PARAMETRIC ANALYSIS OF TELECOMMUTING EFFECTS ON TRANSPORTATION TAX REVENUES

Mehdi Shokouhian
Paul Schonfeld
Oludare Owolabi
Desmond Amiegbebhor
Zeinab Bandpey
Tao Yang
Fei Wu

Department of Civil Engineering

Morgan State University & University of Maryland (UMD)

Principal Investigator: Mehdi Shokouhian

Contract No.

Contract Amount

February 28, 2023

Final Report

This document is prepared for the

DC Department of Transportation (DDOT)

Copyright / Statement of Originality

We declare that this report describes original work carried out by us. The report gives proper acknowledgement to all contributors other than the authors named above. To the best of our knowledge, the work described in the report is original and no part of it has been copied or taken from other sources without necessary permissions or citations.

Abstract

This study explores the effects of increases in online activities (such as telecommuting and online shopping) as well as shifts toward electric vehicles on transportation revenues for the District of Columbia (DC). The analysis is based on (1) a review of the literature and state of practice regarding online activities, trends in vehicle fleet mix, experience with various forms of revenue collection, methods for estimating trends in online activities and vehicle characteristics, and methods for estimating transportation revenues; (2) a survey of DC transportation users designed to obtain information that is unavailable from existing sources; (3) the adaptation of demand and travel behavior models for quantitatively predicting online activities; (4) a model developed for estimating transportation revenues from various sources, as well as other measures of effectiveness; and (5) parametric studies of the effects of various factors, individually and in combinations, on expected transportation revenues. The results of this study should support decision-making by the DC Government regarding taxation mechanisms and other policies that may be applied to fund its transportation needs.

Acknowledgement, Dedication and Preface

This research was supported by the DC Department of Transportation, DDOT. We thank our colleagues from DDOT who provided support, insight and expertise that greatly assisted the research.

Table of Contents

List of Tables/Figures/ Illustrations

Introduction

The District of Columbia (DC), like other major jurisdictions in the United States, is responsible for constructing, maintaining, and operating an extensive multi-modal transportation system, at considerable expense. Besides subsidies from the Federal Government to which DC transportation users contribute substantially, the DC government funds its transportation expenses largely with taxes collected from motorists. It has a Highway Trust Fund, which is funded by the District's tax on gasoline sales, i.e., the "gas tax". The funds are used to provide the match funding for federal highway funds used to build and maintain the District's transportation system. The District of Columbia has 1,057 miles of federal and local roadways with 21.6% in poor and 25.3% in fair conditions, and 244 bridges of which 12% are reported as structurally deficient. The District's Highway Trust Fund is an integral source of funding to maintain this infrastructure. DC's gas tax revenues are significantly affected by some current trends which also raise concerns about future revenues. The major ones among these trends are (1) the increasing adoption of electric and hybrid vehicles and (2) the decrease in vehicle miles travelled on roads due to the increasing substitution of online activities for actual trips. The transition to online activities includes the substitution of telecommuting for work and school trips, the substitution of eCommerce and home deliveries for shopping trips, and the increasing use of electronic communications instead of social and entertainment trips. These trends have started decades ago, but have been greatly accelerated by the Covid-19 pandemic. There are significant concerns that as people learn and get used to relying on such substitutes for travel, serious declines may persist in revenues from gas taxes as well as in the use of mass transit vehicles and in property values (and hence taxes) of downtown office buildings. There may also be countervailing trends resulting from advances in vehicle automation, which may increase people's willingness to make long trips.

Together, these trends have created considerable uncertainties about (1) how sufficient revenues may be generated in the future to fund the transportation system and (2) how these trends may affect the amounts of revenue needed. This study focuses on the first of those questions and explores how much revenue could be generated in the future through different funding mechanisms and under a wide range of circumstances.

The study is exploring the revenue generation alternatives through the following steps:

1. A literature review examines the important trends affecting transportation revenues and the tax mechanisms which may be useful in generating revenues. To accomplish that the study assesses the important trends affecting revenues and the methods through which those effects may be forecast.

