lengths to drivers we draw on the labor and behavioral economics literature which has studied the labor supply decisions of taxi drivers, including drivers for Uber and Lyft. A central finding in this literature is that the majority of TNC drivers operate part-time and that the value of flexible work is significant for drivers (103; 122). Camerer et al (1997) (123) have also shown previously than taxi drivers make labor supply decisions "one day at a time" and have loose income targets in mind for the day.

TNC data for *vehicles* from the city of Chicago also indicates this. We ran exploratory tests for the share of vehicles that drove a certain number of hours in each of the months that are covered in our sample days in Table 5.1. In each case we found that a large majority of drivers are driving part-time, or less than 8 hours per day. Hall and Krueger (2018) (103) estimate that just under 15% of Uber drivers may be driving full time in the city of Chicago. However, their analysis was conducted based on data in October 2015 and there is increasing evidence to suggest a slight increase in the share of full-time drivers since then as Uber has consistently sought to increase its share of full time drivers, particularly since the pandemic receded and driver shortages are more common (124). A study of TNC drivers in Seattle in 2020 found that a third of drivers work more than 32 hours per week (125). Given these different data points, we adopt the following distribution for driver shift lengths. We assume 20% of drivers work full time, i.e. more than 7 hours of work per day. 50% of drivers are assumed to work between 1-3 hours per day and the remaining 30% between 4-6 hours of work per day. We do not correlate the likelihood of a particular type of driver (full-time vs part-time) to the time in the day they are driving as there is no data to guide this assumption.

A larger share of part-time versus full-time drivers works in our model by leading to more vehicles being required to come online to fulfil demand. This naturally leads to an increase in traffic externalities as drivers enter the simulation at random points in the map (see Chapter 4, Section 4.4.2). Drivers then have to drive to where demand is if they cannot find a ride, increasing VMT. More vehicles also leads to greater battery and vehicle manufacturing related externalities (see next subsection). More part-time drivers however reduces the need for DC charging infrastructure, as drivers are likely to only need DC fast-charging once during their shift, and can charge at the end of their shift using L2 chargers.

5.3.7 External cost calculations

To estimate the GHG emissions associated with charging events we make use of the National Renewable Energy Laboratory's Cambium Tool which provides long run marginal emission factors (LRMEFs) of CO_2 equivalents for 134 balancing areas in the United States from 2020 to 2050 (126). We set the year as 2030 to correspond to the deadline Uber and Lyft have set to fully electrify their fleet. In the baseline scenario we draw on the mid-case projection for the evolution of the U.S. electricity grid. Then given time t in seconds, the LRMEF in kg/MWh denoted by λ, and charger power draw in MW denoted by P, we have emissions E_{charging} in each time step across all charging ports as:

$$E_{\text{charging, t}} = \sum_{i=1}^{I} \lambda_{\text{t}} * P_{\text{i,t}} * (1/3600) \tag{5.1}$$

where i represents each individual charging port.

Total charging emissions for the 24h period are then simply the summation of emissions across all time steps:

$$E_{\text{charging, total}} = \sum_{t=1}^{T} E_{\text{charging, t}} \tag{5.2}$$

To estimate vehicle lifecycle emissions we use vehicle and battery manufacturing emission estimates from Argonne National Laboratory's Greenhouse gases, Regulated Emissions, and Energy use in Transportation model (GREET) (127).

If ρ is the battery manufacturing emissions factor for vehicle j per kWh, β is the vehicle manufacturing emissions factor per kWh, L_V is the total expected lifetime mileage of each vehicle, and S is the battery pack size, then manufacturing emissions associated to each vehicle given its shift VMT as V are calculated as:

$$E_{\text{m, total}} = \sum_{j=1}^{J} \frac{V_{\text{j}} * S_{\text{j}} * (\rho_{\text{j}} + \beta_{\text{j}})}{L_{\text{V}}} \tag{5.3}$$

All parameters used in Equation 5.3 are summarized in Table 5.2 below. Total expected lifetime mileage for EVs is assumed as 120,000 miles given modelling suggests it as the threshold for batteries to reach 80% capacity (6).

Table 5.2: **Parameters for estimating GHG related externalities**

Symbol/Name	**Description**	**Value**	**Units**
ρ	Battery manufacturing emissions factor	0.1	tCO_2/kWh
β	Vehicle manufacturing emissions factor	0.125	tCO_2/kWh
L_V	Expected lifetime mileage	120,000	miles
SCC	Social Cost of Carbon	62	\$/tCO_2

To calculate GHG related external costs G_c we then multiply total emissions from both battery manufacturing and charging events by the social cost of carbon (SCC):

$$G_{\text{c, total}} = (E_{\text{charging, total}} + E_{\text{m, total}}) * SCC \tag{5.4}$$

The assumed value for the SCC in year 2030 is $62 per ton of CO_2 in 2020 dollars. This is obtained from the U.S. Government Executive Order 13990, based on the 3% average discount rate (128).

Finally, we estimate non GHG related externalities including collisions (9.0¢/mile), congestion (6.0¢/mile), and noise (0.02¢/mile) as per-mile externalities following previous studies (117; 129).

5.4 Results

The results are organized as follows. First, we present results for fleet metrics including the share of deadheading for vehicles, the median time to start charging once a vehicle starts looking for a charger, the percentage of trips aborted, total taxis, and total fleet VMT. Next, we present the results of total societal costs, including both GHG externalities and VMT based externalities. Finally, we present profiles for the charging power demand across different scenarios showing how the level of public fast-charging infrastructure influences the timing of peak demand as well as where that demand accrues (L2 vs fast-chargers).

5.4.1 Fleet metrics

Table 5.3 below shows the fleet metrics across 3 different days which represent pre-COVID, during COVID, and the post-vaccinations phase. For each day we present four scenarios as discussed previously: two levels of battery pack sizing (40kWh vs 60kWh) and two levels of fast-charger infrastructure (1 dedicated fast-charger per 2 square miles vs 1 dedicated fast-charger per square mile).

Table 5.3: **Fleet Metrics for Selected Days**

Day	Pack Size (kWh)	DC Port / Sq-mile	Total trips	Total taxis	Median pickup time (mins)	Median dead-heading (%)	Fleet VMT (miles)	Trips Aborted (%)	Median time to find a DC port (mins)
7th Nov. 2019	40	0.5	285559	38479	2	36%	2832174	3%	63
7th Nov. 2019	60	0.5	285559	33045	2	37%	2780636	3%	87
7th Nov. 2019	40	1	285559	30220	2	38%	2517612	3%	31
7th Nov. 2019	60	1	285559	27632	2	39%	2472732	3%	27
12th Jan. 2021	40	0.5	76986	8375	4	30%	601218	2%	25
12th Jan. 2021	60	0.5	76986	7390	4	31%	575003	2%	26
12th Jan. 2021	40	1	76986	6806	4	33%	532912	2%	5
12th Jan. 2021	60	1	76986	6866	4	33%	529232	2%	4
20th Feb. 2022	40	0.5	121869	14944	3	28%	994297	1%	31
20th Feb. 2022	60	0.5	121869	12940	3	28%	952621	1%	35
20th Feb. 2022	40	1	121869	11595	3	29%	843284	1%	5
20th Feb. 2022	60	1	121869	11597	3	29%	836310	1%	4

Note that the ratio of open taxis to open customers was set as 25 taxis to 10 customers for 7th November 2019 and 15 taxis to 10 customers for the other 2 days shown in Table 5.3. As described previously in Chapter 4, the ratio is calculated and fixed every 2 minutes, forcing new taxis to come online if the ratio has fallen below the pre-determined threshold. The threshold itself is determined by seed runs which seek to limit the percentage of trips aborted to less than 5%. Trips are considered aborted if customers have to wait more than 20 minutes. Different levels of comparisons are possible from the data presented in Table 5.3. We first discuss the comparison across a single day. We can see that for a particular day, while other metrics change across different scenarios for battery pack size and charger infrastructure, median pickup times and the percentage of trips aborted remain the same. This is because these two metrics are most heavily influenced by the open taxis - open customers ratio discussed above, and the ratio is fixed for a particular day. Next, for any particular day, the number of taxis involved in the 24h period varies significantly across the different scenarios, with the most number of vehicles required with small battery size and low levels of charging infrastructure. The intuition here is straightforward, with smaller batteries and low charging infrastructure, more vehicles are out of circulation either because they spend a long time trying to find an empty fast-charger (see the column on median time to find a DC port which represents the amount of time it took a vehicle to start charging after it started to look for a port) or because they abort their shift as their SOC falls below a critical threshold (10% SOC). As these vehicles are not available, and the ratio of open taxis to open customers is fixed and adjusted back every 2 minutes, new vehicles are forced to come online. The real world representation of this could be an imbalance in supply of taxis and rider demand which forces TNCs to attempt to provide incentives for drivers to come online. Increased vehicles naturally leads to increased emissions from vehicle manufacturing (discussed further in the next section). As the battery size is increased keeping the level of charging infrastructure constant at 1 charger per 2 square miles, the number of vehicles required falls as larger battery packs enable taxis to do more trips before they have to go charge and go out of circulation. On average we found that most vehicles were able to complete their shift with just one visit to a fast-charger with a 60kWh battery pack, while a 40kWh battery pack required two visits to a fast-charger. Note that once charging infrastructure is set to 1 dedicated fast-charger per square mile, the trade-off between pack size and number of vehicles is not present on 2 of the 3 days shown in Table 5.3, as time to find a charging port drops drastically and vehicles are able to quickly recirculate back into the fleet and be available. The exception is 7th November 2019, a heavy pre-pandemic day, where even 1 dedicated charger per square mile is not sufficient for quick turnarounds, and median vehicles spend about 30 minutes either driving around or queuing before they can start charging, leading to a noticeable difference in the number of vehicles between the two pack sizes. This shows how infrastructure needs can be strongly influenced by peak demand, although with the caveat that ridesourcing demand may never recover to these pre-pandemic levels as explained previously in Section 5.3.2.

Total fleet VMT is another metric that is heavily influenced by the level of pack size and charging infrastructure. Fleet VMT noticeably falls for larger batteries and more chargers, as vehicles spend less time looking for a DC port *and* have to go to charge less often during their shift. This suggests that if the goal is to minimize VMT related externalities such as congestion, larger batteries may be preferable.

Finally, on deadheading ratios. We define deadheading as the share of miles a vehicle drives looking for a ride or going to pickup a customer, divided by its total VMT in the simulation. First, we note that as the level of charging infrastructure increases, the share of miles spent driving around looking for a ride or going to pickup a customer increases, simply because the number of miles driving looking for a charger (which is included in total VMT in the denominator) drops significantly. Second, we note that deadheading ratios are almost the same across different scenarios for a particular given day, again suggesting that what influences deadheading the most is the ratio we set for open taxis to open customers and the particular spatial layout of customer demand on a given day. Increasing the ratio leads to more deadheading, as there is increased competition for customers among taxis and taxis may not immediately find customers (over supply). How do our deadheading ratios compare to the existing literature on TNC operations? Anair et al (2020) (130) summarize the range of deadheading ratios spanning 20-47% with Chicago specific estimates at 41-45% (131). There are a few reasons why our estimated ratios might be lower than existing estimates. First, because a non trivial amount of miles is spent looking for and going to a fast-charger, particularly compared to what drivers might drive to go to a gas station, the total miles increases which reduces the share of deadheading in our simulation. Second, we pull vehicles off the road completely in our model when they hit 10% SOC and additionally direct them to go charge at 20-30% SOC depending on peak/off-peak hours which can be considered conservative, especially compared to ICE vehicles which might be less concerned about wasted fuel and more inclined to keep driving even if they do not find a ride as they are confident in being able to find a gas station rapidly when needed. Third, TNC services in the real world have a tiered offering with certain vehicles classed as premium (e.g. Uber Black), i.e. the more expensive service where drivers may have a lower frequency of rides and therefore responsible for increasing deadheading in the fleet. Our model considers all vehicles to offer the same type of service and there is no differentiation among vehicles. Fourth and finally, one of the heuristics for drivers in our model when they do not find a ride is to go towards the customer in the 5km radius they did not match with (as described in Chapter 4 Section 4.4.3), which is clearly not a data point available to real world drivers, who must rely only on their instinct and experience rather than real time data of where consumers are. This might further increase deadheading.

5.4.2 Societal costs

We first present average GHG external costs across all the sample days in our dataset, across the four scenarios we consider for battery pack size and fast-charging infrastructure. Figure 5.3 below shows that if GHG costs are all that matter, lowest costs are obtained with a smaller battery pack size of 40kWh combined with 1 dedicated fast-charger per square mile. The worst outcome from a climate standpoint is with large battery packs and limited charging infrastructure - this results in high vehicle manufacturing emissions and higher charging related emissions as vehicles waste a lot of energy driving trying to find charging ports.

The averaged results in Figure 5.3 while useful mask heterogeneity across different days. Figure C.1 in Appendix C shows the plot of GHG costs for each of the 7 days in our sample. We see that the order of rankings remain the same across different days, lowest costs are obtained with a smaller battery pack size of 40kWh combined with 1 dedicated fast-charger per square mile, followed by a smaller battery pack size of 40kWh combined with 1 dedicated fast-charger per 2 square miles and with the worst outcome being a larger battery pack size of 60kWh combined with 1 dedicated fast-charger per 2 square miles. However, the magnitude of the external costs vary by as much as 5¢ per trip across different days, for the same combination of pack size and charging infrastructure. GHG external costs are usually

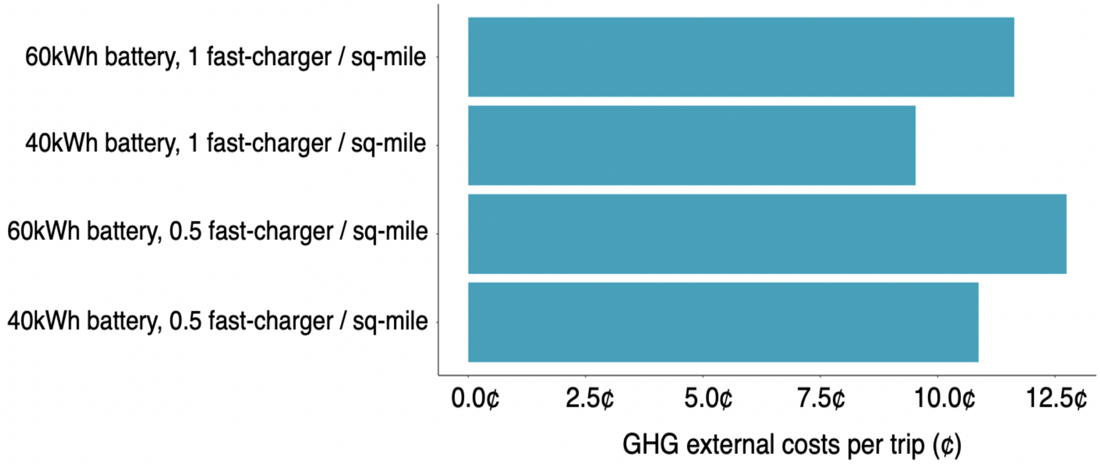

Figure 5.3: **GHG external costs per trip averaged across the sample days.** Lowest GHG costs are obtained for a combination of smaller pack size of 40kWh with 1 dedicated fast-charger per square mile. This results in a 12% reduction in GHG costs compared to the second best scenario and a 25% reduction compared to the highest cost scenario.

higher in Summer and Winter days, reflecting the energy penalty of heating and cooling needs which increases energy/charging requirements per trip and therefore emissions.

Next, we consider the total external costs by including traffic externalities (noise, congestion, crashes). Figure 5.4 below shows the average across the seven days for total external costs for the four scenarios considered for battery pack size and charging infrastructure. Interestingly, we find that while the lowest costs continue to be offered by a combination of smaller battery pack size and a higher level of fast-charging infrastructure, the remaining rankings shift considerably. Due to the much higher VMT induced by limited charger availability, a combination of the 40kWh pack size and just 1 dedicated fast-charger per 2 square miles results in the worst outcome for total external costs, even though previously it reflected the second best outcome for GHG externalities.

In fact, adding in costs for VMT externalities results in both battery pack sizes having almost the same level of externalities, if we keep the level of fast-charging infrastructure fixed (Figure 5.4). This reflects the fact that higher GHG emissions are compensated by lower VMT externalities and vice-versa. The difference in total external costs between the lower and higher level of charging infrastructure is roughly 12% on average. Note that we do not consider the cost of installing charging infrastructure itself because once that is amortized

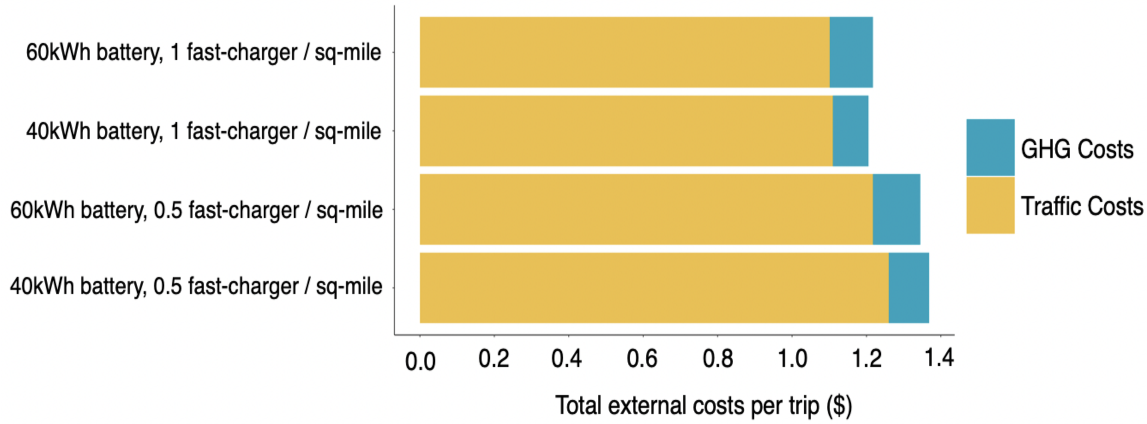

Figure 5.4: **Total external costs per trip averaged across sample days.** Lowest external costs are obtained for a combination of smaller pack size of 40kWh with 1 dedicated fast-charger per square mile. Total external costs are 11% lower with 1 dedicated fast-charger per square mile compared to 1 dedicated fast-charger per 2 square miles.

across every trip, the costs of the investment are small. Does that simply mean investing in an extraordinarily high amount of public fast-chargers is the way forward to reduce externalities? Apart from the fact that there are other barriers to installation including grid connectivity and land acquisition, the fleet metrics in Table 5.3 suggest not. Apart from the pre-pandemic day which might not represent future travel demand as TNCs increasingly pursue profitability, other days see a drastic fall in the time taken to start charging once a vehicle starts looking for a charger as soon as dedicated fast-charging infrastructure is doubled from 0.5 to 1 charger per square mile. If vehicles are already able to obtain a charging port within 4-5 minutes (Table 5.3), comparable to finding a gas station today, it is difficult to see large reductions in VMT externalities accruing from even higher levels of infrastructure.

Total external costs for each individual sample day are shown in Figure C.2 in Appendix C. Across all days considered, the ranking of scenarios remains the same with lowest costs seen with a 40kWh battery pack size and 1 dedicated fast-charger per square mile. There is heterogeneity in total externalities across different days for the same scenario, for e.g. ranging from less than $1 per trip to nearly $1.5 per trip for the lowest cost scenario.

What do our results suggest about future TNC externalities if the electricity grid is zero-carbon? Clearly, if charging and battery manufacturing based GHG emissions both fall to zero, VMT externalities will be the sole concern. Our results suggest that in such a case, larger battery packs are preferable, as they reduce the need for the use of public fast-charging infrastructure during shifts, and therefore reduce additional VMT. If however a zero carbon electricity grid only impacts charging emissions and not vehicle manufacturing emissions as batteries and vehicles are imported from other jurisdictions, lowest costs are once again obtained with a smaller battery pack size and higher level of charging infrastructure (see Figure C.3 in Appendix C).

5.4.3 Charging power profiles: Level 2 vs fast-charging infrastructure

While the previous sub-section focused on societal costs, here we look at the power demand for vehicle charging across both slow and fast-chargers for the different scenarios we consider. Note that as outlined previously, we consider L2 chargers to be widely available near each taxis' home location. We assume that taxis drive back home (their initial start location) at the end of their shift and plug into the charger upon arrival and charge until they reach back up to their initial SOC. Finally, a minority of vehicles do not charge upon arriving back at home as they may have used a fast-charger during their shift and when their shift ends, their final SOC might be higher than what they started the shift with.

Figure 5.5 below shows the charging power draw from both slow and fast-chargers on a Spring weekend during COVID (1st May 2021). We show the charging demand between 12 noon and 12 midnight and not the full 24h to avoid sudden jumps or drops due to our simulation start and end points. With a lower level of dedicated fast-charging infrastructure (Figure 5.5a and Figure 5.5b), there is a stronger reliance on L2 charging. However, with the higher level of charging infrastructure shown in Figure 5.5c and Figure 5.5d, drivers are able to easily access fast-chargers which means they do not need to abort their shift and reduces the need for further L2 charging at home at the end of their shift.