2. Current models for estimating transportation tax revenues are analyzed.
3. The effects of telecommuting and some other online activities on travel characteristics and transportation tax revenues are investigated.
4. A survey was conducted to help understand the options and preferences of employers and employees regarding online activities. It will be designed to provide information that is not available from existing sources. This survey is expected to be conducted after we ascertain the information solicited through the survey is not available from other sources.
5. Models have been developed for estimating demand characteristics, travelers' choices and potential revenues obtainable through different kinds of transportation taxes. These models will be combined into one integrated model that will estimate the revenues obtained from various subgroups and revenue-generating mechanisms.
6. After the above integrated model was finalized, it was used to conduct extensive sensitivity analyses regarding input parameters, individually and in combinations. Scenarios were specified to represent a range of possible future circumstances through combinations of assumptions and input parameters. The effects of these various scenarios on demand characteristics and tax revenues were then explored.
7. Finally, a report was prepared documenting the study's findings from the literature and analysis of survey results. It will also document the methods used for estimating future demand characteristics and transportation revenues, as well as the results obtained with those methods. The report offers recommendations regarding the effectiveness of various tax mechanisms and other policies in satisfying future transportation revenue needs.

The results of this study should support decision-making by the DC Government regarding taxation mechanisms and other policies that may be applied to fund its transportation needs.

1. LITERATURE REVIEW

1.1. Transport infrastructure and decline in sources of funding

America's transportation infrastructures are largely funded by state, local, and federal motor fuel taxes. States have levied taxes since 1919, and by 1932, when the federal tax was introduced, the then-48 states and the District of Columbia were collecting taxes on motor fuel (Liz 2015). Today, a combination of increased telecommuting, growth in the sales of electric vehicles (EVs), and inflation has raised concerns about the sustainability of these taxes as a funding mechanism for transportation infrastructure.

States are also exploring other revenue sources for funding road investment, including mileage-based user fees. With continued improvements in vehicle fuel efficiency and the popularity of hybrid and electric vehicles, mileage-based user fees could present an opportunity for a long-term funding alternative to the motor fuels taxes. However, while legislative and voter action has allowed some states to maintain or increase local sources of roadway funding, federal funding remains a significant portion of overall road funding. Put another way, federal partnership for roadway infrastructure is still required to maintain and modernize the system. Additionally, the COVID-19 pandemic has led to a sharp decline in vehicle miles travelled and therefore gas tax receipts in 2020, and the full impact of this revenue loss for state transportation budgets, could be as much as $37 billion over 2020 and 2021 (USDOT, US Vehicles-Miles" Bureau of Transportation Statistics. 2021, TRIP, "Bumpy Road Ahead: America's Roughest Rides and Strategies to Make Our Roads Smoother,". 2018, TRIP, "Restoring the Interstate Highway System," Texas A&M Transportation Institute, 2019 Urban Mobility Report 2019, TRIP, "Key Facts About the U.S. Surface Transportation System," A National Transportation Research Non-Profit 2020, F. USDOT, "Status of the Nation's Highways, Bridges, and Transit Conditions and Performance Report," 2020)

America's roads are critical for moving ever increasing volumes of people and goods. However, these vital lifelines are frequently underfunded, and over 40% of the system is now in poor or mediocre condition. As the backlog of rehabilitation needs grow, motorists are forced to pay over $1,000 every year per motorist in wasted time and fuel. Additionally, over 36,000 people are still dying on the nation's roads every year, and the number of pedestrian fatalities is on the rise. Federal, state, and local governments will need to prioritize strategic investments dedicated to improving and preserving roadway conditions that increase public safety on the system we have in place, as well as plan for the roadways of the future, which will need to account for connected and autonomous vehicles (ASCE 2021).

The Washington Metropolitan Area Transit Authority (WMATA or Metro) and the District Department of Transportation (DDOT) provide public transit services in the D.C. metro region over a network of heavy rail transit and bus components supported by circulator bus, paratransit and streetcar elements. WMATA has been challenged by an aging infrastructure with increasing State-of-Good-Repair (SOGR) and safety-related needs during a period of steadily declining ridership. Infrastructure and vehicle investments have been significant in recent years. However, additional funding and new sources are needed to address the $1.8 billion required to comply with safety and security directives and for upgrades or replacements, all while a $6.6 billion SOGR backlog persists. In addition, WMATA needs a robust plan to infuse innovation into its system and develop innovative approaches focused on diversifying operations and increasing ridership (DDOT 2017, Metro Rail Fleet Plan, 2010, WMATA, Metrobus Fleet Management Plan, Office of Bus Planning, Final Report Version 2.1 2017, MCNA 2008, MWCG, Metropolitan Washington Council of

Governmenets, State of the Region Infrastructure Report. 2015, MWCG, WMATA's Funding Needs, The Magnitude and the Effect, Updated to Reflect WMATA's Proposed FY 2018 Budget, Presentation to the DC Building Industry Association Microsoft, (2012). "Ordinary Or Extraordi 2017).

The District of Columbia is home to more than 1,150 miles of roads, of which less than 10% are rated as "poor" according to the Pavement Condition Index, a noteworthy improvement from five years ago. With a 43-minute average commute (pre COVID-19), the third highest in the country, D.C. workers spent 60% more time commuting than the national average of 27-minutes. This translates to an annual cost per worker of $2,015 spent sitting in traffic. The COVID-19 pandemic significantly reduced commuter traffic—with 55% employees working full-time remotely—though with projected population and job growth, levels of congestion are projected to return to what they were before the pandemic. D.C. recently raised its gas tax by 10 cents, a step towards generating local funds for preservation and maintenance of the surface transportation network. However, despite a District-wide initiative to reduce pedestrian deaths to zero, known as Vision Zero, pedestrian deaths are on the rise, further indicating that D.C. must find ways to increase its investment in roads for congestion relief, and to improve the safety of drivers and pedestrians alike (Kelly 2019, Eliza 2020, USBTS 2019).

The Delaware Department of Transport set up a Task Force to examine the required needs of the Transportation Trust Fund for the maintenance of the entire transportation program for the period Fiscal Year 2012 –2023 and concluded that total spending for transportation expenses over the period can reasonably be estimated to total $12.4 billion and that current revenue streams will support only 70% of those needs (DDT, Report on the condition, planning and revenues options for the support of the transportation trust fund" Transportation Trust Fund Task Force 2011). This situation has been worsened with the continuous shortfall in tax revenues as a result of the outbreak of COVID-19 pandemic resulting in increased need for virtual services and telecommuting as well as the growth in the number of EVs.

There are more than 617,000 bridges across the United States. Currently, 42% of all bridges are at least 50 years old, and 46,154, or 7.5% of the nation's bridges, are considered structurally deficient, meaning they are in "poor" condition. Unfortunately, 178 million1 trips are taken across these structurally deficient bridges every day. Estimates show that there is a need to increase spending on bridge rehabilitation from $14.4 billion annually to $22.7 billion annually, i.e., by 58%, to improve their condition. At the current rate of investment, it will take until 2071 to make all of the repairs that are currently necessary, and the additional deterioration over the next 50 years will become overwhelming (F. USDOT, MAP-21 Comprehensive Truck Size and Weight Limits Study, Freight Management and Operations, 21st Century Operations Using 21st Century

Technologies 2016, ARBA 2020, USDOT, U.S. Department of Transportation, National Bridge Inventory Management and Preservation 2020).

D.C. has 244 highway bridges, 208 of which are owned by the D.C. Department of Transportation (DDOT) and 36 of which are owned by the National Park Service (NPS). The average age of these bridges is 62 years which is well over the national average of 44 years and approximately 30% of them will need to be rehabilitated in the next 10 years. Though DDOT and NPS have made significant strides in replacing or rehabilitating old bridges, about three percent of bridge conditions are still classified as poor. Even after the rehabilitation of the Arlington Memorial Bridge, more than 200,000 trips will be taken every day over bridges in poor condition (Washington 2021).

1.2. The concept of telecommuting

The telecommuting term appeared in the 1970s to describe work-related substitutions of telecommunication and related information technologies for travel (Caves 2004). The concept of telecommuting grew in the '80s, when IBM installed "remote terminals" in several employees' homes in Europe, as some companies began officially experimenting with telecommuting and work from home. Telecommuting refers more specifically to work undertaken at a location that reduces commuting time. These locations of remote terminal installations can be inside the home. Telecommuting is also regarded as a sustainable travel-demand management strategy. Telecommuting yields environmental advantages by reducing greenhouse gas emissions and traffic congestion (Davis 2021).In the 1990s, telecommuting became the subject of popular culture and attention in Europe and US. In 1995, the motto that "work is something you do, not something you travel to" was coined (Woody 1995). Variations of this motto include: "Work is something we DO, not a place that we GO (Microsoft 2012) and "Work is what we do, not where we are (GSA 2012). Telecommuting has been adopted by a range of businesses, governments and not-for-profit organizations. Organizations may use telecommuting to reduce costs (telecommuting employees do not require an office or cubicle, a space which needs to be rented or purchased, and incurs additional costs such as lighting, climate control, etc.). Some organizations adopt telecommuting to improve workers' quality of life, as teleworking typically reduces commuting time and time stuck in traffic jams. Along with this, teleworking may make it easier for workers to balance their work responsibilities with their personal life and family roles (e.g., caring for children or elderly parents). Some organizations adopt teleworking for environmental reasons, as telework can reduce congestion and air pollution, with fewer cars on the roads.