These trends are also confirmed on a summer weekday post-vaccinations (26th July 2021) shown in Figure 5.6. Again, increasing the level of fast-charging infrastructure reduces the power draw from L2 chargers. Note that L2 power demand is far less fluctuating than fast-charging demand due to the much slower speed of charging (5kW vs 150kW) which means that vehicles stay plugged in for extended number of hours and power demand can

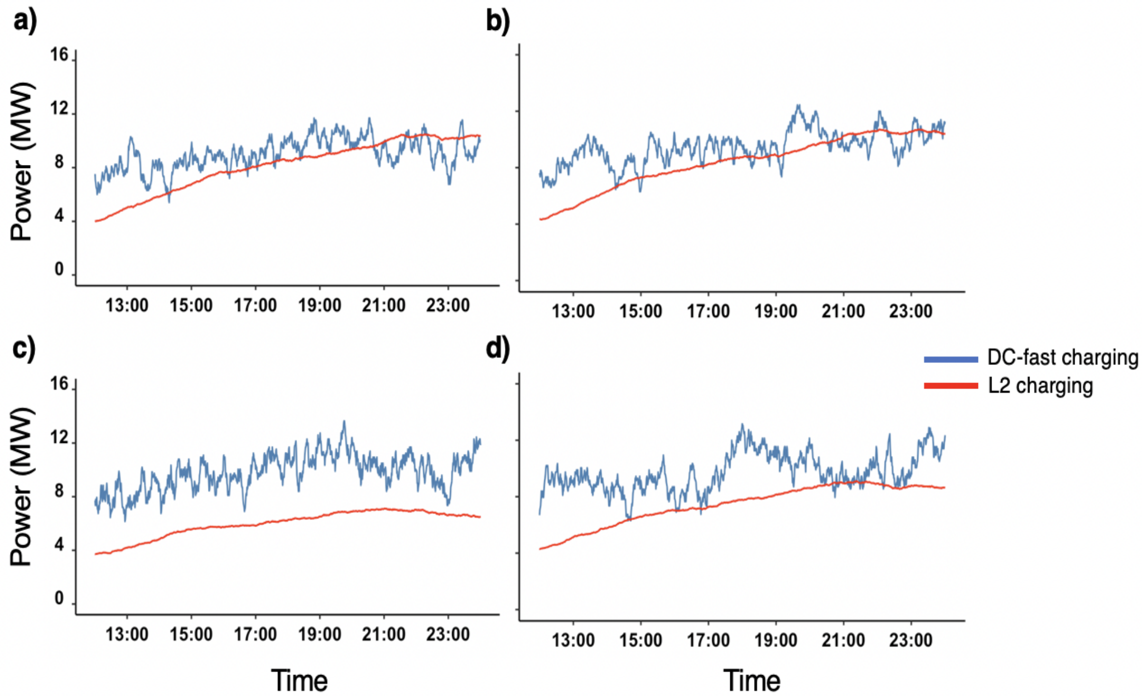

Figure 5.5: **Fast-charger and L2 charger power draw for 1st May 2021 a)** and **b)** represent fast-charging infrastructure of 1 dedicated fast-charger per 2 square miles and the small and large battery pack respectively. **c)** and **d)** represent fast-charging infrastructure of 1 dedicated fast-charger per square mile and the small and large battery pack respectively.

cumulatively build up as more and more vehicles plug in during the day at the end of their shifts.

In Figure 5.6 we can also see the effect of peak weekday evening hours of trip demand between 5pm and 7pm. There is a dip in fast-charging demand during those hours as vehicles stay out driving longer to maximize revenues given the lower SOC threshold set for peak hours (see Chapter 4 Section 4.4.2). Power demand however climbs back up quickly towards the end of the peak-demand period.

Finally, the most extreme case reflecting the trade-off in demand from slow vs fast-chargers is shown in Figure C.4 in Appendix C. Here we consider the heaviest trip demand day in our sample set, which is the pre-pandemic weekday of 7th November 2019. With lower level of fast-charging infrastructure, L2 power demand is double that from fast-chargers as more and more drivers cannot find fast-chargers and end their shift to use slower home charging.

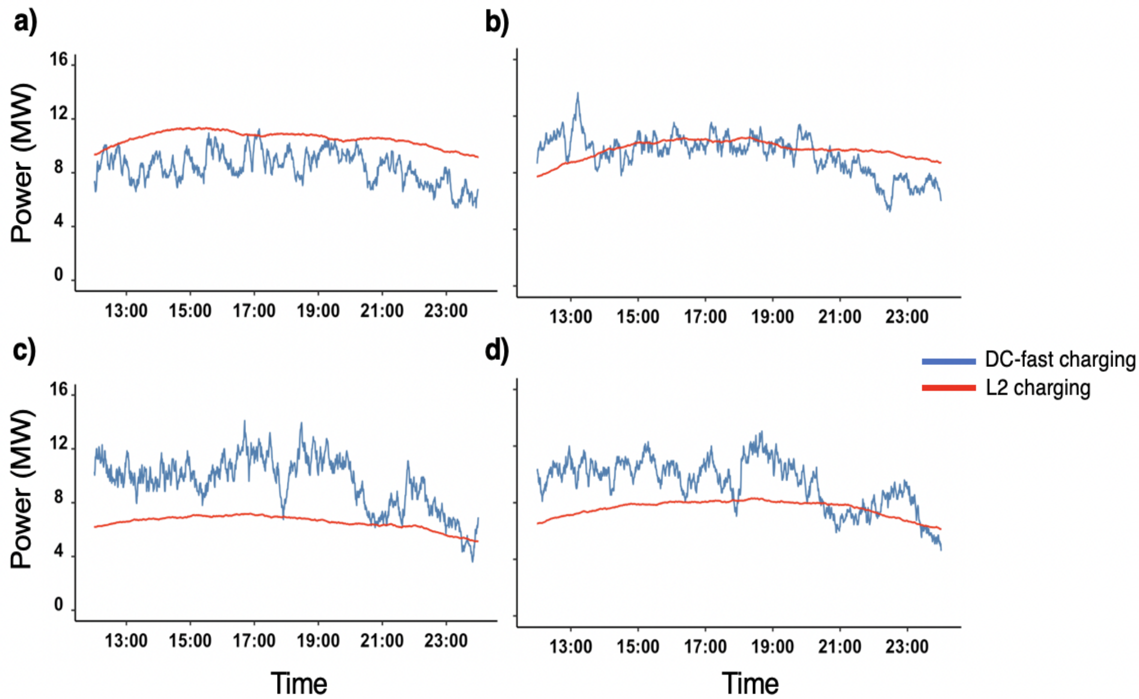

Figure 5.6: **Fast-charger and L2 charger power draw for 26th July 2021 a)** and **b)** represent fast-charging infrastructure of 1 dedicated fast-charger per 2 square miles and the small and large battery pack respectively. **c)** and **d)** represent fast-charging infrastructure of 1 dedicated fast-charger per square mile and the small and large battery pack respectively.

5.5 Sensitivity Analysis

We undertake a brief sensitivity analysis around two major inputs in our estimates of external costs.

5.5.1 95th Percentile SCC

The first sensitivity analysis is related to the valuation of climate damages. While we use a central estimate of \$62/tCO_2 for the SCC in our base analysis, existing literature has shown the wide variations in estimates of the SCC. We use the 95th percentile estimate for the SCC in 2030 with a 3% discount rate, which corresponds to \$187/tCO_2 in 2020 dollars.

Increasing the SCC by 3x clearly increases total external costs. Figure C.5 in Appendix C shows total external costs across the four scenarios averaged across the sample days. We find that the lowest and second-lowest total external costs continue to be with the case of a 40kWh and 60kWh battery pack size respectively with 1 dedicated fast-charger per square mile. However, the ranking of the third and worst option changes compared to the base case. Due to the higher valuation of climate damages, a 40kWh battery with 1 dedicated fast-charger per 2 square miles is preferred to the larger battery with the same level of charging infrastructure.

Total average externalities per trip increase by 16-18% for each scenario as a result of the higher SCC. Overall, ascribing higher valuations for the social cost of GHG externalities does not significantly change our results.

5.5.2 Adjusting driver supply

In Chapter 4 there was a detailed discussion on the adjustment of driver supply during a 24h period, i.e. how the ratio of open taxis to open customers is set at a pre-determined threshold through seed runs. The number of taxis is increased every 2 minutes if needed, i.e. every 2 minutes new taxis are brought online if required, in order to keep customer waiting times down. As described in Chapter 4, TNCs do not have any way to bring drivers online when demand for rides exceeds vehicle supply, apart from providing incentives and nudges to drivers to come online and start driving. While our method of checking the ratio every 2 minutes offers reasonable results as verified by model calibration to real world fleet metrics (see Section 5.4.1 and Table 5.3, here we double the adjustment time to 4 minutes. What might this change? If the supply of drivers is only brought back up to the set ratio every 4 minutes instead of 2, it doubles the time intervals during which demand might be far exceeding supply. However, increasing the adjustment time might also mean that vehicles which were previously engaged have more time to come back online, therefore reducing the possibility of over-supply of drivers. In terms of our research question, reducing the possibility of over-supply would mean reducing deadheading, as taxis have less competition with each other to be assigned customers. It would also reduce the number of vehicles on the road and therefore vehicle manufacturing related emissions. On the other hand, increasing the adjustment time might mean greater periods of under-supply, which might prolong customer waiting times but reduce deadheading. There is also the possibility that these dynamics cancel each other out, and overall fleet metrics remain the same as when supply is adjusted every 2 minutes.

We ran the simulations for the most recent day in our dataset, 20th February 2022, which is our best approximation of post-pandemic demand. Comparison of fleet metrics across all four scenarios for the base case and the case where we double the adjustment time are shown in Table C.2 in the Appendix. We find limited variations in key fleet metrics such as deadheading and customer waiting times for the same scenario, across the two different adjustment times. This suggests that periods of under or over-supply are not overtly extreme in only one direction. External costs are also virtually unchanged (less than <1% difference) for each scenario of battery pack size and charging infrastructure, across the two different model setups. Overall, this suggests that our results are not sensitive to the parameters used in our method of adjusting driver supply.

5.6 Discussion

We estimate the external costs in terms of GHG emissions and traffic externalities such as congestion from fully-electric ridesourcing in Chicago. Total external costs are in the range of $1.2-1.4 per trip, with the large majority induced from traffic externalities such as congestion, crashes, and noise. GHG based externalities are in the range of 10-13¢ per trip.

A smaller battery pack size of 40kWh combined with a charging infrastructure level of approximately 1 dedicated fast-charger for TNC vehicles per square mile yields the lowest external costs overall as well as the lowest GHG emissions. Overall costs from this scenario are very close to a larger battery pack size of 60kWh with the same level of charging infrastructure, despite the higher battery and vehicle manufacturing related emissions. This is because inclusion of traffic externalities means that larger battery packs gain in relative preference, as they reduce the need for frequent visits to fast-charging stations. In as much as real world deployment would be a mix of larger and smaller battery pack sizes and not a homogeneous fleet, we find that there is no sizeable benefit in terms of societal costs for policy intervention to promote one option over the other.

Power demand for TNC fast-charging is significant, in the order of a dozen MW, and shows patterns of peaking in the evening post the 5pm-7pm traffic rush hour. This might mean that TNC demand is complementary to fast-charging demand from personal owned vehicles, which would reduce need for fast-charging infrastructure to be dedicated for TNCs.

Local electricity distribution systems may require upgrading to handle peak demand periods and planning for such will require closer collaboration between TNCs and utilities.

How do our estimates for the external costs of fully-electric ridesourcing compare to existing estimates for ICE vehicles? Ward et al (2021) (129) found that GHG external costs are roughly 25-30¢ per trip for Chicago, indicating that EVs can reduce the GHG externalities of Uber and Lyft by 50-60%. In the case of 1 dedicated fast-charger per square mile which for most days in our sample set results in vehicles finding available DC fast-chargers in 4-5 minutes, which is clearly comparable to finding a gas station for an ICE vehicle, our estimates for traffic based externalities are about $0.9-1.2 per trip. This is within the range of Ward et al (2021) (129) who estimate roughly $1 per trip for Chicago.

5.7 Limitations and future work

We note a few limitations with this work. As described previously in Chapter 4, we do not consider pooling in this model. Trips in the Chicago TNC dataset which are pooled are treated as solo rides. Ignoring pooling has the effect of increased VMT which both increases energy use and therefore charging emissions, while also increasing traffic externalities. However, on a per mile basis, externalities such as congestion and noise are far greater than emissions. As such, inclusion of pooling is likely to further strengthen the case for the smaller 40kWh battery size and 1 dedicated fast-charger per square mile. As noted previously, future developments of the ABM will prioritize the inclusion of pooling capabilities.

Second, the current iteration of this work ignores local air pollutants, i.e. particulate matter based externalities. In a future iteration we will include this analysis, using reduced complexity air pollution models to estimate the health impacts from charging and battery manufacturing emissions. Previous work on the externalities of ridesourcing has shown local air pollution to be a significant share of total external costs (107; 129).

Finally, results presented in this work are specific to Chicago. Other cities with different geographical layouts, patterns of trip demand, and composition of electricity grids will incur different external costs. There is value in regionally specific analysis and future work must involve extending the modelling presented here to other cities in the U.S. with public datasets on TNC demand including New York City, San Francisco, Boston and others. If more local governments require ridesourcing companies to share their data publicly, this will facilitate more generalizeable insights.

5.8 Data Availability

All datasets used in this study are publicly available. The TNC trip-level data for Chicago can be accessed via https://data.cityofchicago.org/Transportation/Transportation-Network-Providers-Trips/m6dm-c72p.

CHAPTER VI

Conclusions and future work

6.1 Findings and policy implications

I began work on my PhD in Fall 2018. At the time, it seemed a consensus had formed around the view that the future of mobility would be electric, shared, and automated (132). The evidence behind this vision for the future of road transport was, however, limited upon closer interrogation, and drew mostly from industry claims regarding 1) the speed at which these developments could materialize; and 2) the willingness of the public to accept such technologies. A number of events since then have proven conclusions about how passengers and goods will travel in the future as premature. For starters, despite several more optimistic industry claims on deployment, fully automated (driverless) vehicles are only commercially available today as geo-fenced taxi services in certain areas, for e.g. Waymo's service in Phoenix, Arizona, and Cruise's operations in San Francisco, California. A series of high profile setbacks including a fatal crash involving an Uber testing vehicle in 2018 have led to several companies exiting the race entirely, while others have been forced to merge forces. Second, the COVID-19 pandemic has added further uncertainty into the mix, as visions of shared mobility have come under scrutiny. Third, the majority of companies engaged in testing automated vehicles use hybrid-electric vehicles, raising questions about how they might transition to fully-electric while still powering the computation and sensors required for self-driving.

Regulators have struggled to keep pace with ongoing technological developments, evidenced most clearly by the limited oversight of partial automation systems that are misleadingly marketed by certain companies as fully capable of self-driving. Academic literature on these topics has rarely straddled both the policy and technical realms. My work aims to fill that gap. In Chapter 2, we draw on engineering methods to show how the energy use of automation might impact EV range by 10-15%, illustrating the trade-off between two emerging technologies. Since that work was published, which showed that the energy penalty of automation may not be as large as some feared, California published a first of its kind legislation (SB-500) which requires self-driving companies to fully-electrify light automated vehicles by 2030 (133). Such legislation might show the way for other jurisdictions to follow suit and motivate self-driving companies to invest in energy-efficient computing, aerodynamic integration of sensor stacks, and consider the broader, systems-level impact of their technology. Policymakers might also consider providing permits for testing on the basis that companies commit to deploying fully-electric automated vehicles in the commercial versions of their technology.

Long haul trucking has always been seen as the low-hanging fruit ripe for automation. The vast majority of miles are on restricted access highways, which is a less complex domain with limited types of vehicles and simpler driving rules. While a limited number, the truck operators we spoke to themselves indicated that their job has been reduced to only driving, and that they can envisage automation completing highway legs. Given that truck driving is one of the most common professions in the United States for men with a high school diploma, employment impacts have always loomed large in discussions about truck automation, even making their way into presidential primaries. Companies engaged in developing this technology have claimed that drivers will still be required on the short-haul legs to and from highways to urban distribution centers, and that the sheer increase in volume of freight routed through cheaper automated trucks will more than compensate for the operator-hours lost on the highway. We interrogated the employment impacts from automation of highway driving in Chapter 3, and found that contrary to industry claims, there is no feasible scenario in which an increase in short haul driving hours makes up for the hours lost to automation on the highway. Further, automation of just the highway legs will result in up to 94% of operator-hours lost to automation, if the technology can function in all conditions. However, most companies are currently restricted to testing in good weather in states such as Arizona, Texas, and Florida, which would constitute just 10% of current operator-hours. Our work is the first to consider scenario-based deployment, assessing different levels of technological capabilities, and showing the share of operator-hours that are impacted as the capabilities of automation mature. Splitting up the journey into long and short haul legs can entail significant opportunities for electrification and therefore reduce the environmental impact of trucking. While electrifying the long haul leg over highways is still challenging due to battery weight constraints, short haul legs of 25-50 miles near cities can be electrified with existing battery technology. A majority of the population lives near cities and is exposed to pollution from heavy diesel trucks, making a strong case for differentiated powertrains in trucking. In this case, automation may actually strengthen the potential for electrification. Policymakers might therefore demand that automated trucking companies commit to partial electrification, in return for the societal license to operate automated trucks and reap the significant savings in labor costs.

Finally, agent based modelling of taxi fleets has received significant attention in the academic literature. However, most studies focus on shared, automated fleets that serve all travel demand within a city. In the near-medium term, widespread electrification of ridesourcing is more likely than automation. I build an ABM that can answer a number of research questions on the impacts of EV deployment in ridesourcing. We find that for the city of Chicago, fully-electrifying the ridesourcing fleet across different levels of battery pack size and charging infrastructure lead to GHG externalities of 10-13¢ per trip, with lowest emissions associated with a smaller battery pack size. This would reduce emissions externalities in Chicago by 50-60% compared to today's fleet based predominantly on ICE vehicles. Interestingly, however, when traffic externalities based on miles of travel such as congestion and crashes are considered, larger battery packs gain in preference as they reduce the need for frequent visits to fast-charging stations. In as much as real world deployment would be a mix of larger and smaller battery pack sizes, we find that there is no notable benefit in terms of societal costs for policy intervention to promote one option over the other.

Overall, the work presented in this dissertation attempts to carefully disentangle the uncertainty around societal impacts from deployment of emerging technologies - automation and electrification - in road transport. Considerable uncertainty remains. Helping policymakers navigate it will require an active role for researchers and hopefully, an increasing role for the types of analyses presented here.

6.2 Methodological insights

6.2.1 Importance of model prototyping

Working with large datasets and incorporating a high level of geo-spatial detail both come at significant computational cost. As detailed in Chapter 4, the ABM that I developed requires several hours to run on a personal laptop despite drawing on advanced parallel computing techniques to speed up performance. As such, in the model development stage it is vital to consider model prototyping at a much smaller scale. For example, in the case of agent based modelling, I first developed the zeroth iteration of the model without any mapping detail, instead just using a simple grid network. Then, once I had basic agents and attributes set up and tested, I increased the level of geo-spatial detail to model on a small, 1 square mile cutout of the full Chicago map. This small section had just 500 nodes compared to the full OpenStreetMap for Chicago, which as described previously has nearly 300,000 nodes. I was able to develop and verify significant chunks of model code and complex agent-agent interactions just through this small model. The small scale also enabled fast visualization of the model stepping in Julia, allowing me to quickly identify bugs, troubleshoot, and verify. Data outputs from the model were also significantly smaller in size, allowing for much quicker post-processing and analysis of results. I finally scaled up to the larger map when I had full confidence in the model performing as expected across a number of tested edge-cases. Scaling up brought its own challenges, some in terms of agent-agent interactions and model functioning, but mostly in terms of model speed and performance. But at this stage, due to the extensive prototyping work undertaken previously, which helped to quickly address remaining modelling issues, I was able to fully focus on speed improvements and achieve dramatic speed-ups through dedicated efforts on those lines.

The work on long haul trucking involved a similar experience. The analysis was conducted in R and involved processing the CFS dataset which involves several million rows of data as well as Google API based routing of shipments. I first worked with a random sample of a smaller subset of shipments, developing and verifying code and model processes until the pipeline from data input to results was working smoothly. Taking this process to the full dataset was then a straightforward extension of what had been already developed.

6.2.2 Value of public datasets

Publicly available datasets are the lifeblood of research on public policy. Data-heavy methods, such as those presented in this dissertation, are only as good as the quality of data that feed into them. I leveraged both the Commodity Flow Survey dataset on trucking shipments as well as a TNC dataset on ridesourcing trips for Chicago. In particular, the latter was released as part of a mandate by the City of Chicago that forced TNCs to release this data as part of their licenses to operate in the city. As I noted in Chapter 5, this has spawned a wide and rich literature on the operations of companies such as Uber and Lyft, particularly their societal impacts and helped identify the levers for public policy to shape their operations towards greater societal benefit. Going forward, it would be even more valuable if such data could also contain additional information. Currently, it only tracks at a granular level the trips that riders take, vehicle and driver information is much more coarse and aggregated at the monthly level. This prevents an understanding of important questions regarding how many drivers were active in any given period supplying their labor, and their patterns of driving (particularly the share of deadheading). Insights into factors such as elasticity of demand are also obstructed because we do not have any visibility on which trips were rejected by drivers, canceled by customers, or where customers might have declined to order a ride based on the price they saw in the app. Waiting times for customers are also not available, which complicates calibration for the type of model that I presented in Chapter 4.

The CFS dataset while valuable, requires extrapolation to the general population as it is a survey of responding establishments. It also does not offer detailed data on truck drivers or vehicles. Private datasets certainly exist for long haul trucking, which track driver routes and operations at a granular level. Enabling access to such data for researchers for academic purposes would be critical to better model and understand how emerging technologies might change the nature of long haul freight.

6.3 Future work

The previous chapters have identified questions that warrant further research. There are a number of research projects that are direct extensions to the studies presented here.

6.3.1 Electrification of short haul trucking

If automation of highway legs becomes possible, freight routed through trucks will involve multiple vehicles, both automated and human-driven, as described in the transfer-hub model in Chapter 3. Short haul journeys to and from distribution centers and highways could be electrified with existing battery technology. Given that the majority of people live in urban areas, what might be the health benefits of reducing exposure to local air pollutants from diesel trucks? How do these benefits vary regionally and how do they compare to the costs of electrification?

6.3.2 Ride-pooling and external costs

Including ride-pooling will enable greater generalization of the results on external costs presented in Chapter 5 and will also allow the study of other research questions. As highlighted previously, the model presented in Chapter 4 can be extended to different cities with publicly available data on ridesourcing trips. It would be interesting to compare the value of ride-pooling and therefore, the role for policy intervention in reducing externalities across different cities, given regionally unique mobility patterns and road networks.

6.3.3 Flexible charging from TNC vehicles

Full electrification of ridesourcing will involve non-trivial levels of additional load for the electricity grid. Is there a role for public policy to incentivize certain charging patterns that minimize GHG emissions? How should TNC charging demand factor into infrastructure planning by utilities? The level of spatial and temporal detail incorporated in the ABM can be useful for companies building fast-chargers as well as utilities, to plan distribution level infrastructure to manage peak loads and appropriately design peak demand tariffs. Incorporating detailed power demand results from the ABM presented here into capacity expansion models can be a fruitful avenue for future work.

APPENDIX A

Chapter 2 Manuscript

nature energy

ARTICLES

https://doi.org/10.1038/s41560-020-0644-3

Check for updates

Trade-offs between automation and light vehicle electrification

Aniruddh Mohan[1], Shashank Sripad[2,3], Parth Vaishnav[1,3] and Venkatasubramanian Viswanathan[2,3] ✉

Weight, computing load, sensor load and possibly higher drag may increase the energy use of automated electric vehicles relative to human-driven electric vehicles, although this increase may be offset by smoother driving. Here, we use a vehicle dynamics model to evaluate the trade-off between automation and electric vehicle range and battery longevity. We find that automation will likely reduce electric vehicle range by 5–10% for suburban driving and by 10–15% for city driving. The effect on range is strongly influenced by sensor drag for suburban driving and computing loads for city driving. The impact of automation on battery longevity is negligible. While some commentators have suggested that the power and energy requirements of automation mean that the first automated vehicles will be gas–electric hybrids, our results suggest that this need not be the case if automakers can implement energy-efficient computing and aerodynamic sensor stacks.

Electric vehicles (EVs) form an increasing share of new vehicle sales around the world. Several countries are seeking to phase out vehicles with internal combustion engines. Simultaneously, automated vehicles (AVs) are being tested on public roads. Automation could reduce vehicle energy use through smoother driving, platooning, shared mobility and optimal routing[1–3]. There is limited literature on the trade-offs between automation and electrification[4]. Early automated electric vehicles (AEVs) may be heavier, need extra computing and sensor power, and (due to the possible need for protruding sensors) be less aerodynamic than EVs. It has been suggested that if these qualities substantially reduce driving range, the first AVs will be gas–electric hybrids[5], which could slow the electrification of the light vehicle fleet.

In this paper we compare the vehicle-level energy use, range and battery life of a vehicle equipped to attain Society of Automotive Engineers (SAE) levels 4–5 automation to human-driven EVs, by undertaking a careful consideration of the effect on vehicle-level energy use of the different components needed for automated driving, as well as the potential increase in drag from LiDAR (light detection and ranging). We find that automation is likely to reduce EV range, and that the aerodynamic impact of LiDAR is a notable driver of this reduction with suburban driving, while computing loads have a greater effect for city driving. Across different EVs, we find a median reduction in the range of 3–5% for a suburban drive profile and no drag impacts from LiDAR. Including LiDAR increases the median range reduction to between 8–12%. A reduction in range will lead to more frequent charging, speeding up battery degradation, but we find that this impact on battery longevity is negligible.

AEV energy consumption

Sripad and Viswanathan[6] developed a physics-based vehicle dynamics model to estimate the energy demands of an EV given a realistic driving profile. They determined the battery size needed for a given vehicle range or equivalently, the range of an EV given the battery size. Using a realistic velocity profile with a 1 s temporal resolution, the model calculates the instantaneous power needed each second to overcome vehicle inertia, aerodynamic drag and road friction. We extend this model for AEVs by adding the weight of the different components to the mass of the vehicle and battery pack, increasing the drag coefficient for automated solutions with a roof-based spinning LiDAR. If no LiDAR is used, or if solid-state LiDAR that is incorporated into the aerodynamic profile of the vehicle is used, the increase in drag is zero. We also modify the velocity profile to account for potentially smoother driving and add the computing and sensor loads at each second. Keeping track of the total energy used, we repeat the driving profile until the battery is fully depleted. This gives us an estimate of the AEV range for a given battery capacity. We then compare this AEV range to that of the EV to understand how automation affects vehicle range. We run this simulation for two types of velocity profile: the California Unified Cycle Driving Schedule, which is a composite profile that is a city–highway mix, and the Urban Dynamometer Driving Schedule, which is a city-only profile (Supplementary Fig. 1). Full details of the physics model are provided in the Methods.

There are several different combinations of sensor hardware that are currently being tested on vehicles that aim to achieve full automation. For example, some developers are using solutions that include LiDAR, while others are relying solely on cameras and radar. There are also differences in the choice of vendor for LiDAR or radar, and in the number of sensors. Given the numerous possible combinations, we assume a uniform distribution for the sensor and connectivity load. We bound this between 30 W and 150 W. The lower case represents a low-powered LiDAR solution such as the 15 W Ouster OS1 system[7] along with two Bosch mid-range radars (MRRs) of 4.5 W each[8], three Point Grey Dragonfly cameras of 1.5 W each[9] and 1.5 W for connectivity. Some industry developers have also suggested that a sensor package without LiDAR is sufficient for high-level automation. As such, the lower case could also represent an optical-only system with no LiDAR and nine cameras (1.5 W each), two MRRs of 4.5 W each and the remaining 7.5 W for communication and connectivity. The upper estimate

[1]Department of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA, USA. [2]Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh, PA, USA. [3]Wilton E. Scott Institute for Energy Innovation, Carnegie Mellon University, Pittsburgh, PA, USA. ✉e-mail: venkvis@cmu.edu

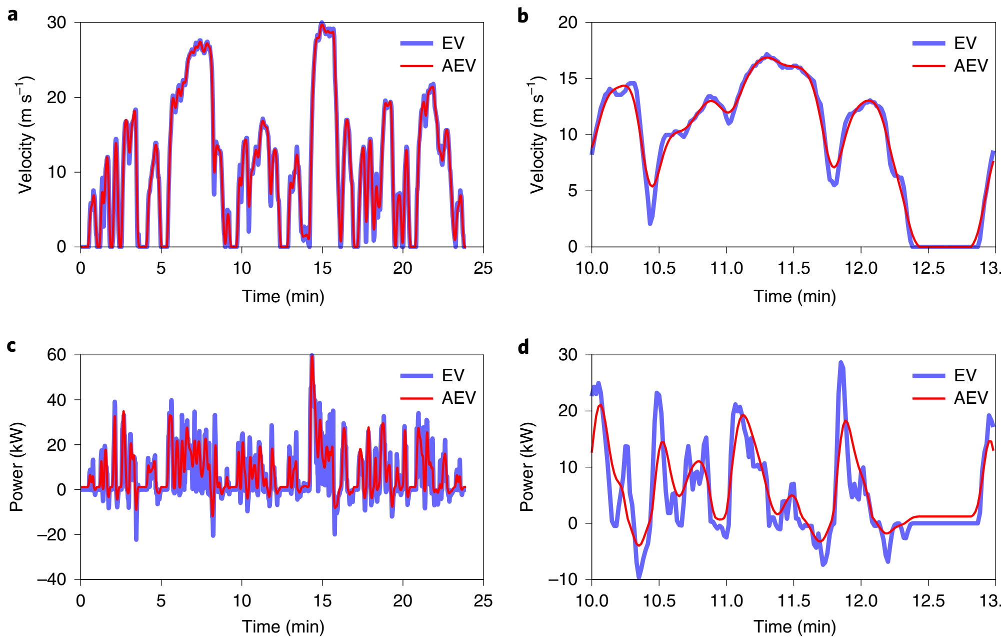

Fig. 1 | EV and AEV velocity and power profiles. a, The composite drive cycle for an EV versus an AEV with 10% energy savings. **b**, Detail of the smoothed drive cycle from minutes 10 to 13 showing the effect of smoothing on braking and acceleration. **c**, The corresponding power profiles for a Tesla Model 3 EV and AEV, assuming a 1,000 W computing load, 150 W sensor load, 25% increase in drag from LiDAR and 10% energy savings from smoother driving. **d**, Detail of the power profile from minutes 10 to 13.

of 150 W represents a system with two Velodyne 64 LiDAR domes of 60 W each[10], along with two Bosch MRRs (9 W)[8] with six Point Grey Dragonfly cameras (9 W)[9] and expanded communications and connectivity (12 W). We were unable to find reliable estimates of the power draw of communication systems, perhaps due to the nascent nature of the technology. We assume a power draw of 1.5 W for our low estimate and 12 W for our high estimate. Modern smartphones consume less than 1 W of power for streaming, connectivity, communication and other functions[11]. Many EVs already have the Global Positioning System (GPS) and other connectivity installed; so it is not clear that the communications load will be additional to existing loads. The Institute of Electrical and Electronics Engineers (IEEE) standard for dedicated short-range communication (DSRC) limits the equivalent isotropically radiated power to less than 2 W (ref. [12]), which precludes a large power draw for the transmitter.

Estimates of computing load for automated driving in the literature show power requirements from a few hundred watts[13] to several thousand watts[14,15]. The Nvidia Pegasus system has been advertised as capable of level 5 autonomous driving and has a power load of 500 W (ref. [16]), while the recently released Tesla Full Self Driving (FSD) claims to require only 150 W (ref. [17]). We therefore bound our estimates between 150 W to 1,000 W with the upper bound representing the higher estimates from the literature. Higher values of computing load are possible given that the technology is nascent, regulators might require redundant systems and cooling requirements for computing may have been underestimated. On the other hand, over the long term, improvements in chip design and computing efficiency should see power loads fall substantially. We also assume a linear relationship between the computing and sensor loads between their respective bounds as increased data flow from the sensors to the computing platform will require a concomitant increase in computing capacity and therefore power draw.

We are aware of no publicly available, empirical estimates of the effect of roof-based LiDAR on the vehicle drag coefficient[1]; so we approximate this effect by using data from wind tunnel tests of the drag impacts of roof add-ons such as police sirens, signs and racks[18,19]. We therefore estimate a lower bound value of 15% increase in drag (which corresponds to the drag increase from a taxi sign) to an upper bound of 40% (which would impose the same aerodynamic penalty as a barrel). In the case of solid-state LiDAR or AVs operating with cameras only, there would be no increase in drag. We therefore consider a second separate case: if there is no LiDAR, or a solid-state LiDAR system, we model no increase in drag. In the longer term, it is also likely that developers will incorporate LiDAR into the vehicle in a way that does not result in additional drag, as consumers may not want to purchase vehicles that have prominent external sensors, as is the case with the AVs being tested today.

Finally, to simulate AEV drive cycles that are smoother than human drivers, and the associated energy savings, we apply a smoothing spline function similar to that of Liu et al. (ref. [20]) to the composite and city drive profiles. The smoothing function can be adjusted to yield different levels of energy savings. We bound this between 5% to 25%, in line with values from the literature that have estimated the energy savings from the smoother driving of AVs (refs. [2,21–23]). Details of the smoothing spline function are provided in the Methods. Our method of smoothing makes no assumptions about the vehicle drive-train. We smooth the velocity profile and calculate the energy savings from first principles, using our physics-based model. Figure 1 shows the original versus smoothed velocity profile for the composite drive cycle for an illustrative case of 10% energy savings from smoother driving, along with the corresponding power profiles for a particular configuration of automation. The effect of smoothing for both drive cycles for different levels of energy savings is shown in Supplementary Figs. 2–4. We also find possible trade-offs between the energy savings from smoother driving and safety, particularly at higher levels of energy savings wherein the vehicle may be at rest in the original profile but has a non-zero velocity in the smoothed profile. However, we find almost no change in our estimates of AEV range when we constrain our smoothed drive cycles so that the velocity is zero at all instances where the velocity in the original profile is zero. For a detailed

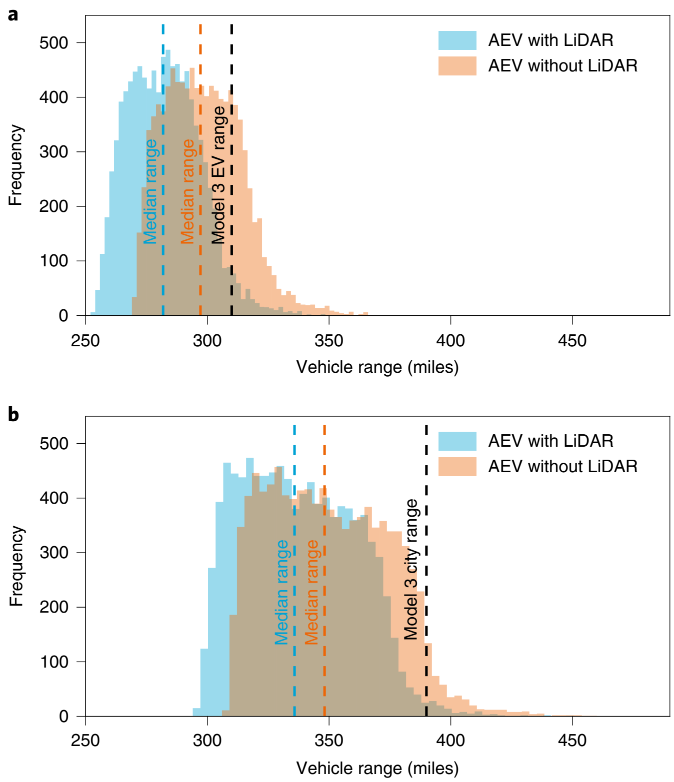

Fig. 2 | Histogram of range results for composite and city drive profiles for a Tesla Model 3 with an 80 kW h battery pack. a, Results of simulating the vehicle range of an automated Model 3 given the composite driving profile and input parameters for automation with and without LiDAR. The median loss in range is 9% with LiDAR and 4% without. In 90% of our MC simulations, the range impact is between −16% and −1% with LiDAR and between -11% and 5% without LiDAR. This is compared to the Tesla Model 3 EV-only range. **b**, Results of simulating the vehicle range of an automated Model 3 given the city driving profile and input parameters, for automation with and without LiDAR. The median loss in range is 14% with LiDAR and 11% without. In 90% of our MC simulations, the range impact is between −22% and −4% with LiDAR and between -19% and 0% without LiDAR. This is compared to the Tesla Model 3 EV-only range for city driving.

discussion of this, see Supplementary Notes 1 and 2, Supplementary Figs. 5–8 and Supplementary Table 1.

Effect of automation on EV range

We use vehicle design parameters based on a Tesla Model 3 with 310 miles of range and an 80 kW h battery pack as our base EV (ref. [24]). We use a combination of scenario analysis and Monte Carlo (MC) simulations to understand how the deep uncertainty in our input parameters affects our estimates of reduction in range. We consider two broad deployment scenarios for automation: with or without LiDAR. We treat these scenarios separately for two reasons. One, there is considerable disagreement among the manufacturers as to whether LiDAR is essential[25], with developers deploying both LiDAR-based and LiDAR-free technologies. Two, LiDAR can impose a considerable energy and aerodynamic penalty; eliminating it would have a large effect on energy use. Within each scenario, we use MC analysis, since there are too many plausible combinations of other parameters to reasonably justify one scenario over another in the current early stages of AV development. MC simulations allow us to explore the entire parameter space[26].

We find that adding automation with LiDAR would result in reduced range or require a larger battery in more than 95% of simulations (Fig. 2a). This indicates that while there is uncertainty around technology choices and the power loads of automation-enabling equipment, LiDAR-based solutions that have a drag penalty will almost certainly reduce range compared to human-driven EVs. Even if we assume no power consumption from LiDAR and no change in drag, energy use increases in 76% of simulations for the composite drive profile, relative to a human-driven EV. We find that even with the computing load and sensor load each at their lower bounds and no increase in drag, AEVs will need energy savings from smoother driving of at least 4–5% over a human driver to have the same range as EVs. If savings from smoother driving are higher, this will lead to an increase in EV range. In 90% of our MC simulations for the composite drive cycle, the reduction in range is between 1% and 16% for an automated Tesla Model 3 with LiDAR compared to an EV with the same battery. The median estimate is a 9% loss in range. Next, we apply the same analysis to estimate the range impacts for other EV models, such as the Tesla Model S, Nissan Leaf, Chevy Bolt and Xpeng G3. The results for the composite drive cycle for all EV models considered are summarized in Fig. 3a.

In 90% of our MC simulations, the range penalty for the city drive cycle is between 4% and 22% for a Model 3 AEV with LiDAR (Fig. 2b). The median estimate is a 14% loss in range. Results for different EV models are shown in Fig. 3b.

Table 1 shows the energy demand and range impacts for the different autonomous configurations considered. We see that the drag impacts of spinning LiDAR are more keenly felt at higher speeds on the highway. Keeping all other autonomy parameters fixed, the drag increase from having roof-based spinning LiDAR results in an additional 5% range penalty for the composite profile compared to 3% for the city profile. Our results therefore show that developers may need to consider automated solutions that either eliminate LiDAR or reduce its drag impact. Solid-state LiDAR could reduce drag but would impose an additional computing load. However, this additional computing load is unlikely to be as important as the impact of additional drag, as even an increase of 100 W in the computing load will only decrease the range by 1% (see Supplementary Note 3 and Supplementary Fig. 9). The higher range penalty in the city cycle (Fig. 3b) compared to the composite profile (Fig. 3a) across all the EVs we analysed is due to the greater effect of computing loads (which we assume applies at each second) on range as a result of longer trip times in the city for the same distance travelled. As such, we see in Table 1 that moving to a more efficient compute stack provides a greater benefit in the city than for suburban driving. Developers who wish to deploy robo-taxis in urban environments therefore have strong incentives to bring down computing power needs in order to maximize vehicle utilization. This could be achieved over time as standardized algorithms are built into efficient specialized circuits or if developers design dedicated chips for neural network computation.

Negligible impact of automation on battery longevity

A decrease in range could lead to more frequent charging or longer charging times, which leads to faster battery degradation. We use the approach developed by Sripad and Viswanathan[6] to model the longevity of the battery for each type of vehicle, that is, to estimate the number of total miles for which the vehicle can be driven until the battery is unable to charge to more than 80% of its original capacity. To realistically assess battery degradation, we model a series of 24 h periods, in each of which the vehicle drives for 50 miles for the composite profile (or 30 miles for the city profile), charges until the battery is full and then rests until it is driven again the following day. For the case of the composite profile and the base AEV with LiDAR, the loss in battery longevity for the median reduction in range is 5% or 5,500 miles, and the fifth percentile case with a range reduction of 16% would result in a loss of 9% or 10,000 miles, as shown

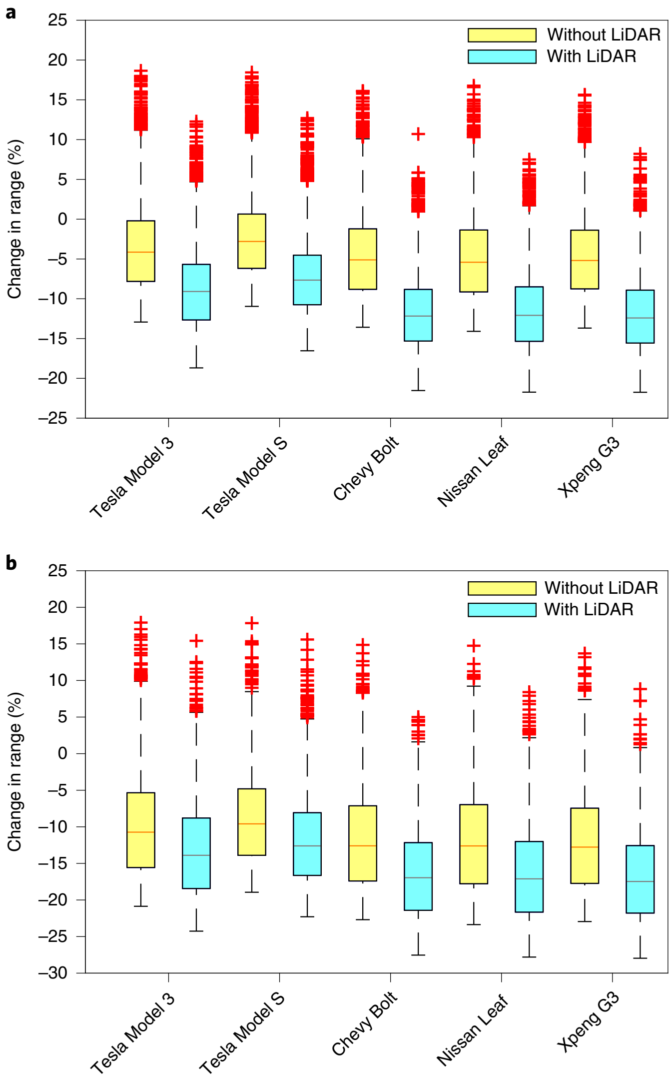

Fig. 3 | Results of the MC simulation of AEV range for different EV models. The *y* axis shows the percentage change in range for the AEV compared to the base EV, $(\text{Range}_{AEV} - \text{Range}_{EV})/\text{Range}_{EV}$. The horizontal red line is the median change in range, and the red data points are the data that are beyond the whisker where the maximum whisker length is 1.5 times the interquartile range. The *x* axis lists the five EVs considered in the analysis. **a**, Box plot showing the range impact for the composite drive profile and for automated solutions with and without LiDAR. **b**, Box plot showing the range impacts for the city drive profile and for automated solutions with and without LiDAR.

Table 1 | Energy demand and range impacts for different autonomous configurations

Drive cycle	Technology	Range impact	W h mile^{-1}
Composite	150 W compute + no LiDAR	6%	244
	500 W compute + no LiDAR	0%	259
	500 W compute + LiDAR	-5%	272
City	150 W compute + no LiDAR	5%	195
	500 W compute + no LiDAR	−4%	214
	500 W compute + LiDAR	−7%	221

All scenarios assume 10% energy savings from smoother driving and a sensor load of 30 W. Increase in drag from LiDAR is fixed at 25%.

in Fig. 4a. Given that the longevity of state-of-the-art batteries is constantly improving, this effect is likely to be more negligible in the future[27]. The battery degradation simulations were performed using a battery-pack model where the cells are simulated using a pseudo two-dimensional electrochemical model[28] that contains a degradation sub-model[29,30]. The details of the degradation model are discussed in the Methods.

The 95th percentile AEV shows a small (1%) increase in battery longevity compared to the EV due to the effect of smoother driving, which lowers the discharge rate of the battery as well as the recharging rate through regenerative braking. Moreover, due to the higher energy supplied by regenerative braking segments for the EV, during each daily round trip, the average state of charge of the battery pack for the EV is about 0.6% higher than that of the 95th percentile AEV. Since a higher state of charge increases the rate of battery degradation[27], the EV battery degrades faster than the 95th percentile AEV, since this state-of-charge effect outweighs the very small increase in the duration of charging time from the 1% loss in range compared to the EV. Battery degradation results for the case without LiDAR are shown in Fig. 4b. Battery degradation results for the base AEV with and without LiDAR using the city velocity profile and a daily drive schedule of 30 miles are shown in Supplementary Fig. 10.

Conclusions

While there is considerable uncertainty with respect to the technologies that will enable fully automated driving, our model provides a way to grapple with this uncertainty and derive useful insights about the vehicle-level energy use of AVs. We find that high-level automation will likely reduce EV range by about 5–10% in the case of the suburban drive profile. Given that the specific energy of lithium-ion batteries is increasing at roughly 5% each year without a corresponding increase in cost or a step change in technology, automation will have the same effect on range as a 1–2 yr time lag on improvements in battery-specific energy. Will consumers value full automation more than the modest loss in range we estimate? Future work must assess the benefits of automation (for example, using travel time productively, increased safety) against its cost (for example, loss in range, increased capital cost, public perceptions of safety)[31]. Our study provides a basis for more realistic assessments of consumer preferences, since we estimate the change in EV range associated with automation. In conclusion, to minimize the effect of automation on EV range, developers should focus on bringing down power requirements for computing and on incorporating sensors in aerodynamic designs.