Teleworkers in the 21st century often use mobile telecommunications technology such as a Wi-Fi-equipped laptop or smartphones to work from coffeeshops; others may use a desktop computer and a landline phone at their home. According to a Reuters poll, approximately "one in five workers around the globe, particularly employees in the Middle East, Latin America and Asia, telecommute frequently and nearly 10 percent work from home every day (Patricia 2012). In the 2000s, annual leave or vacation in some organizations was seen as absence from the workplace rather than ceasing work, and some office employees used telework to continue to check work emails while on vacation.

The practice became much more mainstream during the COVID-19 pandemic, when millions of workers were forced to start remote working for the first time (San 2020). Although the concepts of "telecommuting" and "telework" are closely related, there is a difference between the two. All types of technology-assisted work conducted outside a centrally located workspace (including work undertaken in the home, outside calls, etc.) are regarded as telework. Telecommuters often maintain a traditional office and usually work from an alternative work site from 1 to 3 days a week (Hill, et al. 1998). Telecommuting refers more specifically to work undertaken at a location that reduces commuting time. These locations can be inside the home or at some other remote workplace, which is facilitated through a broadband connection, computer or phone lines (Ellison 2004) or any other electronic media used to interact and communicate (Gajendran and Harrison 2007). As a broader concept than telecommuting, telework has four dimensions in its definitional framework: work location, that can be anywhere outside a centralized organizational workplace; usage of (information and communication technologies) as technical support for telework; time distribution, referring to the amount of time replaced in the traditional workplace; and the diversity of employment relationships between employer and employee, ranging from contract work to traditional full-time employment (Garret and T 2007).

A person who telecommutes is known as a "telecommuter", "teleworker", and sometimes as a "home-sourced", or "work-at-home" employee. A telecommuter is also called a "telecommuting specialist", as a designation and in a professional context. Many telecommuters work from home, while others, sometimes called "nomadic workers" work at coffee shops or other locations.

A study by (USDOE 1994) estimated that telecommuting in the 339 largest US cities (accounting for two-thirds of the US population) could eliminate the need for 7300±11 200 lane-miles of freeways and major arterials for an (undiscounted) cost savings of $13±20 billion. Another study (USDOT, "Transportation Implications of Telecommuting" US DOT, Washington, DC 1993)estimated that, nationwide, telecommuting could result in 408,815 lives saved and 58,850 accidents avoided by the year 2002 due to reducing travel. The same study estimated travel time savings by telecommuters at 826 million to 1.7 billion hours in 2002.

The empirical literature on telecommuting has grown significantly over the decade of the 1990s and into the 2000s. With the availability of more and better data for analyses, published articles have proliferated. Early studies relied on relatively small samples of telecommuters and individual places of employment that had adopted organized telecommuting programs. Often these were state government agencies. Data collected often included information on only telecommuting frequency and commute VMT (Margaret and Elena 2014).

To look at factors that influence the likelihood of telecommuting, (Yen, Mahmassani and Herman, Employer attitudes and stated preferences toward telecommuting: An exploratory analysis. Transp. Res. Rec. 1463, 15–25. 1994) used a stated-preference approach, in which survey respondents are asked about their preferences and what they would do in certain circumstances, not what they actually do. They surveyed 545 employees in selected organizations in Austin, Houston, and Dallas. The employees were presented with four alternatives: (1) not working from home at all, (2) possibly working from home, (3) working from home several days per week, and (4) working from home every day. The employees were also given seven program scenarios, which included 5% and 10% increases and decreases in salary and some increases in costs due to equipment purchases. The estimation results show, not surprisingly, that if the individual is given a salary increase when he telecommutes, he is more likely to telecommute, whereas a decrease in salary makes it less likely he will telecommute. If there are additional costs to working from home, the individual is less likely to telecommute. Employees with children under 16 at home are more likely to say that they would telecommute, as were those with personal computers at home and thos with higher computer proficiency levels. The greater the distance from home to workplace, the more likely the employee is to say that she will telecommute. Among job characteristics, the authors find that the more face-to-face communication with coworkers the employee says he needs, the lower the probability of telecommuting.

The implications of telecommuting and other combined factors in reducing gas tax revenues by reducing the need to travel and resulting fuel consumption have in recent years captured the attention of public planners and policymakers. The application of telecommuting offers particular appeal since it addresses a number of other policy issues such as the `family friendly' workplace (Gordon, ''Clinton calls for `family friendly' work arrangements'', Telecommuting Review, 13(8), p. 16. 1996) and employment opportunities for mobility limited sectors of the labor force (Hesse 1995).

While it has been largely accepted that telecommuting reduces commute travel, there are acknowledged viewpoints about its overall impact on fuel tax revenues.

(Mokhtarian and Salomon, Modeling the desire to telecommute: The importance of attitudinal factors in behavioral models. Transp. Res. Part A: Policy Practice 31, 35–50. 1997) lists examples of policy statements supporting telecommuting from the state governments of California, Washington, Florida and Virginia, as well as the federal government (Bush administration). Since that time, similar laws, resolutions, and proclamations have been adopted by the states of Arizona (Gordon, ''Clinton calls for `family friendly' work arrangements'', Telecommuting Review, 13(8), p. 16. 1996), New Jersey (Gordon, ''Trip-reduction legislation signed into law in New Jersey, Telecommuting Review: The Gordon Report'', 1 August, p. 11. 1992), Georgia (Gordon, ''Georgia Legislature passes commuter efficiency resolution, Telecommuting Review'' The Gordon Report, May, p. 3. 1993) and Minnesota (proclamation by Governor Carlson declaring the week of 13 May 1991 to be `Telecommuting Week'), among other activities at local, state and federal levels. At the federal level, the Clinton administration released the President's Management Council National Telecommuting Initiative Action Plan (November 1995), which called for an increase in the number of federal government telecommuters from about 4000 to 60 000 by the end of year 1998 (about 3 per cent of the civilian federal workforce).

Many studies have found results supporting the hypothesis that telecommuting can reduce daily trip rates, travel distance, and VMT. Researchers have distinguished between telecommuting penetration (the percentage of workers who telecommute), and the number of telecommuting occasions (the number of days on which an employee works entirely at home). Both statistics can be useful, but it is the latter that is critical for assessing the effects, including on VMT, congestion, and emissions, of telecommuting (Walls, Safirova and Jiang 2007).

There have been many studies researching the drivers and constraints of telecommuting. They all have some similar findings with many stating that land use patterns, internet infrastructure, socio-demographic characteristics, access to high-speed internet, the presence of children at home, public transport access and cost of travel and fuel can influence rates of telecommuting (Caulfield 2015, Choo, Mokhtarian and Salomon, . Does telecommuting reduce vehicle-miles traveled? An aggregate time series analysis for the US. Transportation 32,37–64. 2005).

A widely discussed benefit of telecommuting is the reduction in travel time, cost, congestion and emissions. These have had varying levels of success, depending on the country in which the research has been implemented. (Choo, Mokhtarian and Salomon, . Does telecommuting reduce vehicle-miles traveled? An aggregate time series analysis for the US. Transportation 32,37–64. 2005) argued that more people choose to telecommute in opposition of fuel taxes and congestion charges, and that not only will telecommuting reduce the number of work-related trips, but also non- work-related trips for commuters and their immediate family members.

Financial experts anticipated a reduction in transportation funding due to teleworking, online shopping, and home delivery and the pandemic has only exacerbated the collection of revenue from transportation's traditional funding streams. TRB's recent report ''Renewing the National Commitment to the Interstate Highway System: A Foundation for the Future'' calls on the Congress to identify and adopt new transportation funding mechanisms that are equitable and efficient, that do not unduly impose the burden of payment on future generations or on less financially equipped groups, and do not disadvantage or divert resources from other highways and modes of passenger and freight transportation (NASEM 2020).

Telecommuters partially or entirely replace their out-of-home work activities by working at home or at locations close to home. In general, telecommuting offers more flexibility to workers by relaxing the temporal and spatial work-related constraints.

Over the past decades, several overviews of the impacts of telecommunications on travel have appeared, both conceptual (Salomon 1986, Mokhtarian, Handy and Salomon, Methodological issues in the estimation of the travel, energy, and air quality impacts of telecommuting. Transp. Res. Part A: Policy Pract. 29, 283–302 1995) and empirical (Nilles 1989, Mokhtarian, Handy and Salomon, Methodological issues in the estimation of the travel, energy, and air quality impacts of telecommuting. Transp. Res. Part A: Policy Pract. 29, 283–302 1995). Most of the empirical research has focused on telecommuting, probably because it has been feasible for longer than most other `tele-applications' (such as videoconferencing or on-line shopping), it has the appealing benefits, and the prospect of eliminating or reducing the peak-period commute trip is especially attractive.