Methods

Physics-based model. The total force experienced by a vehicle can be written as follows:

$$F_{\text{total}} = F_{\text{drag}} + F_{\text{friction}} + F_{\text{inertia}} + F_{\text{gradient}} \quad (1)$$

We ignore the gradient term and focus on the inertial, friction and drag forces[6]. The power at any time t can be written as follows:

$$P(t)_{\text{drag}} = \left(\frac{1}{2} \times \rho \times C_d \times A \times v(t)^2\right) \times v(t) \quad (2)$$

$$P(t)_{\text{friction}} = (\mu_{rr} \times \text{mass} \times g) \times v(t) \quad (3)$$

$$P(t)_{\text{inertia}} = \left(\text{mass} \times \frac{dv}{dt}\right) \times v(t) \quad (4)$$

where $v(t)$ is the velocity at time t from the drive cycle data, dv/dt is the acceleration, ρ is the density of air, C_d is the drag coefficient of the vehicle, A is the frontal surface area, mass is the weight of the vehicle including the battery pack,

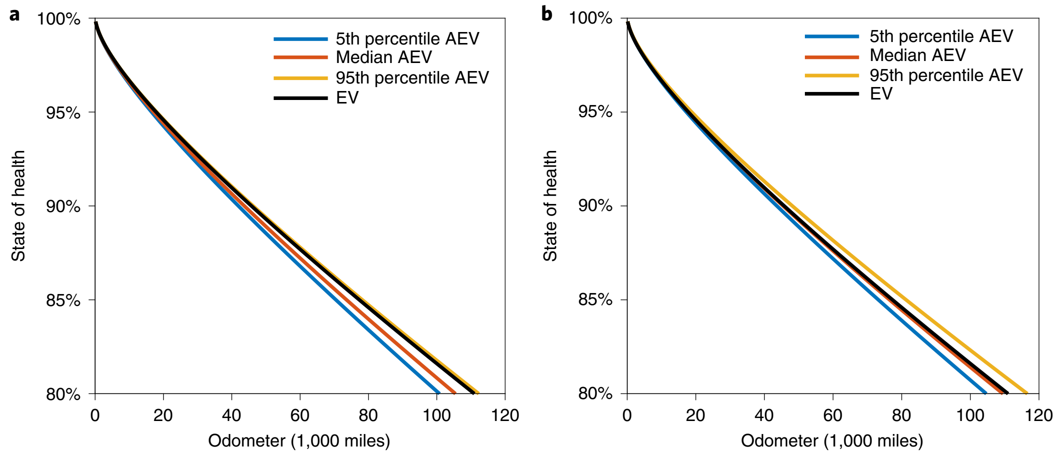

Fig. 4 | Battery degradation results for our base EV and AEV for the composite drive cycle. The simulations were performed for a daily drive schedule of 50 miles. A state of health or a remaining useful capacity of about 80% is considered the end of life of the battery pack. A series of 24 h power profiles that include 50 miles of daily driving was used as an input to the battery-pack model. **a**, The base AEV with LiDAR. The median loss in battery longevity with LiDAR is 5,500 miles or under four months of driving. The fifth percentile case with LiDAR leads to a loss of 10,000 miles or over half a year of driving. **b**, The base AEV without LiDAR. The median AEV without LiDAR lasts for approximately 110,000 miles, and the fifth percentile AEV lasts for approximately 105,000 miles. The median loss in battery longevity compared to the EV is 1,300 miles or a month of driving. The fifth percentile case leads to a loss of 6,000 miles or four months of driving compared to the EV.

g is the acceleration due to gravity and μ_{rr} is the coefficient of friction (rolling resistance). Therefore the total power is given as follows:

$$P(t)_{\text{total}} = \frac{P(t)_{\text{drag}} + P(t)_{\text{friction}} + P(t)_{\text{inertia}}}{\eta_1 \times \eta_2} + \frac{P(t)_{\text{compute}}}{\eta_2} + \frac{P(t)_{\text{sensor}}}{\eta_2} \qquad (5)$$

where η_1 and η_2 are the drive-train and battery efficiencies, respectively; P_{compute} is the computing load of the platform, and P_{sensor} is the load of all the sensors and communications on the vehicle.

The power requirements can be integrated over the duration of a trip to estimate the total energy needed for the trip, given a velocity profile. This energy requirement can be compared with the energy capacity of the battery pack, and accordingly, the duration of the drive, t, that is possible for a given battery-pack size can be calculated. This time t can then be used along with the velocity profile to calculate the range of the vehicle by integrating velocity with respect to time.

We first perform these calculations for a human-driven EV, where the mass would be just the mass of the vehicle plus the battery pack and where C_d would be simply the drag coefficient of the EV; velocity would be not smoothed, and P_{compute} and P_{sensor} would both be zero. This gives us the range of the EV.

In the case of the human-driven Tesla Model 3, our model predicts the range to be 309 miles for the city–highway drive profile, compared to the Environmental Protection Agency (EPA)-rated range of 310 miles. For the city drive profile, the model predicts the human-driven Tesla Model 3 range to be 393 miles, compared to the EPA-rated range of 391 miles. The ranges that our model predicts for other EVs are similarly close to their EPA-rated ranges. This suggests that it is appropriate for us to compare our estimates of AEV range to the EPA's estimates of EV range to estimate the reduction in range associated with the introduction of automation.

To model the AEV, we then 'turn on' the automation for the vehicle, by doing the following: adding the weight of the different components to the mass of the vehicle and battery pack, increasing the drag coefficient due to LiDAR (for automated solutions with spinning roof-based LiDAR; otherwise the increase in drag is zero), applying a different velocity profile due to smoothing (see below) and applying the computing and sensor loads at each second. This gives us a new energy requirement and accordingly a new range. We can then compare the new range to the old range to understand the energy impacts of automation on an electric vehicle's range. We also run this simulation for both drive profiles: a city–highway mix, and a city-only profile.

Velocity smoothing. Smoothing of the velocity profiles is accomplished using splines[20] to minimize the objective function:

$$\left[\sum_{i=1}^{n}\left(y_i - g(x)_i\right)^2 + \lambda \int \left(g''(x)\right)^2 \mathrm{d}x\right] \qquad (6)$$

where $g(x)_i$ is the spline function, which yields the velocity of the vehicle at time x_i, and λ is the smoothing parameter. The first term is the mean square error (MSE), with y_i equal to the value of y, the velocity of the vehicle, at the ith data point in the original driving cycle, $i = 1, 2,\ldots, n$; and $g(x_i)$ is the predicted value of the velocity of the vehicle in the smoothed cycle, after a period of time x_i corresponding to the time elapsed at the ith point in the original driving cycle. Minimizing the function ensures that, with the smoothed drive cycle, the vehicle is at approximately the same points at the same times as it was with the original drive cycle. The $g''(x)$ is the second derivative of g with respect to x (that is, the rate of change of acceleration of the vehicle at the ith point in the drive cycle); λ is a smoothing parameter that determines the extent to which we are willing to trade off fidelity to the original drive cycle (the first term) against the objective of reducing energy use (the second term). Therefore, by parameterizing λ, we vary the energy savings that result from the AEV's smoother drive cycle. The level of energy savings considered reasonable is based on estimates from the literature, as outlined in the section on 'AEV energy consumption'. We use the Fit function in MATLAB, which takes an input parameter 'SmoothingParam', to model the spline function. Drive profiles corresponding to different levels of energy savings are shown in Supplementary Figs. 2 and 4. Potential trade-offs between energy savings and safety are discussed in Supplementary Note 2.

Vehicle characteristics. We use vehicle design parameters based on a Tesla Model 3 with 310 miles of range and an 80 kW h battery pack as our base EV. The battery-pack weight was estimated at 553 kg from the model from Sripad and Viswanathan[6]. We used this battery-pack mass and capacity to estimate the range of the AEV using the physics-based model described above. The drag coefficient of the Model 3 is 0.23 and the frontal area is 2.22 m^2 (ref. [32]). The weight of the car excluding the battery pack is 1,200 kg (refs. [24,32]), and we additionally include the weight of a passenger assumed to be 80 kg. The additional weight of the sensor and computing package is assumed to be 20 kg in total. We assume a total battery-to-wheels efficiency (including both the battery and drive-train efficiency) of 0.855 and a coefficient of friction of 0.0071, which is also viewed as a fitting parameter, while benchmarking the EPA range for the vehicle. The same process was then repeated for other EV models.

Battery degradation modelling. The degradation processes are modelled within the battery-packmodel's[28–30,33] degradation sub-model[33], shown as follows:

$$j_{\text{SEI}} = -S \times F \times k_{0,\text{SEI}} \times c_{\text{solvent}}^{\text{surface}} \times \exp\left[-\frac{\alpha_{\text{c,SEI}} \times F}{R \times T} \times \left(\phi_s - \phi_l - I \times R_{\text{film}} - U_{\text{SEI}}\right)\right] \qquad (7)$$

$$j_{\text{PL}} = -S \times i_{0,\text{PL}} \times \exp\left[-\frac{\alpha_{\text{c,PL}} \times F}{R \times T} \times \left(\phi_s - \phi_l - I \times R_{\text{film}}\right)\right] \qquad (8)$$

$$\frac{\mathrm{d}\epsilon_{\text{AM}}}{\mathrm{d}t} = -k_{\text{AMI}} \times I \qquad (9)$$

where j_{SEI} is the side reaction current for the solid–electrolyte interphase (SEI) layer formation, j_{PL} is the side reaction current for lithium plating and equation (9) captures rate of active material isolation (AMI) based on the rate of change of active material porosity (ϵ_{AM}). The total current is given by I, F is Faraday's constant and S is the surface area of the electrode particle. The other constants from the degradation sub-model are the rate constants ($k_{0,SEI} = 1\times10^{-12}\,m\,s^{-1}$ (ref. [33]) and $k_{AMI} = 2\times10^{-14}\,m\,s^{-1}$) and the exchange current density ($i_{0,PL} = 0.001\,A\,m^{-2}$)[33]. The α symbols are the cathodic transfer coefficients, and $c_{solvent}^{surface}$ is the concentration of the solvent on the surface of the anode. The φ symbols are the potentials of the solid electrode (φ_s) and liquid electrolyte (φ_l). R_{film} is the resistance of the SEI layer. U_{SEI} is the equilibrium potential at which the SEI layer forms. The cells are assumed to have an $Ni_{0.80}Co_{0.15}Al_{0.05}$ cathode and a graphite anode. Ambient temperature is set at 298 K. We model a 24 h period that includes a drive of 50 miles for the composite profile, or 30 miles for the city profile, followed by charging, and then the appropriate rest segment. We assumed 30 miles for daily travel in the city profile on the basis of the American Driving Survey (2013–14)[34]. The Federal Highway Administration's highway statistics show a ratio of approximately 0.7 for urban miles to urban and rural total miles driven[35]. We therefore picked a round number of 50 miles for our city–highway profile. Incidentally, varying the daily miles driven has no effect on our results—for instance, we ran a case for a daily drive schedule of 100 miles for the composite profile, and the difference in battery longevity between the EV and AEV was similar to that of the 50 mile case. The simulations were performed using AutoLion-ST v6.3, Build 2.

We derive the power profiles for the drive from running the physics-based model described above. Due to regenerative braking, the battery charges when the vehicle brakes. The rate at which this charging occurs is masked by the fact that the battery is also discharging at these moments, in order to provide power for computing and the sensors. We account for both the charging and discharging phenomena in the following way: we transfer the computing and sensor loads for the seconds where the vehicle is braking to the nearest second where the vehicle is stationary, travelling at constant speed or accelerating. This allows us to account for the full effect of battery degradation that occurs due to regenerative braking. While this artificially increases the computing and sensor loads and therefore the discharge rate of the battery during braking-adjacent moments in the drive cycle, the effect of this on battery longevity is negligible given the low computing and sensor loads compared to the power requirements for vehicle acceleration, and given that the rate of charging is likely to have a greater effect on battery longevity than small variations in the rate of discharge. The degradation model is implemented within a full-pack battery electrochemical–thermal model, which is described elsewhere[29,30,28]. Note that the adjustment described above is made only to model battery degradation. We model vehicle range without making these modifications to the power profile.

Data availability

All underlying data are publicly available at https://github.com/battmodels/Automation-EV-Range. Source data are provided with this paper.

Code availability

The custom code for the MC simulation of the physics-based model presented in the paper is publicly available on GitHub at https://github.com/battmodels/Automation-EV-Range.

To allow readers to engage with our research, we have created a web applet (available at https://tinyurl.com/avrange) that allows users to select different combinations of radar, computing and LiDAR load, cameras, energy savings and EVs, to assess the effect of different assumptions about automation on vehicle range. The input files for the battery degradation simulations on AutoLion-ST v6.3, Build 2 will be provided on request. Source data are provided with this paper.

Received: 7 August 2019; Accepted: 1 June 2020;
Published online: 29 June 2020

References

1. Taiebat, M., Brown, A. L., Safford, H. R., Qu, S. & Xu, M. A review on energy, environmental, and sustainability implications of connected and automated vehicles. *Environ. Sci. Technol.* **52**, 11449–11465 (2018).
2. Stephens, T. et al. *Estimated Bounds and Important Factors for Fuel Use and Consumer Costs of Connected and Automated Vehicles* (National Renewable Energy Laboratory, 2016).
3. *Study of the Potential Energy Consumption Impacts of Connected and Automated Vehicles* (US Energy Information Administration, US Department of Energy, 2017).
4. Offer, G. J. Automated vehicles and electrification of transport. *Environ. Sci. Technol.* **8**, 26–30 (2015).
5. Hawkins, A. J. Not all of our self-driving cars will be electrically powered—here's why. *The Verge* https://www.theverge.com/2017/12/12/16748024/self-driving-electric-hybrid-ev-av-gm-ford/ (2017).
6. Sripad, S. & Viswanathan, V. Evaluation of current, future, and beyond Li-ion batteries for the electrification of light commercial vehicles: challenges and opportunities. *J. Electrochem. Soc.* **164**, E3635–E3646 (2017).
7. Ouster. *OS1 Datasheet* http://data.ouster.io/downloads/OS1-lidar-sensor-datasheet.pdf (2019).
8. *Chassis Systems Control: Mid-range Radar Sensor (MRR) for Front and Rear Applications* (Robert Bosch, 2015).
9. Point Grey Dragonfly2 0.8 MP Mono FireWire 1394a Board Level (Sony ICX204) (Point Grey, 2018).
10. Velodyne HDL-64E (Velodyne, 2019).
11. Carroll, A. et al. An analysis of power consumption in a smartphone. In *Proceedings of the 2010 USENIX Conference on USENIX Annual Technical Conference* Vol. 14, 21–21 (USENIX Association, 2010).
12. Sepulcre, M., Gozalvez, J. & Hernandez, J. Cooperative vehicle-to-vehicle active safety testingunder challenging conditions. *Transp. Res. Part C Emerg. Technol.* **26**, 233–255 (2013).
13. Gawron, J. H., Keoleian, G. A., de Kleine, R. D., Wallington, T. J. & Kim, H. C. Life cycle assessment of connected and automated vehicles: sensing and computing subsystem and vehicle level effects. *Environ. Sci. Technol.* **52**, 3249–3256 (2018).
14. Liu, S., Tang, J., Zhang, Z. & Gaudiot, J.-L. Computer architecture for autonomous driving. *Computer* **50**, 18–25 (2017).
15. Lin, S.-C. et al. The architectural implications of autonomous driving: constraints and acceleration. In *Proceedings of the Twenty-Third International Conference on Architectural Support for Programming Languages and Operating Systems* 751–766 (ACM, 2018).
16. Nvidia Drive AGX Pegasus (Nvidia, 2018).
17. Walz, E. Tesla shares details of its new self-driving chipset at 'Autonomy Investor Day'. *FutureCar* https://www.futurecar.com/3153/Tesla-Shares-Details-of-its-New-Self-Driving-Chipset-at-Autonomy-Investo (2019).
18. Chowdhury, H., Alam, F., Khan, I., Djamovski, V. & Watkins, S. Impact of vehicle add-ons on energy consumption and greenhouse gas emissions. *Procedia Eng.* **49**, 294–302 (2012).
19. Chen, Y. & Meier, A. Fuel consumption impacts of auto roof racks. *Energy Policy* **92**, 325–333 (2016).
20. Liu, J., Kockelman, K. M. & Nichols, A. Anticipating the emissions impacts of smoother driving by connected and autonomous vehicles, using the MOVES model. In *Transportation Research Board 96th Annual Meeting* (2016).
21. Prakash, N., Cimini, G., Stefanopoulou, A. G. & Brusstar, M. J. Assessing fuel economy from automated driving: influence of preview and velocity constraints. In *ASME 2016 Dynamic Systems and Control Conference* (American Society of Mechanical Engineers, 2016).
22. Mersky, A. C. & Samaras, C. Fuel economy testing of autonomous vehicles. *Transp. Res. Part C Emerg. Technol.* **65**, 31–48 (2016).
23. Gao, Z. et al. Evaluation of electric vehicle component performance over eco-driving cycles. *Energy* **172**, 823–839 (2019).
24. *Certificate Summary Information Report*, Tesla Motors: Model 3 Long Range, 02/28/2019 (United States Environmental Protection Agency, 2019).
25. Templeton, B. Elon Musk's war on LIDAR: who is right and why do they think that? *Forbes* https://www.forbes.com/sites/bradtempleton/2019/05/06/elon-musks-war-on-lidar-who-is-right-and-why-do-they-think-that/7fe42c4f2a3b (2019).
26. Morgan, M. G., Henrion, M. & Small, M. *Uncertainty: a Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis* (Cambridge University Press, 1992).
27. Harlow, J. E. et al. A wide range of testing results on an excellent lithium-ion cell chemistry to be used as benchmarks for new battery technologies. *J. Electrochem. Soc.* **166**, A3031–A3044 (2019).
28. Srinivasan, V. & Wang, C.-Y. Analysis of electrochemical and thermal behavior of Li-ion cells. *J. Electrochem. Soc.* **150**, A98–A106 (2003).
29. Fang, W., Kwon, O. J. & Wang, C.-Y. Electrochemical–thermal modeling of automotive Li ion batteries and experimental validation using a three-electrode cell. *Int. J. Energy Res.* **34**, 107–115 (2010).
30. Smith, K. & Wang, C.-Y. Power and thermal characterization of a lithium-ion battery pack for hybrid-electric vehicles. *J. Power Sources* **160**, 662–673 (2006).
31. Anderson, J. M. et al. *Autonomous Vehicle Technology: a Guide for Policymakers* (Rand, 2014).
32. Tesla Model 3. *Tesla* https://www.tesla.com/model3 (2019).
33. Yang, X.-G., Leng, Y., Zhang, G., Ge, S. & Wang, C.-Y. Modeling of lithium plating induced aging of lithium-ion batteries: transition from linear to nonlinear aging. *J. Power Sources* **360**, 28–40 (2017).
34. AAA Foundation for Traffic Safety. American driving survey: methodology and year 1 results, May 2013–May 2014. *AAA* https://newsroom.aaa.com/wp-content/uploads/2015/04/FACT_SHEET_American_Driving_Survey_Methodology_and_year_1_results_May_2013_to_May_2014.pdf (2015).
35. Annual vehicle distance traveled in miles and related data – 2017(1). *Federal Highway Administration* https://www.fhwa.dot.gov/policyinformation/statistics/2017/pdf/vm1.pdf (2017).

Acknowledgements

This work was supported by the Carnegie Mellon University (CMU) College of Engineering, Department of Engineering and Public Policy; Scott Institute for Energy

Innovation; Center for Climate and Energy Decision Making (SES-1463492; through a cooperative agreement between the National Science Foundation and CMU); the Block Center for Technology and Society at CMU; Technologies for Safe and Efficient Transportation University Transportation Center; and Mobility21, A United States Department of Transportation National University Transportation Center. A.M. thanks A. Dongare and A. Bhat at CMU for their helpful comments on the power demands of AV computing platforms.

Author contributions

A.M., S.S., P.V. and V.V. designed the research and conceived the paper; A.M., S.S., P.V. and V.V. developed the physics-based model for vehicle energy use; S.S. and V.V. developed the battery degradation model; A.M., S.S., V.V. and P.V. performed the analysis and created the figures; and A.M., S.S., P.V. and V.V. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41560-020-0644-3.

Correspondence and requests for materials should be addressed to V.V.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

APPENDIX B

Chapter 3 Manuscript

Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-022-01103-w OPEN

Check for updates

Impact of automation on long haul trucking operator-hours in the United States

Aniruddh Mohan[1] & Parth Vaishnav[1,2]✉

Automated long haul trucking is being developed for commercial deployment in the United States. One possible mode of deployment for this technology is a "transfer-hub" model where the operationally less complex highway driving is automated, while human drivers drive the more complex urban segment of the route. We study the possible net impacts on tractor-trailer operator-hours from this mode of deployment. Using data from the 2017 Commodity Flow Survey, we gather information on trucking shipments and the operator-hours required to fulfill those shipments. We find that up to 94% of long haul trucking operator-hours may be impacted as the technology improves to operate in all weather conditions. If the technology is however restricted to the southern states where the majority of companies are currently testing automated trucking, we find that only 10% of operator-hours are impacted. We conduct interviews with industry stakeholders including tractor-trailer operators on the feasibility of such a system of deployment. We find that an increase in short haul operation is unlikely to compensate for the loss in long haul operator-hours, despite public claims to this effect by the developers of the technology. Policymakers should consider the impact of different scenarios of deployment on the long haul trucking workforce.

[1] Department of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA, USA. [2] School for Environment and Sustainability, University of Michigan, Ann Arbor, MI, USA. ✉email: parthtv@umich.edu

Introduction

Automated driving technology is currently being tested on public roads in the United States in both the light vehicle and heavy duty segments. Given the likely reduced operational complexity involved in highway driving, several companies are currently working on developing automation for long haul trucking, which is designed to work as per a "transfer-hub" model (Hickman et al., 2018; Transport Topics, 2018). This would involve an automated truck (AT) completing the highway leg of the route and human drivers undertaking the more complex suburban-urban segments at both the starting and end points of the journey. Truck ports near highways would be used to switch out the trailer from the prime mover and enable this switch at both ends. For a schematic of the possible transfer-hub model see Fig. 1.

The promise of ATs has led to widespread concern about job losses in long haul trucking, which is a common profession in the United States, particularly for men with high-school educations (Wertheim, 2020). On the other hand, it has also been noted, often by the companies developing this technology, that long haul trucking currently faces a labor shortage and automation will create new short haul jobs, which will more than make up for the long haul jobs lost. As a result of these conflicting claims, as well as the uncertainty over the technology itself and its limitations, there is little clarity on how automated trucking will be deployed and its economic and political ramifications, such as the impact on the long haul trucking labor market.

We use data from the Commodity Flow Survey (CFS) 2017 (United States Census Bureau, 2020), which is a dataset jointly produced by the U.S. Census Bureau, U.S. Bureau of Transportation Statistics, and the U.S. Department of Commerce. CFS produces a sample of shipments in the United States including data on the type, origin and destination, value, weight, modes of transportation, and distance shipped. We estimate the operator-hours required for different routes using origin and destination information for trucking shipments from the CFS dataset, which enables us to estimate highway and (sub)urban splits for each shipment. We consider the technological constraints of automated trucking to make more refined estimates of the possible near term impacts of automation on long haul trucking operator-hours in the United States by assessing different scenarios of deployment. We find that contrary to strong claims by companies developing this technology (Gilroy, 2019), the loss in long haul operator-hours is unlikely to be compensated for by an increase in demand for short haul drivers. We find large labor impacts from the deployment of the transfer-hub model, contingent on the pace of progress of automation technology's ability to function in different operating conditions. We compare the capabilities of automation with the tasks truck drivers are required to perform through discussions with industry stakeholders, including long haul operators. This allows us to have a realistic assessment of the tasks the technology will be required to perform over the highway legs and explore the feasibility of such deployment scenarios. In this framework, where jobs are considered as a bundle of tasks, our work builds on an extensive literature in the social sciences (Acemoglu and Autor, 2011; Acemoglu and Restrepo, 2019; Arntz et al., 2016). Through our limited sample of interviews, we discover reluctance to shift to new modes of operation such as short haul driving that will result from any application of the transfer-hub mode of deployment. The rest of the paper is structured as follows. Section "Literature review" reviews the existing literature on job impacts of AT. Section "Methods" describes our methods for both the analysis of operator-hour impacts and semi-structured interviews. Section "Results" presents the results and section "Conclusion and policy implications" concludes.

Literature review

Given the relatively recent advances in automated driving technology, there are a limited number of studies that have been conducted on the job impacts of automation on truck driver jobs. Viscelli (2018) has analyzed the impact of automation based on a similar "transfer-hub" operational model assumption shown in Fig. 1, where highway driving is automated while human drivers perform the urban segments of the route. Using revenue data from major trucking carriers and estimates of average per driver revenue, Viscelli (2018) estimates nearly 300,000 long haul jobs to be at risk. As per the analysis, these at-risk jobs are primarily concentrated in dry van and refrigerated trucking, which are characterized by high turnover and low wages (Viscelli, 2018).

Similarly, Gittleman and Monaco (2020) have undertaken a study analyzing the potential job losses from automation. They also find that contrary to several estimates in the media, the upper bound on job losses from automation are likely to be 400,000 jobs, far less than popular estimates of 1–2 million jobs at risk. Gittleman and Monaco (2020) perform this analysis using 2002 Vehicle Inventory and Use Survey (VIUS) data to gather estimates of the operational use of heavy trucks. They then apportion the number of long haul truck drivers using Occupational Employment Statistics data from the Bureau of Labor Statistics. Their study also notes that only a specific aspect of trucking—long haul highway driving—is ripe for automation and that operators who perform other tasks or are involved in customer

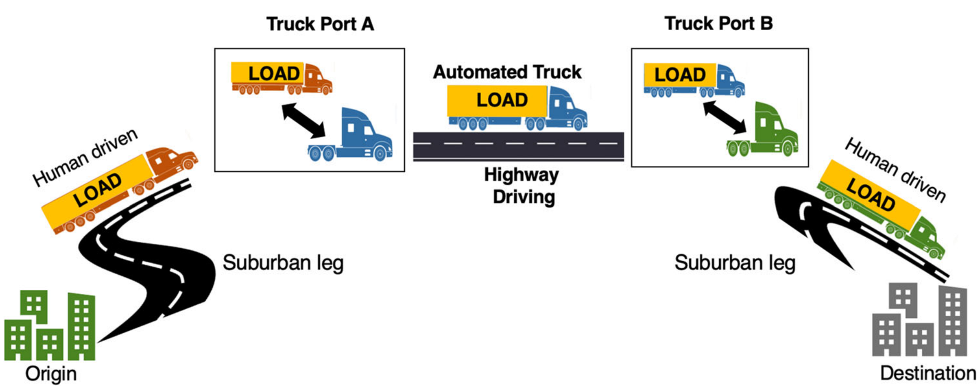

Fig. 1 Schematic showing the possible operation of a transfer-hub model where a human driven truck drives the load from the origin to the first truck port where the load is switched to an automated prime mover shown in black. This completes the highway leg of the journey to another truck port where a human driven truck takes the load from the port to the final destination.

facing roles such as in package delivery services are unlikely to face job losses (Gittleman and Monaco, 2020).

Groshen et al. (2019) consider different scenarios of adoption of automation and analyze the job impacts in different sectors of the economy using simulations and consultations with industry experts. They estimate 60–65% of heavy truck and tractor-trailer truck driving jobs to be eliminated with full implementation of automation (Groshen et al., 2019).

Waschik et al. (2021) leverage a dynamic model of the U.S. economy to simulate the macroeconomic impacts of automated trucking in the economy. They consider three different speeds of adoption: slow, medium and fast and find that employment in for-hire trucking falls by 20–25 percent and in private trucking by 4–5 percent across the scenarios considered (Waschik et al., 2021). Broader benefits in terms of productivity and employment in the economy are found to outweigh the employment impacts on long haul operators and the study assumes that long haul operators will switch to short haul jobs, minimizing the overall effect on the workforce (Waschik et al., 2021).

Finally, future of work studies around the deployment of automated vehicles have also delved into the issue of job losses in trucking from automation. For instance, Leonard et al. (2020) suggest that automation will create new roles such as remote management of trucks, dispatching, and field support while disrupting day driving jobs. However, they do not estimate the number or share of jobs at risk.

We found no examples in the literature of analyses that used trucking routes and shipment data or examined the specific capabilities of the technology (e.g., only operating in favorable weather conditions) to estimate the labor impacts of automated trucking. We also found no studies that undertook interviews with stakeholders including long haul operators to understand their perspectives on the tasks required over the highway legs.

Outside of the specific case of automated trucking, we also find a broad literature in economics on the impacts of automation on employment. Frey and Osborne (2013) model the probability of computerization for different occupations and find that workers in transportation and logistics are among those at high risk of automation, partly due to their low wages and low levels of educational qualifications. However, their methods, including considering that entire occupations rather than single tasks within jobs can be automated have faced criticism. Arntz et al. (2016) have shown that jobs are a bundle of tasks and while tasks may be easy to automate, jobs are often not. As such, forecasts of a large number of job-losses from automation, such as the study by Frey and Osborne (2013) can overestimate the impact of automation. One of the questions that therefore emerges here is the composition of tasks involved in long haul trucking. We attempt to unpack this through our semi-structured interviews with truck operators.

Borland and Coelli (2017) study the impact of automation on employment in Australia and find that total work available did not decrease following the introduction of computerization and that job turnover in the labor market has also not increased due to computer-based technologies. Mokyr et al. (2015) have analyzed the history of automation and job losses, explaining how there has always been anxiety on how technological progress will cause substitution of machines for human workers, leading to unemployment. However, such scenarios do not come to pass because the long-run effects of technology are beneficial in terms of net job-creation (Autor, 2015) and technology only impacts the type of jobs available and what they might pay. Technology can also complement labor resulting in increased productivity, earnings, and demand for labor (Autor, 2015; Bessen, 2019).

Nevertheless, perceptions of large job losses from automation are present in the social and political discourse. In terms of the perception of automation in trucking, Dodel and Mesch (2020) have shown how workers in occupations involving a greater number of manual or physical tasks, such as in the case of long haul trucking, can have more negative perceptions regarding the impact of automation on their livelihood. Orii et al. (2021) analyzed discussions related to automation among the members of the r/Truckers subreddit and found that <1% of comments had positive views on automation.

Methods

We contribute to the literature in the field by drawing on the CFS data (United States Census Bureau, 2020) to get a reliable estimate of the density of long haul trucking in different regions and the operator-hours required for different routes. Our primary analysis is centered in the use of freight data along with routing and operator-hour algorithms to estimate the share of operator-hours that may be lost to automation. We complement this quantitative analysis with a limited number of interviews with long haul trucking stakeholders to understand the feasibility of a transfer-hub mode of deployment. Overall, our mixture of quantitative and qualitative methods is based on the triangulation method (Jick, 1979), useful for analyzing socio-technical transitions and emerging technologies. We further elaborate on our methods below.

Data. The Commodity Flow Survey is a well known dataset for transportation planning and research, produced every five years by the U.S. Bureau of Transportation Statistics, U.S. Census Bureau, and the U.S. Department of Commerce. The latest iteration, CFS 2017, is a sample of 5.9 million shipments from approximately 60,000 responding establishments (United States Census Bureau, 2020). We disregard inter-modal shipments and focus on shipments delivered through for-hire trucks and private trucks. We only consider shipments routed over >150 miles as those are commonly classified as long haul (FMCSA, 2020; Viscelli, 2018). This subset contains nearly 1.5 million trucking shipments detailing origin and destination states, shipment distance, weight, and financial quarter. The data also contain a weighting factor, which can be used to estimate the total number of shipments of that type in the population.

Routing. We draw on the Google Maps API and the GGMAP package (Kahle and Wickham, 2013) in R to estimate highway and (sub)urban splits for each shipment. For the purposes of routing we categorize two types of shipments in the dataset: intrastate and inter-state. Intrastate shipments are those where the shipment does not cross state borders. Inter-state shipments are those where the state of origin is different to the state of destination.

We apply a differentiated methodology to calculate the highway and (sub)urban splits for each shipment depending on whether the shipment is within a state or across it and depending on whether the shipment has listed origin and destination Metropolitan Statistical Areas (MSAs). MSAs are listed for many but not all shipments in the dataset. If MSAs are not provided, we use the closest approximation of origin or destination location.

The different types of shipments and the methods used to calculate the highway and (sub)urban splits and highway and urban average speeds are shown in Table 1.

For inter-state shipments we proceed as follows. Where possible we use the origin and destination MSAs from the CFS dataset in Google Maps and estimate the highway and urban distance ratios (details provided in the next subsection) for those routes, which we then apply to the actual shipment distance from the dataset. To do this, we assume that the precise origin or

Table 1 Truck routing calculation methods.

Shipment type	Case	Highway/urban ratio calculation	Highway, urban average speeds calculation	CFS Ton-miles share
Inter-state	Both origin & destination have MSA specified	Origin MSA area to destination MSA area in Google Maps	Google Maps	34%
	One of origin/destination has no MSA specified	MSA area to Rest of State centroid in google maps	Google Maps	42%
	Both origin & destination have no MSA specified	Rest of State centroid to Rest of State centroid in Google Maps	Google Maps	15%
Intrastate	Both origin & destination have MSA specified	Origin MSA area to destination MSA area in Google Maps	Google Maps	3%
	One of origin/destination has no MSA specified	FAF (2012)	Fixed to 55,25 mph	4%
	Both origin & destination have no MSA specified	FAF (2012)	Fixed to 55,25 mph	2%

destination is the centroid of the MSA. For shipments that specify either the origin or destination MSA (but not both), or specify neither origin nor destination metropolitan areas, we use the rest of the state centroid, which is defined as the centroid of all other areas of the state that are not listed MSAs. Note that this is an approximation that affects the estimate of the highway and sub(urban) split but *does not* affect the distance of the shipment, which is provided in the CFS dataset.

For intrastate journeys we apply the same method for shipments, which have specified origin and destination MSAs. For those that do not, we apply the highway and sub(urban) split derived from the Freight Analysis Framework dataset (Bureau of Transportation Statistics, 2012) by splitting the roads into those have average speeds below and above 50 mph.

Let the place of origin be designated as $p_{\mathrm{o,i}}$ and place of destination be designated as $p_{\mathrm{d,i}}$ where i is a shipment. Then, consider a shipment from $p_{\mathrm{o,i}}$ to $p_{\mathrm{d,i}}$ where $p_{\mathrm{o,i}}$ and $p_{\mathrm{d,i}}$ are set as per the cases listed in Table 1. The Google Maps API where applicable then provides us with detailed route directions, which list the amount of time driven for any stretch of road before the next turn and so on. This allows us to calculate speeds for each section and then split the drive into segments, which are greater than or equal to 50 mph (classified as highway) and below 50 mph (classified as urban or suburban). Note that the route suggested by Google Maps may be different depending on the time of day that the API request is sent. We therefore ran several iterations of the routing algorithm at different times of day and found no discernible difference to our results.

Let the highway segment of this journey be h_{i} and the urban sections u_{i}. Let the origin-destination distance be d_{i}. The highway to total ratio r_{i} is then defined as:

$$r_{\mathrm{i}} = \frac{h_{\mathrm{i}}}{d_{\mathrm{i}}} \tag{1}$$

We then use these calculated ratios for each origin-destination combination and apply them to the actual shipment distance from the CFS dataset. This allows us to calculate the highway and urban leg lengths $D_{\mathrm{S,H,i}}$ and $D_{\mathrm{S,U,i}}$ for the shipments in the dataset.

Let the shipment distance be $D_{\mathrm{S,i}}$. Then,

$$D_{\mathrm{S,H,i}} = D_{\mathrm{S,i}} * r_{\mathrm{i}} \tag{2}$$

and then,

$$D_{\mathrm{S,U,i}} = D_{\mathrm{S,i}} * (1 - r_{\mathrm{i}}) \tag{3}$$

Operator-hours calculation. The final step involves the calculation of urban and highway operator-hours. We assume the urban legs are equally split at the two ends of the journey with the highway leg in between. We apply a constraint of 11 h of daily driving as per hour of service (HOS) regulations (FMCSA, 2020) and then calculate the operator-hours required for the highway and urban legs of the journey. Using this information and the aforementioned weighting factor we are then able to calculate the total operator-hours as well as the share of highway and urban operator-hours.

Let day_1 hours be the number of hours remaining that can be driven on day 1 of the trip after completing the initial urban leg. Let O be operator-hours described for both highway leg O_{H} and urban leg O_{U}. Let highway and urban driving time be T_{H} and T_{U}, respectively, which can be calculated from the average velocities V_{H} and V_{U} for the respective segments also derived from the Google Maps API where applicable. Then for shipment i:

$$T_{\mathrm{U,i}} = \frac{D_{\mathrm{S,U,i}}}{V_{\mathrm{U,i}}} \tag{4}$$

and similarly

$$T_{\mathrm{H,i}} = \frac{D_{\mathrm{S,H,i}}}{V_{\mathrm{H,i}}} \quad (5)$$

Then, our algorithm to estimate the operator-hours is described below. Note that $\lceil(x)\rceil$ denotes the ceiling of x and $x\%y$ denotes the remainder of x when divided by y.

Algorithm 1. i from $1{:}I$

$day_{1,\mathrm{i}} = 11 - \frac{T_{\mathrm{U,i}}}{2}$
if $day_{1,\mathrm{i}} \geq T_{\mathrm{H,i}}$ **then**
  $O_{\mathrm{H,i}} = T_{\mathrm{H,i}}$
else
  if $day_{1,\mathrm{i}} < T_{\mathrm{H,i}}$ **then**
    $O_{\mathrm{H,i}} = T_{\mathrm{H,i}} + \left\lceil \left(\frac{T_{\mathrm{H,i}} - day_{1,\mathrm{i}}}{11}\right) \right\rceil * 10$
  end if
  if $\left(\frac{T_{\mathrm{H,i}} - day_{1,\mathrm{i}}}{11}\right)\%11 + \frac{T_{\mathrm{U,i}}}{2} > 11$ **then**
    $O_{\mathrm{U,i}} = T_{\mathrm{U,i}} + 10$
  else
    $O_{\mathrm{U,i}} = T_{\mathrm{U,i}}$
  end if
end if

The algorithm can be explained as follows. If the highway driving time is less than the number of driving hours remaining on day 1, then the shipment is simply completed on the day and the highway operator-hours are equal to the highway driving time. However, if the highway driving time exceeds this then the driver undertakes the journey over the following days with 10 h of rest following 11 h of driving as mandated by law. The urban driving time is simply the time taken to drive the urban leg if the second and final urban segment can be completed staying within the HOS requirements, else it is completed with a day of rest.

With the calculated urban and highway operator-hours for each trip we can then estimate the total operator-hours across both highways and urban areas using the trip weighting factor provided by the CFS dataset. The weighting factor is the estimate of the true number of trips of such type in the actual population and is available for each shipment in the CFS dataset. Let the weighting factor be Π. Further let shipment weight be W. Then for the total operator-hours O_{Total} we have:

$$O_{\mathrm{Total}} = \sum_{i=1}^{I} (O_{\mathrm{H,i}} + O_{\mathrm{U,i}}) * \Pi_{\mathrm{i}} * \frac{W_{\mathrm{i}}}{\mathrm{TL}} \quad (6)$$

where TL is truckload or the total weight that can be carried on one fully loaded semi truck.

Then the urban and highway share of the total operator-hours, US and HS, is simply:

$$\mathrm{HS} = \frac{\sum_{i=1}^{I} O_{\mathrm{H,i}} * \Pi_{\mathrm{i}} * \frac{W_{\mathrm{i}}}{\mathrm{TL}}}{O_{\mathrm{Total}}} \quad (7)$$

$$\mathrm{US} = \frac{\sum_{i=1}^{I} O_{\mathrm{U,i}} * \Pi_{\mathrm{i}} * \frac{W_{\mathrm{i}}}{\mathrm{TL}}}{O_{\mathrm{Total}}} \quad (8)$$

Note that the highway share (HS), across both inter-state and intrastate trucking, is the share of operator-hours at risk from automated highway trucking. US represents the share of hours that must still be driven by a human driver.

Notice that if the truckload is a constant, such as for, e.g., fully loaded class 8 semi trucks, then it cancels in both the numerator and denominator of equations (7) and (8) and is therefore irrelevant to our results. More information on methods including the limitations of our approach are provided in Supplemental Information (SI) Section 1.

Interviews. In order to obtain some assurance about the validity of the assumptions underlying the transfer-hub model, we undertook semi-structured interviews with stakeholders in the trucking industry using a purposeful sampling methodology (Robinson, 2014), which was formulated through authors' prior work in the automated vehicle domain (Mohan et al., 2020), as well as prior informal conversations with companies and researchers in this area, which helped us identify relevant questions and stakeholders. For our conversations with drivers in particular, we used snowball sampling (Naderifar et al., 2017): we identified an initial set of drivers who have a public profile (e.g., have podcasts or YouTube Channels about trucking) and asked them to introduce us to their colleagues. We stopped when we achieved data saturation (Guest et al., 2006): that is, when conversations with new drivers did not introduce us to new concepts or phenomena (Hennink et al., 2017). We spoke with stakeholders across automated trucking startups (2), truck drivers (5), trucking logistics operators (1), and labor union representatives (1). In terms of our selection of different interviewees, we deliberately sought to elevate the voices of truck drivers in our sample, relative to other actors such as automated trucking startup CEOs or logistics operators. This is because of two reasons. Firstly, operators were best placed to provide us with the operational challenges and opportunities for the transfer-hub model of automation, given they are currently in charge of the major task that automation may replace (driving). Second, much of the narrative and coverage around automated trucking in the popular media has focused on the claims made by private operators, without much consideration of whether long haul drivers themselves believe that a switch to automation is feasible. Note that our sample size was not designed to enable generalization to all the stakeholders in long haul trucking. Our sample size and qualitative method (semi-structured interviews) were instead selected with an idiographic approach (Robinson, 2014), focused on gathering detailed insights into the tasks truck drivers performed on a journey. The interviews also highlight interesting areas for future research. Most importantly, as part of the triangulation method (Jick, 1979) we use to analyze automated trucking, the interviews complement our quantitative analysis of the CFS data and the routing and operator-hour algorithms we present by providing a feasibility check on the deployment modes assumed in this paper and which have been promoted by technology companies. The full list of interviewees is provided in Table 2.

Results

Our analysis finds that up to 94% of operator-hours for truck drivers are impacted if the technology is deployed across the continental U.S. in all conditions. However, if restricted to the states where testing is currently taking place, only 10% of operator-hours are impacted. The capabilities of the technology and decisions around where and how AT should be deployed will therefore determine the extent of impacts on the long haul operator labor market. Below we first discuss the findings from our exploration of possible scenarios of deployment, and their associated impacts on operator-hours, if realized. We analyze the extent of the possible increase in short haul jobs if AT delivers cost and time savings in freight delivery and show that this is unlikely to outweigh the hours lost to automation. Then, we present the takeaways from our semi-structured interviews with stakeholders in long haul trucking. Finally, we end with a brief discussion on the labor impacts in sectors associated with long haul trucking.

Table 2 List of interviewees.

Number	Job Title	Affiliation	Date
1	Chief Technology Officer	Automated Trucking Startup A	24th April 2019
2	Legislative Representative	Teamsters	14th August 2019
3	Chief Operating Officer	Automated Trucking Startup B	22nd November 2019
4	Senior Manager	Large Trucking Logistics Firm	10th February 2020
5	Trucker	JB Hunt	22nd September 2020
6	CDL Coordinator	CCAC	22nd September 2020
7	Trucker	Holland	15th October 2020
8	Trucker	Holland	20th October 2020
9	Trucker	Poly Trucking	21st October 2020
10	Trucker	American Service Line	12th November 2020

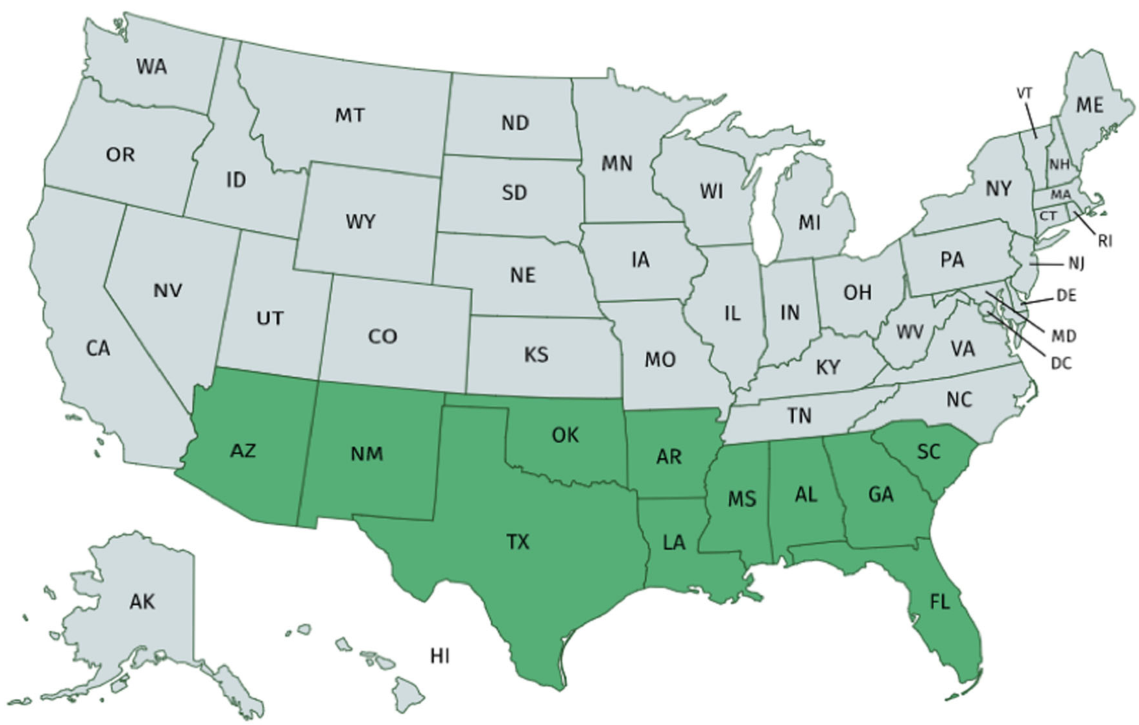

Fig. 2 Highlighted southern sun-belt states where a transfer-hub model may first be deployed.