To demonstrate the impact of Covid-19 on travel, charts were developed to illustrate how the COVID-19 pandemic was impacting travel in the Metropolitan Washington Region.

The charts were prepared by COG/TPB staff using Continuous Count Station (CCS) data collected by the District of Columbia, Maryland, and Virginia as well as enplanement data provided by the Metropolitan Washington Airports Authority (MWAA) and BWI Thurgood Marshall Airport. The intention was to update these charts on a regular basis as data become available (Transportation Planning 2021). The underlying assumption was that people were getting accustomed to working from home.

A key finding from the study conducted by Pew Research in October 2020 (Parker and Horowitz, How the coronavirus outbreak has–and hasn't–changed the way Americans work. Pew Res Center. Tavares, A.I., 2017. Telework and health effects review. Int. J. Healthcare 3, 30. 2020) was that workers were highly divided: only 54% of working adults would like to work from home once the pandemic is over. This finding is significant; while several studies (Tavares, 2017, Ollo-Lopez ́ et al., 2020) have shown positive impact of the option to telework and of actual telework, the experience from the pandemic has been mixed for many.

Thus, the extent of continued future adoption of telework when it is an available option remains an open question for employers and policy makers in a post-pandemic world. On the positive side of the argument, we note that the resources that corporations have spent during the pandemic to make teleworking easier, increased schedule flexibility, and inclusion aspects of telework may permanently change the way Americans expect to work, and this may lead to maintaining high levels of telecommuting (Igeltjørn and Habib 2020, Bjursell, Bergmo-Prvulovic and Hedegaard 2021). On the other hand, the current level of adoption may not be sustained in the wake of growing evidence related to decline in innovation and productivity (Miglioretti, et al. 2021, Song and Gao 2020) and lack of clearly defined boundaries between work and private life (Lewis 2017, Pluut and Wonders 2020). This is further complicated by the fact that the pandemic forced organizations to suddenly adopt remote work, sometimes without providing employees with the necessary skills and support to thrive in the remote work environment (Errichiello and Pianese 2021).

While we note that employer strategies will play a major role in defining the future forms and adoption of telework, employee preferences and constraints, such as access to appropriate technology or environment to work from home, are also going to be extremely important factors. Overall, there is consensus that different remote work models will persist and that hybrid forms of work will be sustained post COVID-19 pandemic (Gurchiek 2021). Yet, there is a need for further research to understand employee perceptions, barriers and assets related to remote work, as well as the variation among different employee groups. The resulting behavioral insight will be an important input to establishing the forms and strategies to maintain productivity, worker well-being and company culture in a remote work world.

The broad and durable nature of telework adoption during the pandemic across sectors and user-groups presents a rare and unique opportunity to study telework. Most studies prior to the pandemic treated teleworking as a choice, part of an intentional telework program from the employer's end. Instead, analysis of remote work in the COVID-19 era needs to account for the fact that the pandemic broadly forced employers and workers to adopt telework for an extended period except for individuals for whom onsite presence was essential.

In past research, telework has been considered as a means to reduce congestion and the environmental impact of the transportation sector for several decades (Lari 2012, Matthews and Williams 2005, Irwin 2004, Larson and Zhao 2017, Gareis and Kordey 1999, Mokhtarian, Handy and Salomon, Methodological issues in the estimation of the travel, energy, and air quality impacts of telecommuting. Transp. Res. Part A: Policy Pract. 29, 283–302 1995, Choo, Mokhtarian and Salomon, Does telecommuting reduce vehicle-miles traveled? An aggregate time series analysis for the US. Transportation 32,37–64. 2005). Employee telework adoption has been tied to schedule flexibility (Shabanpour, et al. 2018), worker age and educational attainment, (Noonan and Glass 2012), and interaction with the employer's expectations (Brewer and Hensher 2000). In terms of attitudes, telework adoption preferences are linked to both constraints (family effects, commuting, job suitability) as well as opportunities (interaction with co-workers) (Yen, Mahmassani and Herman, Employer attitudes and stated preferences toward telecommuting: An exploratory analysis. Transp. Res. Rec. 1463, 15–25. 1994, Yen and Mahmassani, Telecommuting adoption: Conceptual framework and model estimation. Transp. Res. Rec. 1606, 95–102. 1997, Mokhtarian and Salomon, Modeling the desire to telecommute: The importance of attitudinal factors in behavioral models. Transp. Res. Part A: Policy Practice 31, 35–50. 1997, Elldér 2020). A comprehensive understanding of the long-term viability of remote work and related spatially and temporally flexible work arrangements is still taking shape (Nayak and Pandit 2021, Salon, et al. 2021), and many of the earlier findings may need to be revisited in this new context. For example, earlier research suggests that attitudes may be more consistently important than sociodemographic status like presence of children (Mokhtarian and Salomon, Modeling the desire to telecommute: The importance of attitudinal factors in behavioral models. Transp. Res. Part A: Policy Practice 31, 35–50. 1997). Among the unique features shaping the COVID-19 telework situation is the frequent occurrence of multiple members of the same household teleworking simultaneously, including children attending school online. Overlapping telework arrangements potentially impose resource, time, and space restriction on individuals and increased work-life conflicts.