Scenarios. We consider different scenarios of deployment, which correspond to constraints the technology may face in the near-medium term. They are as follows:

- Deployment in southern, sunny states only—To our knowledge, significant AT testing currently takes place in Florida, Texas, and Arizona. We hypothesize that initial highway deployment will be restricted to states in the southern sun-belt (see Fig. 2) to minimize risk of exposure to snow or hail, which may be outside the limits of safe operation of the technology and therefore only consider routes between or within these states.
- Deployment across all states over the financial quarters Q2 and Q3, which encompass the spring and summer months from April 1 to September 30—The CFS dataset lists the financial quarter of shipment allowing us to assess a scenario where all journeys that would be performed in more favorable weather are automated.
- Deployment for journeys above 500 miles only—While AT may be financially optimal at even short distances of around 100 miles (see SI Section 2), the time required for switching trailers and separating urban and highway legs could mean that journeys, which currently cannot be performed in a single day given HOS requirements will be automated first. Long haul operators can comfortably cover 500 miles without stopping so we set this as the threshold.
- Widespread deployment—deployment of transfer-hub-based AT across the United States and automation of all highway driving. While this is likely to take several years, it can be thought of as the most extreme scenario in terms of job losses. Note that most studies in the literature (Gittleman and Monaco, 2020; Viscelli, 2018) only consider this scenario .

Operator-hours. Figure 3 shows the impact on operator-hours across the different scenarios considered, with each scenario considered as additive (cumulative) to the previous scenario.

In scenario 4, with widespread deployment across the continental United States, 94% of current operator-hours may be automated. On the other hand, we find that deployment of automation only in places where companies are currently testing the technology, i.e., the southern states, limits the impact to just 10% of operator-hours (scenario 1). These estimates could represent a potential trajectory, over time, for impacts on

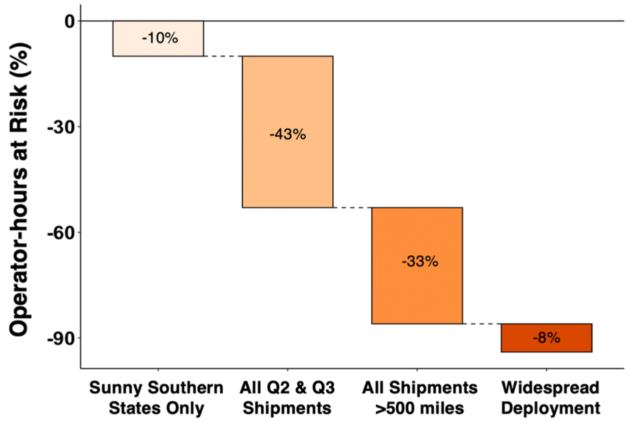

Fig. 3 The impact on operator-hours as each deployment scenario becomes feasible, cumulatively building on the previous deployment. Starting with only a 10% impact on operator-hours, automation can eventually put 94% of current long haul operator-hours at risk.

employment, if the technology improves over time. It is likely that near term deployment, if any, will be in favorable weather conditions and in states with favorable regulation, precisely those states where testing is currently taking place. If this is then further expanded to all parts of the country but only in the favorable weather months, i.e., financial Quarter 2 and 3, then we find that over half of operator-hours could be impacted. If automation is then extended to all shipments over 500 miles, we find that a further 33% of operator-hours are impacted. Other possible deployment scenarios and their impacts are shown in Table S1.

Changes in demand and price elasticity of freight. Automation is likely to dramatically reduce the need for labor, which constitutes roughly 40% of the cost of trucking (Williams and Murray, 2020). It is also likely to eliminate HOS requirements, which currently mean that-unless two drivers work in tandem-a truck remains idle for >60% of the time in order to allow the driver to rest (FMCSA, 2020). As a consequence, automation will reduce the amortized capital cost of the truck, which constitutes another 16% of the cost of trucking (Williams and Murray, 2020). Finally, the elimination of HOS requirements will make trucking faster than it is today, since the trucks will no longer have to stop for driver breaks. Cost and delivery time reductions in the delivery of freight through automated trucking may result in increased demand for trucking shipments. This may happen in two ways.

First, freight that was previously routed through air, train, or inter-modal services may now be shifted to long haul trucking. The modal choice of freight could therefore tilt in favor of trucking. There is considerable variation in modal elasticity estimates in the literature (De Jong et al., 2010). For example, Abdelwahab (1998) has estimated an elasticity of −1.44 and −0.99 for truck modal choice given an increase in trucking shipment time and cost, respectively. More recently, Christidis et al. (2009) reviewed a number of studies and found cross-elasticities of rail and road that range from 0.3 to 2. Dahl (2012) has found limited elasticity for trucking demand to fuel costs.

Second, cost and time reductions in freight delivered through trucking may increase demand for freight services in the economy overall. This increase in demand will increase the overall operator-hours required to fulfill long haul trucking shipments.

While elasticities vary widely in the literature for these factors as described above, we find that any resultant increase in demand for freight from some or all of these factors is likely to have a small impact on operator-hours. Further, this increase is unlikely to offset the overall operator-hours at risk due to automation. For example, even a 50% increase in demand for trucking services, which translates to an overall elasticity of 5, will only offset 5% of the at-risk hours due to automation, dropping the overall share from 94% to 89% of operator-hours at risk. This is simply because even though such an increase will lead to great demand for trucking services, the large majority of all operator-hours will still be needed on the highway (Fig. 3).

A large increase in demand of the magnitude of 50% or more for trucking services is in fact unlikely. Trucking already dominates the freight market with 70% of tonnage in the United States shipped through trucks (Association, 2020a). Freight that is today delivered through competing modes such as rail is often because of reasons other than cost or shipment time, for example, gross weight requirements, or because a close coupling between the location of the rail tracks and siting of various industrial facilities. Overall, we find limited evidence in support of the claim that increases in demand for trucking due to the economic and productivity gains from automation will create short haul jobs that will offset the highway operator-hours lost due to automation.

Jobs. What do these impacts on operator-hours imply for jobs? It is important to note that we do not directly characterize job losses from automation in trucking, focusing instead on the share of operator-hours that will be impacted. This is because data for the number of long haul truck drivers in the United States is not available to a high level of accuracy, owing to the large number of owner operators. Estimates have a wide range, from a few hundred thousand to millions of jobs. Studies that attempt to put a number on the total jobs lost therefore run the risk of a large error and are often not comparable due to the differing assumptions about the baseline number of jobs.

Nevertheless, translating the share of operator-hours impacted to different estimates of the number of long haul trucking operators in the United States can provide some insight. Previous analyses has estimated the number of long haul operators to be between 300,000–400,000 (Gittleman and Monaco, 2020; Viscelli, 2018). Waschik et al. (2021) estimate that there are roughly 550,000 long haul operators, in their modeling of the macro-economic impacts of automated trucking. Our results on the share of operator-hours at risk from automation would therefore mean that anywhere from 30,000 to >500,000 jobs may be impacted, depending on the scenarios presented above.

Impacts on workers in associated sectors. Finally, automation of highway trucking will have impacts on more than just tractor-trailer operators. While outside the focus of this paper, we offer a brief discussion here. First, automation of highway trucking could obviate the need for truck stops on the highway or at least make them far less frequented than before. Employment associated with operating truck stops will therefore be impacted. These truck stops currently employ about 70,000 people (See SI Table S2). On the other hand, the creation of truck ports to facilitate the transfer-hub model of AT will likely create new jobs. These will possibly involve new tasks such as switching trailers between the human and automated prime movers, offering services to human operators that were previously offered at truck stops, and maintenance and safety checks of sensors and other equipment on board automated trucks before their deployment on a route. To a first approximation, our analysis indicates that it is possible that the labor-hours lost at truck stops and other locations on highways could be compensated by new employment opportunities at transfer-hub ports (see SI Section 4 for more details). However, it is unclear whether operators in existing jobs at truck stops will be interested in or qualified for the new jobs that may arise from deployment of AT.

Discussion with Stakeholders. Scholars note that typically, only some of the tasks that constitute a job, are amenable to automation (Arntz et al., 2016; Autor, 2015). As such, tasks are easy to automate; jobs are often not. Existing literature has noted how truck driving jobs have increasingly been reduced to solely the task of driving (Viscelli, 2016, 2018). While this supports automation, one of the major challenges we envisaged to AT deployment on highways is maintenance and repairs. However, upon speaking with actors in the trucking industry (interviews 3–9) we found that employed truck drivers do not perform any significant maintenance and repairs on their truck. Instead, drivers simply call for assistance and trucking companies send out repair teams or arrange for a repair appointment at the nearest service station. ATs will therefore need to be able to send out distress signals and get assistance when needed. While some drivers indicated that they do perform some maintenance on their trucks (interview 7,9) this was restricted to minor fixes, for, e.g., a broken headlight. Often drivers indicated that their companies did not want them to try to repair problems with the truck but instead rely on expert help. Some trucking jobs are union jobs and in those cases the drivers often have clauses in their contracts that restrict them to only driving and not having to perform work such as maintenance and unloading (interview 2).

The benefits of automation in terms of shorter trip times and lower costs could quite easily and quickly be incorporated into trucking operations and logistics, as per one of our interviews with a Senior Manager at a logistics firm (interview 4). AT startup executives acknowledged that both weather and lighter regulation motivated their decision to test in states such as Florida and Texas (interviews 1,3). This raises questions regarding the widespread application of this technology to all parts of the continental United States.

Interestingly, every operator (interviews 5–10) we spoke to said that they could not see any major barriers to ATs performing highway journeys. Several of them (interviews 5,9,10) highlighted the difficulties posed by inclement weather conditions, places where lane markings are absent, and routing if global positioning system (GPS) signal is lost, which could mean that some routes continue to be driven by human drivers. All drivers we spoke to try to use their maximum allowed driving time of 11 h in a day. They all also indicated that there are often several weeks where they do not have any direct contact with their employers and simply perform their tasks as assigned. All drivers also indicated their trucks are tracked with GPS and therefore companies and customers are continuously aware of their location. Our discussions indicate that the job of long haul trucking has indeed effectively been reduced to a single task, which-given conducive external conditions-makes it amenable to automation.

We further offer some propositions for future research to explore. Owing to the limitations of our sample size, the lack of segmentation between different types of drivers, and no consideration of potential confounders, these are not intended as conclusions.

Proposition 1. *A larger volume of shorter trips may not compensate for the loss of work associated with automated long haul trips*. A transfer-hub model will require drivers to shift to short haul jobs. The most common pay structure currently for “truckload" drivers, who haul full truckloads worth of generic containerized freight for a trucking company (and not for a shipper like Walmart or Target), is payment per mile of haulage. They are often not paid for the waiting and paperwork that occurs at the beginning and end of each trip. Truckload drivers seek to maximize the time they keep the cargo moving. A shift to shorter trips would increase the ratio of stationary (unpaid) to driving (paid) time, reducing their wages per hour worked (Viscelli, 2016). In our limited sample there was near consensus among our interviewees that the shorter trips and lower pay that may come from urban driving jobs will be unattractive (interviews 5–9). Therefore, one of the interesting questions for future qualitative research based on a large sample size would be to understand operators' views on shifting to short haul jobs and whether these might hinder or accelerate a shift to the transfer-hub model.

Proposition 2. *Transfer hub deployment could create short haul jobs in locations that are different from where long haul truckers currently live*. A shift to only urban driving will likely require operators to live in sub(urban) areas. At least one operator we spoke with expressed reluctance to shift to short haul trucking for this reason alone (interview 9). Around 40% of older truckers come from rural areas so it's possible that the geographical shift will prove a barrier to transitioning current operators to short haul jobs, which speaks to the cultural significance long haul trucking jobs have carried in the U.S. (Levy, 2015). However, these new jobs may prove attractive to new truckers joining the workforce, who are increasingly from urban areas (Cheeseman Day and Hait, 2019). Again, this is an important question for future work and will require careful study. Shifting of employment from rural to urban areas for trucking will naturally have political implications, particularly given the existing rural-urban divide in the American political landscape (Thiede et al., 2017).

Proposition 3. *Partial automation is viewed negatively by heavy truck operators, as previous studies have suggested* (Slowik and Sharpe, 2018). Partial automation systems have also been criticized in the broader safety literature as they may lead to disengaged and distracted operators who are too quick to trust the technology and will be unable to react in a timely and safe manner should something go wrong (Endsley and Kiris, 1995). Many companies involved in self-driving technology are in fact skipping partial automation systems and focusing solely on full automation (Ayre, 2017, Naughton, 2017, Volvo, 2017). Economically, if companies still have to pay drivers for their labor and also pay for the system, it is difficult to see how this would be attractive compared to the current system of only paying drivers, unless the cost savings from increased safety were significant. All drivers (interviews 5, 7–10) in our limited sample expressed dislike for partial automation systems that they have used or experienced such as lane assist and emergency braking. How might views on such technologies differ depending on the age and experience of truck drivers? We believe that future work on the transition from human driven to automated trucks and the potential role of partial automation systems as a bridge technology must take into account the views of drivers, as they will be the primary users of such technologies.

Conclusion and policy implications

Automation of the major part of the job—in this case highway driving—will naturally put downward pressure on wages in the long haul trucking industry. It is unclear that the labor supply will easily adjust to the new level of prevailing wage and operating requirements (short haul jobs) in the market. Our limited number of interviews certainly highlight the challenges employers may face in transitioning long haul operators to different jobs such as short haul driving. It also suggests that the deployment of AT is being driven by techno-economic considerations alone with limited understanding of the social consequences, consistent with the broader narrative around automated mobility technologies (Bissell et al., 2020). Although companies have claimed that such technologies will benefit truck drivers, our evidence does not suggest that the motivations of truck drivers are part of the designed operation of this technology. Moving away from industry led visions of AT futures will require a greater understanding of the motivations and interests of long haul operators, and a participatory approach to shaping AT deployment (Mladenović, 2019).

In the currently envisioned transfer-hub model, short term adjustment costs are likely and potentially notable. As we show, a significant share of operator-hours will be affected if the technology is deployed in all conditions and locations. Further, we argue that this result is robust to increase in demand for freight delivered through trucking if as assumed, the cost of long haul trucking falls due to automation.

We do caution however that the potential loss of a significant share of operator-hours to automation need not be necessarily viewed as permanent unemployment or as a permanent welfare loss. Long haul trucking has been characterized by turnover rates of nearly 100% in recent years (Association, 2020b). The profession is increasingly unattractive to potential new entrants with most new operators lasting less than a year in the job. This has occurred in substantial part through a concerted effort to make trucking cheaper by paying drivers less; for example, by encouraging many drivers to operate as independent contractors (Viscelli, 2016). Wages may need to increase as these arms-length employment arrangements are challenged in court and as it becomes increasingly difficult to find new drivers. This dynamic may strengthen the economic case for automation.

Historically, technological change has resulted in short term employment shocks but realignment in the labor market means that these shocks have limited impact on the broader economy in the long-run, as new industries grow and workers transition to new jobs with new skill requirements (Bessen, 2019). Long haul operators may therefore move across to different sectors after a period of unemployment, some may transition to lower paying short haul jobs that will be created by the transfer-hub model, others may retire prematurely (Waschik et al., 2021).

The sharp reduction in labor cost makes the economics of ATs compelling but will disrupt livelihoods and, by potentially shifting demand from rail to trucks, likely also increase emissions of greenhouse gasses and other air pollutants (Kaack et al., 2018). The threat of jobs lost due to automation in trucking may also have profound political impacts, the existing literature has found increased support for radical right wing parties as the risk of automation increases (Im et al., 2019). Policymakers could demand that-in exchange for permission to deploy ATs on public roads-truck operators re-invest some of the monetary benefits of reduced labor costs to ameliorate the disruption to employment and in reducing the environmental footprint of the trucking industry (Viscelli, 2020). Ultimately, societal and political choices can determine the mode of deployment of AT capabilities, and accordingly, the winners and losers of any shift to automation of long haul trucking.

Data availability

The Commodity Flow Survey (2017) dataset used in this paper is publicly available from the U.S. Census Bureau.

Received: 28 July 2021; Accepted: 23 February 2022;
Published online: 15 March 2022

References

Abdelwahab WM (1998) Elasticities of mode choice probabilities and market elasticities of demand: evidence from a simultaneous mode choice/shipment-size freight transport model. Transp Res Part E: Logist Transp Rev 34:257–266

Acemoglu D, Autor, D (2011) Skills, tasks and technologies: implications for employment and earnings. In Handbook of labor economics, vol. 4. Elsevier. pp. 1043–1171

Acemoglu D, Restrepo P (2019) Automation and new tasks: How technology displaces and reinstates labor. J Econ Perspect 33:3–30

Arntz M, Gregory T, Zierahn U (2016) The risk of automation for jobs in OECD countries: a comparative analysis. OECD

Association AT (2021). News & Insights. https://www.trucking.org/news-insights (2020). Accessed: 05-Apr-2021

Association AT (2020) News & Insights. https://www.trucking.org/news-insights/truckload-turnover-rises-third-quarter. Accessed: 24-Apr-2021

Autor DH (2015) Why are there still so many jobs? the history and future of workplace automation J Econ Perspect 29:3–30. https://www.aeaweb.org/articles?id=10.1257/jep.29.3.3.

Ayre J (2021) Google/Waymo stopped testing level 3 self-driving tech after testers literally fell asleep while using it, switched to full autonomy. https://tinyurl.com/yhj5cddb (2017). Accessed: 05-Apr-2021

Bessen J (2019) Automation and jobs: when technology boosts employment. Econ Policy 34:589–626

Bissell D, Birtchnell T, Elliott A, Hsu EL (2020) Autonomous automobilities: the social impacts of driverless vehicles. Curr Sociol 68:116–134

Borland J, Coelli M (2017) Are robots taking our jobs? Australian Econ Rev 50:377–397

Bureau of Transportation Statistics (2012). Freight analysis framework version 4. https://www.bts.dot.gov/faf/faf4. Accessed: 20-Nov-2020

Cheeseman Day J, Hait AW (2021) Number of truckers at all-time High. https://tinyurl.com/3uj4urr9 (2019). Accessed: 6-May-2021

Christidis P, Leduc G et al. (2009) Longer and heavier vehicles for freight transport. JRC European Commission. pp. 1–40

Dahl CA (2012) Measuring global gasoline and diesel price and income elasticities. Energy Policy 41:2–13

De Jong G, Schroten A, van Essen H, Otten M, Bucci P (2010) The price sensitivity of road freight transport: a review of elasticities. Applied transport economics, a management and policy perspective, De Boeck, Antwerp

Dodel M, Mesch GS (2020) Perceptions about the impact of automation in the workplace. Inform Commun Soc 23:665–680

Endsley MR, Kiris EO (1995) The out-of-the-loop performance problem and level of control in automation. Human factors 37:381–394

FMCSA (2020) Hours of Service (HOS). https://www.fmcsa.dot.gov/regulations/hours-of-service (2020). Accessed: 20-Nov-2020

Frey CB, Osborne M (2013) The future of employment. Oxford

Gilroy R (2020) Ike maintains automation will add truck driving jobs. https://www.ttnews.com/articles/ike-maintains-automation-will-add-truck-driving-jobs (2019). Accessed: 18-Jun-2020

Gittleman M, Monaco K (2020) Truck-driving jobs: Are they headed for rapid elimination? ILR Rev 73:3–24

Groshen EL, Helper S, MacDuffie JP, Carson C (2019) Preparing US workers and employers for an autonomous vehicle future. WE Upjohn Institute for Employment Research

Guest G, Bunce A, Johnson L (2006) How many interviews are enough? an experiment with data saturation and variability. Field Methods 18:59–82

Hennink MM, Kaiser BN, Marconi VC (2017) Code saturation versus meaning saturation: how many interviews are enough? Qual Health Res 27:591–608

Hickman J, Levy F, Burks S, Viscelli S, Lee J (2018) An NSF workshop report: effect of automated trucks on the truck driver workforce. Tech Rep, Virginia Tech Transportation Institute's Center for Truck and Bus Safety

Im ZJ, Mayer N, Palier B, Rovny J (2019) The "losers of automation": a reservoir of votes for the radical right? Res Polit 6:2053168018822395

Jick TD (1979) Mixing qualitative and quantitative methods: Triangulation in action. Admin Sci Quart 24:602–611

Kaack LH, Vaishnav P, Morgan MG, Azevedo IL, Rai S (2018) Decarbonizing intraregional freight systems with a focus on modal shift. Environ Res Lett 13:083001

Kahle D, Wickham H (2013) ggmap: spatial visualization with ggplot2. The R J 5:144–161

Leonard JJ, Mindell DA, Stayton EL (2020) Autonomous vehicles, mobility, and employment policy: the roads ahead. Massachusetts Inst. Technol., Cambridge, MA, Rep. RB02-2020

Levy KE (2015) The contexts of control: Information, power, and truck-driving work. Inform Soc 31:160–174

Mladenović MN (2019) How should we drive self-driving vehicles? anticipation and collective imagination in planning mobility futures. In: The governance of smart transportation systems. Springer. pp. 103–122

Mohan A, Sripad S, Vaishnav P, Viswanathan V (2020) Trade-offs between automation and light vehicle electrification. Nat Energy 5:543–549

Mokyr J, Vickers C, Ziebarth NL (2015) The history of technological anxiety and the future of economic growth: Is this time different? J Econ Perspect 29:31–50

Naderifar M, Goli H, Ghaljaie F (2017) Snowball sampling: A purposeful method of sampling in qualitative research. Stride Dev Med Educ 14:e67670

Naughton K (2017) Ford's dozing engineers side with Google in full autonomy push. https://tinyurl.com/2h98swvp. Accessed: 05-Apr-2021

Orii L, Tosca D, Kun AL, Shaer O (2021) Perceptions on the future of automation in r/truckers. In: Extended abstracts of the 2021 CHI conference on human factors in computing systems. pp. 1–6

Robinson OC (2014) Sampling in interview-based qualitative research: A theoretical and practical guide. Qual Res Psychol 11:25–41

Slowik P, Sharpe B (2018) Automation in the long haul: challenges and opportunities of autonomous heavy-duty trucking in the United States. International Council on Clean Transportation, Working Paper

Thiede B, Greiman L, Weiler S, Beda S, Conroy T (2017) Six charts that illustrate the divide between rural and urban America. The Conversation

Transport Topics (2018) How drivers and autonomous trucks could work together to move freight. https://www.ttnews.com/articles/how-drivers-and-autonomous-trucks-could-work-together-move-freight. Accessed: 18-Jun-2020

United States Census Bureau. CFS 2017. https://www.census.gov/programs-surveys/cfs.html (2020). Accessed: 20-Nov-2020

Viscelli S (2020) Will robotic trucks be" sweatshops on wheels?". Issue Sci Technol 37:79–88

Viscelli S (2016) The big rig: trucking and the decline of the American dream. Univ of California Press

Viscelli S (2018) Driverless? Autonomous trucks and the future of the American trucker. Center for Labor Research and Education, University of California, Berkeley, and Working Partnerships USA

Volvo (2017) Volvo Cars CEO urges governments and car industry to share safety-related traffic data. https://tinyurl.com/28ezuxds. Accessed: 05-Apr-2021

Waschik R, Friedman D, Taylor C, Boatner J et al. (2021) Macroeconomic impacts of automated driving systems in long-haul trucking. Tech Rep, United States. Department of Transportation. Intelligent Transportation

Wertheim J (2020) Automated trucking, a technical milestone that could disrupt hundreds of thousands of jobs, hits the road. https://tinyurl.com/przw2kj3. Accessed: 18-Jun-2020

Williams N, Murray D (2020) An analysis of the operational costs of trucking: 2020 update

Acknowledgements

The authors would like to thank Venkatasubramanian Viswanathan, Jeremy Michalek, Nicholas Muller, Jessika Trancik, and Jack Stilgoe for their helpful comments on earlier drafts of this paper. This work was supported by the CMU College of Engineering, Department of Engineering & Public Policy; and the Block Center for Technology and Society at CMU.

Author contributions

AM and PV designed the research, conceived the paper, and developed the routing and operator-hours algorithm used in this work. AM performed the analysis and created the figures; and AM and PV wrote the paper.

Competing interests

The authors declare no competing interests.

Ethical approval

Not applicable.

Informed consent

Informed consent was obtained from all interviewees.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1057/s41599-022-01103-w.

Correspondence and requests for materials should be addressed to Parth Vaishnav.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022

APPENDIX C

Supplementary Information for Chapter 5

Table C.1 shows the fleet metrics for 4 of the selected sample days. Metrics for the other 3 days in our sample set were shown previously in Table 5.3.

Table C.1: **Fleet Metrics for Selected Days**

Day	Pack Size (kWh)	DC Port / Sq-mile	Total trips	Total taxis	Median pickup time (mins)	Median dead-heading (%)	Fleet VMT (miles)	Trips Aborted (%)	Median time to find a DC port (mins)
1st May 2021	40	0.5	140558	13965	3	28%	965452	0%	29
1st May 2021	60	0.5	140558	13145	3	28%	924970	0%	27
1st May 2021	40	1	140558	12163	3	29%	837552	0%	4
1st May 2021	60	1	140558	12438	3	29%	840778	1%	3
26th July 2021	40	0.5	88999	11241	4	27%	766678	2%	25
26th July 2021	60	0.5	88999	9803	4	28%	736383	2%	26
26th July 2021	40	1	88999	9242	4	29%	674985	2%	5
26th July 2021	60	1	88999	9071	4	29%	666786	2%	4
6th Sept. 2021	40	0.5	85770	10324	3	28%	767592	2%	19
6th Sept. 2021	60	0.5	85770	9834	3	29%	746800	2%	18
6th Sept. 2021	40	1	85770	9593	3	29%	708521	2%	5
6th Sept. 2021	60	1	85770	9542	3	29%	704437	2%	4
15th Oct. 2021	40	0.5	180094	19826	3	27%	1349738	1%	39
15th Oct. 2021	60	0.5	180094	17938	3	28%	1303656	1%	42
15th Oct. 2021	40	1	180094	16365	3	29%	1155574	1%	5
15th Oct. 2021	60	1	180094	16244	3	30%	1149289	1%	4

Figure C.1 below shows GHG costs per trip for the seven different days and four different scenarios we consider in our analysis. Lowest externalities are obtained with a combination of a smaller battery pack size of 40kWh with 1 dedicated fast-charger per square mile.

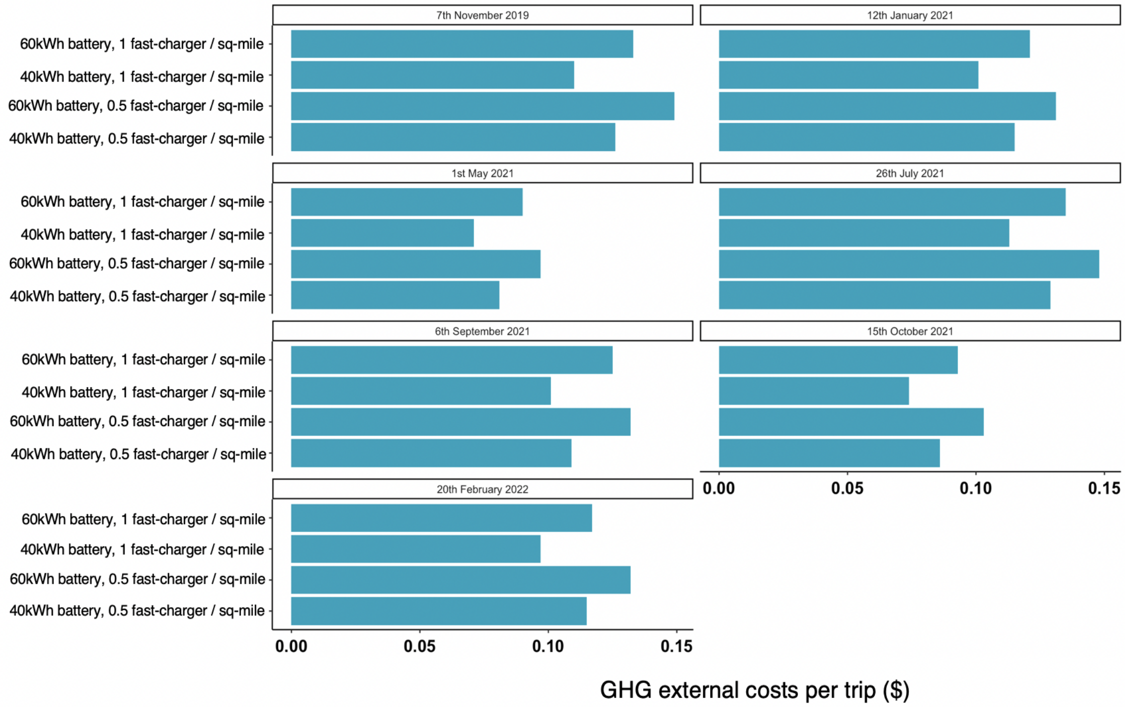

Figure C.1: **GHG costs per trip for each sample day**

Figure C.2 below shows total external costs per trip for the seven different days and four different scenarios we consider in our analysis. Lowest externalities are obtained with a combination of a smaller battery pack size of 40kWh with 1 dedicated fast-charger per square mile. External costs for a larger pack size with the same level of charging infrastructure are almost the same across the different days considered, as increased GHG externalities are compensated by lower VMT based externalities. Total external costs vary significantly across the different days for each particular scenario due to unique ridesourcing demand patterns and the effect of heating and cooling needs on energy use, for example ranging from less than 1 USD per trip to nearly 1.5 USD per trip for the lowest cost scenario.

Figure C.3 shows the average total external costs per trip across the seven sample days when charging related GHG emissions are zero (i.e. a fully zero carbon electricity grid).

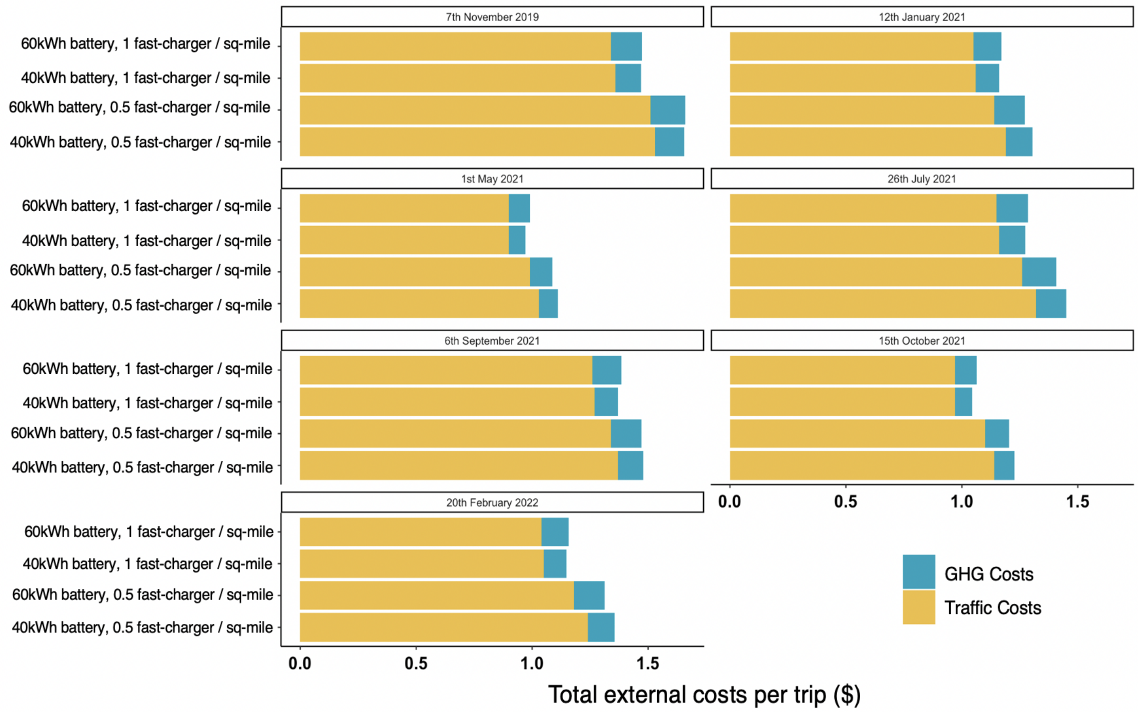

Figure C.2: **Total external costs per trip for each sample day**

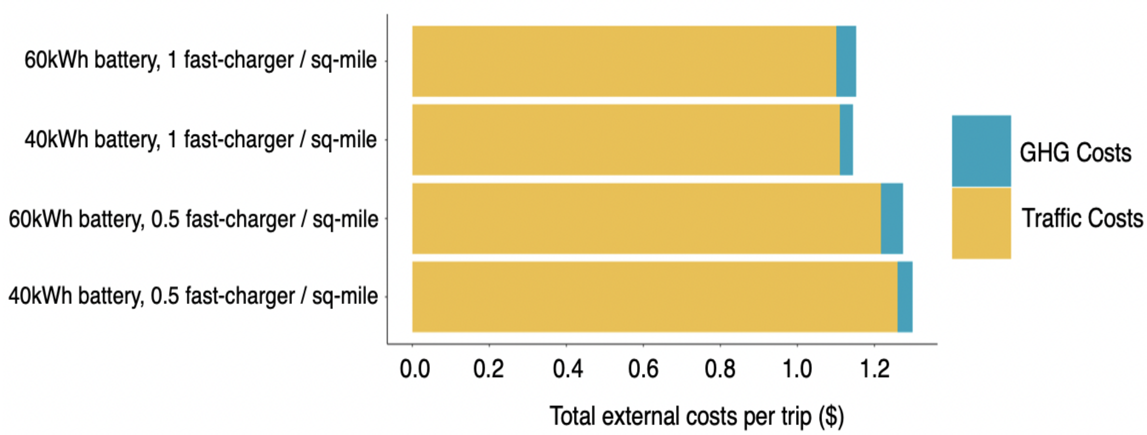

Figure C.3: **Total external costs per trip averaged across all sample days when charging related GHG emissions are zero**

Figure C.4 shows the power demand from fast-charging infrastructure and level 2 chargers for the pre-pandemic day we model, 7th November 2019, with heavy trip demand. A significant number of drivers abort their shift due to unavailable fast-charging infrastructure and increase the power demand from level 2 chargers which rises continually throughout the day albeit with a small dip during peak hours in the evening between 5pm - 7pm when drivers maximize time spent on the road for increased revenues.

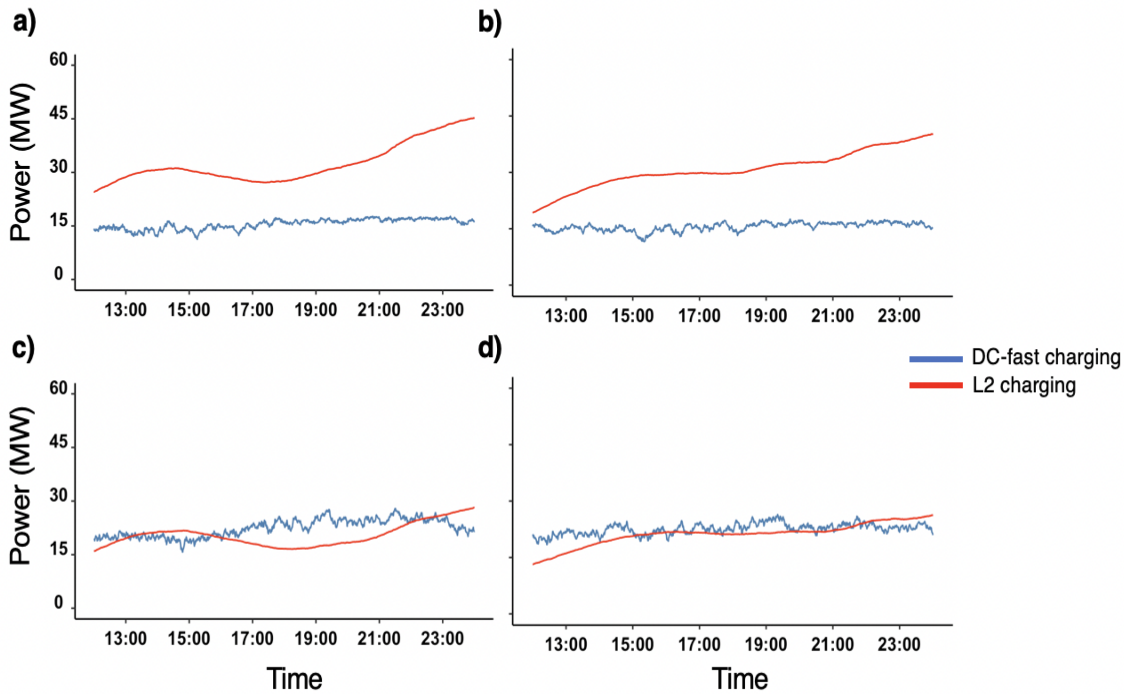

Figure C.4: **Fast-charger and L2 charger power draw for 7th November 2019** **a)** and **b)** represent fast-charging infrastructure of 1 dedicated fast-charger per 2 square miles and the small and large battery pack respectively. **c)** and **d)** represent fast-charging infrastructure of 1 dedicated fast-charger per square mile and the small and large battery pack respectively.

Figure C.5 shows the total external costs per trip averaged across all sample days when we use a higher valuation for the social cost of carbon. Increasing the valuation of climate damages increases total external costs by 16-18% compared to the results presented in Figure 5.4. Further, due to the higher valuation of climate damages, a 40kWh battery is always preferred to the larger 60kWh battery given the same level of charging infrastructure.

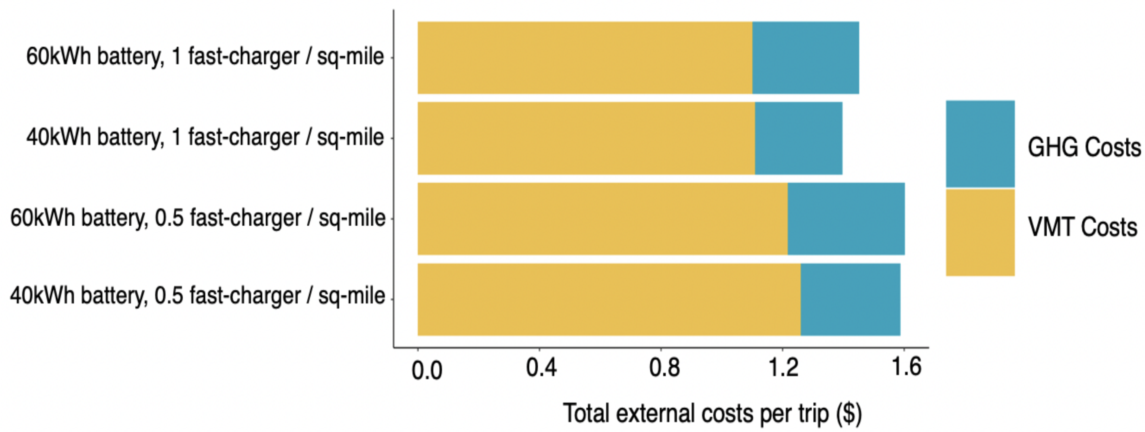

Figure C.5: **Total external costs per trip averaged across all sample days when using the 95th percentile SCC**

Table C.2: **Fleet Metrics for 20th February 2022 across different times for adjusting driver supply**

Adjustment time for supply of drivers	Pack Size (kWh)	DC Port / Sq-mile	Total trips	Total taxis	Median pickup time (mins)	Median dead-heading (%)	Fleet VMT (miles)	Trips Aborted (%)	Median time to find a DC port (mins)
2 minutes	40	0.5	121869	14944	3	28%	994297	1%	31
2 minutes	60	0.5	121869	12940	3	28%	952621	1%	35
2 minutes	40	1	121869	11595	3	29%	843284	1%	5
2 minutes	60	1	121869	11597	3	29%	836310	1%	4
4 minutes	40	0.5	121869	14806	3	28%	996694	1%	31
4 minutes	60	0.5	121869	13037	3	28%	954615	1%	35
4 minutes	40	1	121869	11647	3	28%	838413	1%	5
4 minutes	60	1	121869	11488	3	29%	833549	1%	4

BIBLIOGRAPHY

[1] Davis, S. & Boundy, R. G. Transportation energy data book: Edition 38. Tech. Rep., Oak Ridge National Lab.(ORNL), Oak Ridge, TN (United States) (2020).

[2] Davis, S. J. *et al.* Net-zero emissions energy systems. *Science* **360**, eaas9793 (2018).

[3] Jenn, A. Emissions benefits of electric vehicles in uber and lyft ride-hailing services. *Nature Energy* **5**, 520–525 (2020).

[4] Roca, J. B., Vaishnav, P., Morgan, M. G., Mendonça, J. & Fuchs, E. When risks cannot be seen: Regulating uncertainty in emerging technologies. *Research Policy* **46**, 1215–1233 (2017).

[5] Rayner, S. The novelty trap: Why does institutional learning about new technologies seem so difficult? *Industry and Higher Education* **18**, 349–355 (2004).

[6] Mohan, A., Sripad, S., Vaishnav, P. & Viswanathan, V. Trade-offs between automation and light vehicle electrification. *Nature Energy* **5**, 543–549 (2020).

[7] Taiebat, M., Brown, A. L., Safford, H. R., Qu, S. & Xu, M. A review on energy, environmental, and sustainability implications of connected and automated vehicles. *Environ. Sci. Technol.* **52**, 11449–11465 (2018).

[8] Stephens, T. *et al.* Estimated bounds and important factors for fuel use and consumer costs of connected and automated vehicles. Tech. Rep., National Renewable Energy Lab.(NREL), Golden, CO (United States) (2016).

[9] U.S. Energy Information Administration. Study of the potential energy. consumption impacts of connected and automated vehicles. Tech. Rep., U.S. Energy Information Administration, U.S. Department of Energy (2017).

[10] Offer, G. J. Automated vehicles and electrification of transport. *Environ. Sci. Technol.* **8**, 26–30 (2015).

[11] Andrew J. Hawkins. Not all of our self-driving cars will be electrically powered — here's why. https://www.theverge.com/2017/12/12/16748024/self-driving-electric-hybrid-ev-av-gm-ford/ (2017). Accessed: 30-January-2019.

[12] Sripad, S. & Viswanathan, V. Evaluation of current, future, and beyond li-ion batteries for the electrification of light commercial vehicles: Challenges and opportunities. *J. Electrochem. Soc.* **164**, E3635–E3646 (2017).

[13] Ouster. OS1 Datasheet. https://static1.squarespace.com/static/58dd3a75c534a52312bad724/t/5ca679da15fcc032a353ad04/1554414043557/OS-1+Datasheet.pdf (2019). Accessed: 20-March-2019.

[14] Bosch. Chassis systems control: Mid-range radar sensor (mrr) for front and rear applications. Tech. Rep., Robert Bosch GmbH (2015).

[15] PtGrey. Dragonfly2 0.8 MP Mono FireWire 1394a Board Level (Sony ICX204). https://www.ptgrey.com/dragonfly2-08-mp-mono-firewire-1394a-board-level-sony-icx204 (2018). Accessed: 20-December-2018.

[16] Velodyne. HDL-64E. hhttps://velodynelidar.com/hdl-64e.html (2019). Accessed: 05-January-2019.

[17] Carroll, A., Heiser, G. *et al.* An analysis of power consumption in a smartphone. In *USENIX annual technical conference*, vol. 14, 21–21 (Boston, MA, 2010).

[18] Sepulcre, M., Gozalvez, J. & Hernandez, J. Cooperative vehicle-to-vehicle active safety testing under challenging conditions. *Transportation research part C: emerging technologies* **26**, 233–255 (2013).