In the United States, seven percent of workers worked from a remote location (home or other) before the pandemic. During the height of the pandemic in 2020 and a good part of 2021, this percentage surged to 30-40 percent, implying that everybody who could telework did so (KPMG 2022) among others, reports that 37 percent of all US jobs can be performed remotely). As vaccination rates increased and the worst of the pandemic began to fade in late 2021, workers began to make their way back to the office, resulting in about 25-35 percent of workers working from a remote location (these percentages are largely derived from various polls including: (Gallup 2020, Parker, Horowitz and Minkin, . COVID-19 Pandemic Continues To Reshape Work in America. Pew Research Center. Available at: https://www.pewresearch.org/socialtrends/2022/02/16/covid-19-pandemic-continues-to-reshape-work-in-america/. Acce 2022, KPMG 2022).

There is a large body of research dedicated to the study of telecommuting adoption and telecommuting frequency (Singh, et al. 2013, Zhang, et al. 2020, Astroza, et al. 2020, Heiden, et al. 2021, Nguyen 2021, Danalet, Justen and Mathys 2021, Mohammadi, et al. 2022). In general, the body of research has shown that telecommuting is adopted by higher income, higher educated, technology savvy, and younger workers in urban contexts. There is also an extensive body of literature that has explored the impacts of telecommuting on activity-travel demand, with an interesting mix of findings.

While some researchers have documented a clear inverse (substitution) relationship between telecommuting and amount of travel (Lachapelle, Tanguay and Neumark-Gaudet 2018) others have found a more complementary relationship between telework and travel demand – suggesting that the elimination of the commute results in discretionary time that engenders additional non-work travel (Moeckel 2017, Zhu, et al. 2018, Ollo-Lopez, Goni-Legaz and Erro-Garcés 2020, Caldarola and Sorrell 2022). In the wake of the pandemic, researchers have explored who is teleworking (Nguyen 2021, Danalet, Justen and Mathys 2021, Appel-Meulenbroek, et al. 2022, Asmussen, et al. 2022), how much they are teleworking (Zhang, et al. 2020, Heiden, et al. 2021, Mohammadi, et al. 2022), and the extent to which teleworking has affected individual/household travel characteristics (Zhu, et al. 2018, Ollo-Lopez, Goni-Legaz and Erro-Garcés 2020, Caldarola and Sorrell 2022).

There is also a fairly significant body of literature on the implications of teleworking for labor productivity (Martin, Hauret and Fuhrer 2022, Mohammadi, et al. 2022), worker interactions and wellbeing (Hoffman 2021, Nguyen 2021), and vibrancy of central cities (Adobati and Debernardi 2022). Teleworking continues to be of great interest to the profession due to its potentially transformative implications for mode use (particularly transit), the future of employment centers and the small businesses that depend on them, and the spatial and temporal characteristics of travel demand. In particular, travel demand forecasting models will need to be substantially updated to reflect remote work's adoption, frequency, and location, as the trajectory of human behaviors, choices, and preferences appear to have been forever altered by the pandemic effects (leading to a human adaptation process that engendered the adoption of new habits and routines). Given the importance, rapidly evolving nature, and impacts of this behavioral phenomenon (i.e., telework and its various facets), and the multitude of dimensions that characterize this phenomenon, it is critical for the profession to engage in a continuous stream of telework-related research to understand its evolving nature and incorporate the latest insights into transportation demand forecasting models.