[19] Gawron, J. H., Keoleian, G. A., De Kleine, R. D., Wallington, T. J. & Kim, H. C. Life cycle assessment of connected and automated vehicles: Sensing and computing subsystem and vehicle level effects. *Environ. Sci. Technol.* **52**, 3249–3256 (2018).

[20] Liu, S., Tang, J., Zhang, Z. & Gaudiot, J.-L. Caad: Computer architecture for autonomous driving. *arXiv preprint arXiv:1702.01894* (2017).

[21] Lin, S.-C. *et al.* The architectural implications of autonomous driving: Constraints and acceleration. In *Proceedings of the Twenty-Third International Conference on Architectural Support for Programming Languages and Operating Systems*, 751–766 (ACM, 2018).

[22] Nvidia. Nvidia Drive AGX Pegasus. https://www.nvidia.com/en-us/self-driving-cars/drive-platform/hardware/ (2018). Accessed: 20-October-2018.

[23] Eric Walz. Tesla Shares Details of its New Self-Driving Chipset at 'Autonomy Investor Day'. https://www.futurecar.com/3153/Tesla-Shares-Details-of-its-New-Self-Driving-Chipset-at-Autonomy-Investor-Day (2019). Accessed: 30-April-2019.

[24] Chowdhury, H., Alam, F., Khan, I., Djamovski, V. & Watkins, S. Impact of vehicle add-ons on energy consumption and greenhouse gas emissions. *Procedia Eng.* **49**, 294–302 (2012).

[25] Chen, Y. & Meier, A. Fuel consumption impacts of auto roof racks. *Energy Policy* **92**, 325–333 (2016).

[26] Liu, J., Kockelman, K. M. & Nichols, A. Anticipating the emissions impacts of smoother driving by connected and autonomous vehicles, using the moves model. In *Transportation Research Board 96th Annual Meeting* (2016).

[27] Prakash, N., Cimini, G., Stefanopoulou, A. G. & Brusstar, M. J. Assessing fuel economy from automated driving: influence of preview and velocity constraints. In *ASME 2016 Dynamic Systems and Control Conference*, V002T19A001–V002T19A001 (American Society of Mechanical Engineers, 2016).

[28] Mersky, A. C. & Samaras, C. Fuel economy testing of autonomous vehicles. *Transport. Res. C-Emer.* **65**, 31–48 (2016).

[29] Gao, Z. *et al.* Evaluation of electric vehicle component performance over eco-driving cycles. *Energy* **172**, 823–839 (2019).

[30] United States Environmental Protection Agency. Certificate Summary Information Report, TESLA MOTORS: MODEL 3 LONG RANGE, 02/28/2019. https://iaspub.epa.gov/otaqpub/display_file.jsp?docid=46968&flag=1 (2019). Accessed: 30-April-2019.

[31] Brad Templeton. Elon Musk's War On LIDAR: Who Is Right And Why Do They Think That? https://www.forbes.com/sites/bradtempleton/2019/05/06/elon-musks-war-on-lidar-who-is-right-and-why-do-they-think-that/7fe42c4f2a3b (2019). Accessed: 10-June-2019.

[32] Morgan, M. G., Henrion, M. & Small, M. *Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis* (Cambridge university press, 1992).

[33] Harlow, J. E. *et al.* A wide range of testing results on an excellent lithium-ion cell chemistry to be used as benchmarks for new battery technologies. *J. Electrochem. Society* **166**, A3031–A3044 (2019).

[34] Anderson, J. M. *et al.* *Autonomous vehicle technology: A guide for policymakers* (Rand Corporation, 2014).

[35] Mohan, A. & Vaishnav, P. Impact of automation on long haul trucking operator-hours in the united states. *Humanities and Social Sciences Communications* **9**, 1–10 (2022).

[36] Transport Topics. How Drivers and Autonomous Trucks Could Work Together to Move Freight. https://www.ttnews.com/articles/how-drivers-and-autonomous-trucks-could-work-together-move-freight (2018). Accessed: 18-June-2020.

[37] Hickman, J., Levy, F., Burks, S., Viscelli, S. & Lee, J. An NSF Workshop Report: Effect of Automated Trucks on the Truck Driver Workforce. Tech. Rep., Virginia Tech Transportation Institute's Center for Truck and Bus Safety (2018).

[38] Wertheim, J. Automated trucking, a technical milestone that could disrupt hundreds of thousands of jobs, hits the road. https://tinyurl.com/przw2kj3 (2020). Accessed: 18-June-2020.

[39] United States Census Bureau. CFS 2017. https://www.census.gov/programs-surveys/cfs.html (2020). Accessed: 20-November-2020.

[40] Gilroy, R. Ike Maintains Automation Will Add Truck Driving Jobs. https://www.ttnews.com/articles/ike-maintains-automation-will-add-truck-driving-jobs (2019). Accessed: 18-June-2020.

[41] Acemoglu, D. & Autor, D. Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics*, vol. 4, 1043–1171 (Elsevier, 2011).

[42] Arntz, M., Gregory, T. & Zierahn, U. The risk of automation for jobs in oecd countries: A comparative analysis (2016).

[43] Acemoglu, D. & Restrepo, P. Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives* **33**, 3–30 (2019).

[44] Viscelli, S. Driverless? autonomous trucks and the future of the american trucker (2018).

[45] Gittleman, M. & Monaco, K. Truck-driving jobs: Are they headed for rapid elimination? *ILR Review* **73**, 3–24 (2020).

[46] Groshen, E. L., Helper, S., MacDuffie, J. P. & Carson, C. Preparing us workers and employers for an autonomous vehicle future (2019).

[47] Waschik, R., Friedman, D., Taylor, C., Boatner, J. *et al.* Macroeconomic impacts of automated driving systems in long-haul trucking. Tech. Rep., United States. Department of Transportation. Intelligent Transportation ... (2021).

[48] Leonard, J. J., Mindell, D. A. & Stayton, E. L. Autonomous vehicles, mobility, and employment policy: The roads ahead (2020).

[49] Frey, C. B. & Osborne, M. The future of employment (2013).

[50] Borland, J. & Coelli, M. Are robots taking our jobs? *Australian Economic Review* **50**, 377–397 (2017).

[51] Mokyr, J., Vickers, C. & Ziebarth, N. L. The history of technological anxiety and the future of economic growth: Is this time different? *Journal of economic perspectives* **29**, 31–50 (2015).

[52] Autor, D. H. Why Are There Still So Many Jobs? The History and Future of Workplace Automation. *Journal of Economic Perspectives* **29**, 3–30 (2015). URL https://www.aeaweb.org/articles?id=10.1257/jep.29.3.3.

[53] Bessen, J. Automation and jobs: When technology boosts employment. *Economic Policy* **34**, 589–626 (2019).

[54] Dodel, M. & Mesch, G. S. Perceptions about the impact of automation in the workplace. *Information, Communication & Society* **23**, 665–680 (2020).

[55] Orii, L., Tosca, D., Kun, A. L. & Shaer, O. Perceptions on the future of automation in r/truckers. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–6 (2021).

[56] Jick, T. D. Mixing qualitative and quantitative methods: Triangulation in action. *Administrative science quarterly* **24**, 602–611 (1979).

[57] FMCSA. Hours of Service (HOS). https://www.fmcsa.dot.gov/regulations/hours-of-service (2020). Accessed: 20-November-2020.

[58] Kahle, D. & Wickham, H. ggmap: Spatial visualization with ggplot2. *The R Journal* **5**, 144–161 (2013).

[59] Bureau of Transportation Statistics. Freight Analysis Framework Version 4. https://www.bts.dot.gov/faf/faf4 (2012). Accessed: 20-November-2020.

[60] Robinson, O. C. Sampling in interview-based qualitative research: A theoretical and practical guide. *Qualitative research in psychology* **11**, 25–41 (2014).

[61] Naderifar, M., Goli, H. & Ghaljaie, F. Snowball sampling: A purposeful method of sampling in qualitative research. *Strides in development of medical education* **14** (2017).

[62] Guest, G., Bunce, A. & Johnson, L. How many interviews are enough? an experiment with data saturation and variability. *Field methods* **18**, 59–82 (2006).

[63] Hennink, M. M., Kaiser, B. N. & Marconi, V. C. Code saturation versus meaning saturation: how many interviews are enough? *Qualitative health research* **27**, 591–608 (2017).

[64] Williams, N. & Murray, D. An analysis of the operational costs of trucking: 2020 update (2020).

[65] De Jong, G., Schroten, A., van Essen, H., Otten, M. & Bucci, P. The price sensitivity of road freight transport: A review of elasticities. *Applied transport economics, a management and policy perspective, De Boeck, Antwerp* (2010).

[66] Abdelwahab, W. M. Elasticities of mode choice probabilities and market elasticities of demand: evidence from a simultaneous mode choice/shipment-size freight transport model. *Transportation Research Part E: Logistics and Transportation Review* **34**, 257–266 (1998).

[67] Christidis, P., Leduc, G. *et al.* Longer and heavier vehicles for freight transport. *JRC European Commission* 1–40 (2009).

[68] Dahl, C. A. Measuring global gasoline and diesel price and income elasticities. *Energy Policy* **41**, 2–13 (2012).

[69] Association, A. T. News & Insights. https://www.trucking.org/news-insights (2020). Accessed: 05-April-2021.

[70] Viscelli, S. *The big rig: Trucking and the decline of the American dream* (Univ of California Press, 2016).

[71] Levy, K. E. The contexts of control: Information, power, and truck-driving work. *The Information Society* **31**, 160–174 (2015).

[72] Cheeseman Day, J. & Hait, A. W. Number of Truckers at All-Time High. https://tinyurl.com/3uj4urr9 (2019). Accessed: 6-May-2021.

[73] Thiede, B., Greiman, L., Weiler, S., Beda, S. & Conroy, T. Six charts that illustrate the divide between rural and urban america. *The Conversation* (2017).

[74] Slowik, P. & Sharpe, B. Automation in the long haul: Challenges and opportunities of autonomous heavy-duty trucking in the united states. *International Council on Clean Transportation, Working Paper* (2018).

[75] Endsley, M. R. & Kiris, E. O. The out-of-the-loop performance problem and level of control in automation. *Human factors* **37**, 381–394 (1995).

[76] Naughton, K. Ford's dozing engineers side with Google in full autonomy push. https://tinyurl.com/2h98swvp (2017). Accessed: 05-April-2021.

[77] Volvo. Volvo Cars CEO urges governments and car industry to share safety-related traffic data. https://tinyurl.com/28ezuxds (2017). Accessed: 05-April-2021.

[78] Ayre, J. Google/Waymo Stopped Testing Level 3 Self-Driving Tech After Testers Literally Fell Asleep While Using It, Switched To Full Autonomy. https://tinyurl.com/yhj5cddb (2017). Accessed: 05-April-2021.

[79] Bissell, D., Birtchnell, T., Elliott, A. & Hsu, E. L. Autonomous automobilities: The social impacts of driverless vehicles. *Current Sociology* **68**, 116–134 (2020).

[80] Mladenović, M. N. How should we drive self-driving vehicles? anticipation and collective imagination in planning mobility futures. In *The governance of smart transportation systems*, 103–122 (Springer, 2019).

[81] Association, A. T. News & Insights. https://www.trucking.org/news-insights/truckload-turnover-rises-third-quarter (2020). Accessed: 24-April-2021.

[82] Kaack, L. H., Vaishnav, P., Morgan, M. G., Azevedo, I. L. & Rai, S. Decarbonizing intraregional freight systems with a focus on modal shift. *Environmental Research Letters* **13**, 083001 (2018).

[83] Im, Z. J., Mayer, N., Palier, B. & Rovny, J. The "losers of automation": A reservoir of votes for the radical right? *Research & Politics* **6**, 2053168018822395 (2019).

[84] Viscelli, S. Will robotic trucks be" sweatshops on wheels?". *Issues in Science and Technology* **37**, 79–88 (2020).

[85] Andrew J. Hawkins. Uber pledges to shift to '100 percent' electric vehicles by 2030. https://www.theverge.com/2020/9/8/21427196/uber-promise-100-percent-electric-vehicle-ev-2030 (2020). Accessed: 19-November-2020.

[86] Jing, P., Hu, H., Zhan, F., Chen, Y. & Shi, Y. Agent-based simulation of autonomous vehicles: A systematic literature review. *IEEE Access* **8**, 79089–79103 (2020).

[87] Fagnant, D. J. & Kockelman, K. M. The travel and environmental implications of shared autonomous vehicles, using agent-based model scenarios. *Transportation Research Part C: Emerging Technologies* **40**, 1–13 (2014).

[88] Boesch, P. M., Ciari, F. & Axhausen, K. W. Autonomous vehicle fleet sizes required to serve different levels of demand. *Transportation Research Record* **2542**, 111–119 (2016).

[89] Liu, J., Kockelman, K. M., Boesch, P. M. & Ciari, F. Tracking a system of shared autonomous vehicles across the austin, texas network using agent-based simulation. *Transportation* **44**, 1261–1278 (2017).

[90] Martinez, L. M. & Viegas, J. M. Assessing the impacts of deploying a shared self-driving urban mobility system: An agent-based model applied to the city of lisbon, portugal. *International Journal of Transportation Science and Technology* **6**, 13–27 (2017).

[91] Chen, T. D., Kockelman, K. M. & Hanna, J. P. Operations of a shared, autonomous, electric vehicle fleet: Implications of vehicle & charging infrastructure decisions. *Transportation Research Part A: Policy and Practice* **94**, 243–254 (2016).

[92] Loeb, B., Kockelman, K. M. & Liu, J. Shared autonomous electric vehicle (saev) operations across the austin, texas network with charging infrastructure decisions. *Transportation Research Part C: Emerging Technologies* **89**, 222–233 (2018).

[93] Loeb, B. & Kockelman, K. M. Fleet performance and cost evaluation of a shared autonomous electric vehicle (saev) fleet: A case study for austin, texas. *Transportation Research Part A: Policy and Practice* **121**, 374–385 (2019).

[94] Bauer, G. S., Greenblatt, J. B. & Gerke, B. F. Cost, energy, and environmental impact of automated electric taxi fleets in manhattan. *Environmental science & technology* **52**, 4920–4928 (2018).

[95] Fagnant, D. J. & Kockelman, K. M. Dynamic ride-sharing and fleet sizing for a system of shared autonomous vehicles in austin, texas. *Transportation* **45**, 143–158 (2018).

[96] Lokhandwala, M. & Cai, H. Dynamic ride sharing using traditional taxis and shared autonomous taxis: A case study of nyc. *Transportation Research Part C: Emerging Technologies* **97**, 45–60 (2018).

[97] Farhan, J. & Chen, T. D. Impact of ridesharing on operational efficiency of shared autonomous electric vehicle fleet. Tech. Rep. (2018).

[98] Lu, M., Taiebat, M., Xu, M. & Hsu, S.-C. Multiagent spatial simulation of autonomous taxis for urban commute: Travel economics and environmental impacts. *Journal of Urban Planning and Development* **144**, 04018033 (2018).

[99] Bauer, G. S., Phadke, A., Greenblatt, J. B. & Rajagopal, D. Electrifying urban ridesourcing fleets at no added cost through efficient use of charging infrastructure. *Transportation Research Part C: Emerging Technologies* **105**, 385–404 (2019).

[100] Moniot, M., Borlaug, B., Ge, Y., Wood, E. & Zimbler, J. Electrifying new york city ride-hailing fleets: An examination of the need for public fast charging. *Iscience* **25**, 104171 (2022).

[101] Fitzgerald, R., Reinicke, N. & Moniot, M. The highly integrated vehicle ecosystem (hive): A platform for managing the operations of on-demand vehicle fleets. Tech. Rep., National Renewable Energy Lab.(NREL), Golden, CO (United States) (2021).

[102] NPR. Lyft And Uber Prices Are High. Wait Times Are Long And Drivers Are Scarce. https://www.npr.org/2021/08/07/1025534409/lyft-and-uber-prices-are-high-wait-times-are-long-and-drivers-are-scarce (2021). Accessed: 6-June-2022.

[103] Hall, J. V. & Krueger, A. B. An analysis of the labor market for uber's driver-partners in the united states. *Ilr Review* **71**, 705–732 (2018).

[104] Datseris, G., Vahdati, A. R. & DuBois, T. C. Agents.jl: a performant and feature-full agent-based modeling software of minimal code complexity. *SIMULATION* **0**, 003754972110688 (2022). URL https://doi.org/10.1177/00375497211068820.

[105] Wall Street Journal. Uber, Lyft and the Crucial Challenge for Profitability in 2022. https://www.wsj.com/video/series/george-downs/uber-lyft-and-the-crucial-challenge-for-profitability-in-2022/3979065D-A342-4B05-990C-8447442D2746 (2022). Accessed: 6-June-2022.

[106] Greenblatt, J. B. & Saxena, S. Autonomous taxis could greatly reduce greenhouse-gas emissions of us light-duty vehicles. *Nature Climate Change* **5**, 860–863 (2015).

[107] Bruchon, M. B., Michalek, J. J. & Azevedo, I. L. Effects of air emission externalities on optimal ridesourcing fleet electrification and operations. *Environmental Science & Technology* **55**, 3188–3200 (2021).

[108] Wenig, J., Sodenkamp, M. & Staake, T. Battery versus infrastructure: Tradeoffs between battery capacity and charging infrastructure for plug-in hybrid electric vehicles. *Applied Energy* **255**, 113787 (2019).

[109] Nicholas, M., Slowik, P. & Lutsey, N. Charging infrastructure requirements to support electric ride-hailing in us cities. *Washington DC: Publisher International Council for Clean Transport* (2020).

[110] Moniot, M., Ge, Y. & Wood, E. Estimating fast charging infrastructure requirements to fully electrify ride-hailing fleets across the united states. *IEEE Transactions on Transportation Electrification* **8**, 2177–2190 (2022).

[111] Huang, Y. & Kockelman, K. M. Electric vehicle charging station locations: Elastic demand, station congestion, and network equilibrium. *Transportation Research Part D: Transport and Environment* **78**, 102179 (2020).

[112] Guttenberg, M., Sripad, S., Bills, A. & Viswanathan, V. Incepts: Software for high-fidelity electric vehicle en route state of charge estimation, fleet analysis and charger deployment. *eTransportation* **7**, 100106 (2021).

[113] Yan, X., Liu, X. & Zhao, X. Using machine learning for direct demand modeling of ridesourcing services in chicago. *Journal of Transport Geography* **83**, 102661 (2020).

[114] Ghaffar, A., Mitra, S. & Hyland, M. Modeling determinants of ridesourcing usage: A census tract-level analysis of chicago. *Transportation Research Part C: Emerging Technologies* **119**, 102769 (2020).

[115] Dean, M. D. & Kockelman, K. M. Spatial variation in shared ride-hail trip demand and factors contributing to sharing: Lessons from chicago. *Journal of Transport Geography* **91**, 102944 (2021).

[116] Taiebat, M., Amini, E. & Xu, M. Sharing behavior in ride-hailing trips: A machine learning inference approach. *Transportation Research Part D: Transport and Environment* **103**, 103166 (2022).

[117] Bruchon, M. B. *Analyzing and Optimizing Shared Mobility Fleet Impacts*. Ph.D. thesis, Carnegie Mellon University (2022).

[118] Balding, M., Whinery, T., Leshner, E. & Womerldorff, E. Estimated tnc share of vmt in six us metropolitan regions (revision 1). *Fehr and Peers. https://drive. google. com/file/d/1FIUskVkj9lsAnWJQ6kLhAhNoVLjfFdx3/view* (2019).

[119] Department of Energy. Electric Vehicle Charging Station Locations. https://afdc.energy.gov/fuels/electricity_locations.html#/analyze (2022). Accessed: 08-June-2022.

[120] US Environment Protection Agency. 2022 Nissan Leaf. www.fueleconomy.gov (2022). Accessed: 08-June-2022.

[121] Yuksel, T., Tamayao, M.-A. M., Hendrickson, C., Azevedo, I. M. & Michalek, J. J. Effect of regional grid mix, driving patterns and climate on the comparative carbon footprint of gasoline and plug-in electric vehicles in the united states. *Environmental Research Letters* **11**, 044007 (2016).

[122] Chen, M. K., Rossi, P. E., Chevalier, J. A. & Oehlsen, E. The value of flexible work: Evidence from uber drivers. *Journal of political economy* **127**, 2735–2794 (2019).

[123] Camerer, C., Babcock, L., Loewenstein, G. & Thaler, R. Labor supply of new york city cabdrivers: One day at a time. *The Quarterly Journal of Economics* **112**, 407–441 (1997).

[124] Cheng, M. Can Uber lure back its US drivers without making them full-time employees? https://qz.com/2013562/uber-wants-more-full-time-workers-as-it-faces-a-driver-shortage/ (2021). Accessed: 08-June-2022.

[125] Parrott, J. & Reich, M. A minimum compensation standard for seattle tnc drivers. *Report prepared for the City of Seattle, forthcoming* (2020).

[126] Gagnon, P., Frazier, W., Hale, E. & Cole, W. Long-run marginal co2 emission rates workbooks for 2020 standard scenarios cambium data. Tech. Rep., National Renewable Energy Laboratory-Data (NREL-DATA), Golden, CO (United … (2021).

[127] Argonne National Laboratory. Greenhouse gases, Regulated Emissions, and Energy use in Transportation Model (2019). URL https://greet.es.anl.gov/greet.models.

[128] Group, I. W. *et al.* Technical support document: social cost of carbon, methane, and nitrous oxide interim estimates under executive order 13990. Tech. Rep., Tech. rep., White House. URL https://www. whitehouse. gov/wp-content/uploads … (2021).

[129] Ward, J. W., Michalek, J. J. & Samaras, C. Air pollution, greenhouse gas, and traffic externality benefits and costs of shifting private vehicle travel to ridesourcing services. *Environmental Science & Technology* **55**, 13174–13185 (2021).

[130] Anair, D., Martin, J., Moura, M. & Goldman, J. *Ride-Hailing's Climate Risks: Steering a Growing Industry toward a Clen Transporatation Future* (Union of Concerned Scientists., 2020).

[131] Schaller, B. The new automobility: Lyft, uber and the future of american cities (2018).

[132] Sperling, D. *Three revolutions: Steering automated, shared, and electric vehicles to a better future* (Island Press, 2018).

[133] Bonifacic, I. California makes zero-emission autonomous vehicles mandatory by 2030. https://techcrunch.com/2021/09/24/california-makes-zero-emission-autonomous-vehicles-mandatory-by-2030/?tpcc=ECTW2020 (2021). Accessed: 1-May-2022.