1.3. Electric Vehicles

While traditional motor fuel taxes have been gradually and negatively affected, there is a need for sustainable complete or partial alternatives. One of the biggest threats against the gas tax revenue is electric vehicles (EVs) as these cars do not contribute to the state highway funds or the federal highway funds. There were less than two million EVs in the U.S. out of a vehicle fleet of more than 268 million vehicles, which means missed revenue from EVs is steadily growing (EV 2020). However, one forecast suggests EVs may comprise 17.5 percent of U.S auto sales by 2028 (EV 2020) with 50% of all cars in the US ordered to be Electric Vehicles by 2030 in the Biden Infrastructure Bill. This would further reduce gas tax revenue, but a state like California, where EVs and hybrids already make up almost 8 percent of auto sales, may be impacted sooner.

There has sometimes been resistance to taxes on EVs on the grounds that they are environmentally preferable to traditional gas-powered vehicles and thus create fewer externalities. While this is true, since electric vehicles generate less net carbon emissions, it does not account for the single largest function of the current motor fuel tax regime (funding roads) or the greatest externality associated with driving cars (wear-and- tear). However, comparing the environmental benefits of electric vehicles and revenue generated from fuel tax is still subject to further investigation.

In the District of Columbia (2021), most energy used in electric vehicles comes from natural gas, biomass, and solar energy. The exhaust emissions of EVs and PHEVs that only rely on electricity to operate are zero. Still, emissions are produced where electricity is generated from hydrocarbon and biomass sources.

The (VDM 2021) in DC charges $36/year for Clean Fuel/Electric vehicle (Hybrid) at the first registration. The renewal fee for the second year is different.

The (USEIA 2020) estimated that gas consumption will decline by 19% through 2050. The most influential factor for gas tax revenue will be the penetration of electric vehicles. The increase in fuel economy standards will also cause a decline in vehicle gasoline consumption. According to data from the (ANL 2021), as of March 2021, 75,959 hybrid electric vehicles (HEVs) were sold in the United States as well as 33,370 battery electric vehicles (BEVs) and 12,687 plug-in hybrid electric vehicles (PHEVs). Comparing the number of hybrid electric vehicles sold in the U.S. in March 2021 with July 2020, the increase is as high as 70.1%.

1.4. Vehicle Miles Traveled

DC raised its gas tax in July 2020, increasing it by 10 cents to 23.5 cents/gallon. This is above Virginia's 21.95 cents/gallon gas tax and significantly below Maryland's 36.7 cents/gallon tax. This user fee is in addition to the federal gas tax of 18.4 cents/gallon, which has not been increased since 1993. DC has the 36th highest gas tax in the country. All revenue from DC's gas tax goes toward the Highway Trust Fund. From the foregoing, revenue from the federal motor fuel tax will not fund projected spending at the current tax rate, so the only options for lawmakers are to either appropriate general fund money or increase taxes. According to (CBO, Congressional Budget Office, "Issues and options for a tax on vehicle miles traveled by commercial trucks," 10-23 2019) estimates, the Highway Trust Fund will run out of money by the end of 2021 and the deficit is projected to be almost $70 billion over the first years after the FAST Act funding expires (CBO, Congressional Budget Office, Highway Trust Fund Accounts—CBO's Baseline as of March 6, 2020 2020).

Given the challenges facing the motor fuel tax, one solution, long supported by many economists, is to fund highways by taxing vehicle miles traveled. A vehicle miles traveled tax, also frequently referred to as a VMT tax, VMT fee, mileage-based fee, or road user charge, is a policy of charging motorists based on how many miles they have traveled. It has been proposed in various states in the United States, including Illinois, which are currently following through with implementing this tax, and elsewhere as an infrastructure funding mechanism to replace, or supplement the fuel tax, which has been generating billions less in revenue each year due to increasingly fuel-efficient vehicles (Atkinson 2009). Rather than using taxes on cars or motor fuel as a proxy for transportation, a tax levied directly on the miles gets much closer to capturing the externalities and to approximating the road maintenance cost of each driver. A tax on vehicle miles driven would provide a more direct link to the cost of highway use but, unlike an increase in the tax on motor fuels, would be difficult to implement, requiring new tolls or electronic motoring of vehicles. An advantage of a vehicle mileage tax is that it could be adjusted to reflect the additional costs of congestion by increasing tolls or the tax rate in certain locations and at certain times of the day (TPCBB 2020).

Some motorists are concerned that VMT charging could be an invasion of their privacy, as location information is utilized (Badger 2011). They view the program as "Big Brother" or a "Nanny" state. As any data collection system poses a risk to private information of users, VMT pilot programs across the country have explored various options to protect the privacy of participants.

Oregon's 2012 VMT fee pilot study offered five plans, each with a different technology option and payment method depending on the drivers' privacy preferences (Jaffe 2013). Drivers had the choice to report miles using a smartphone, a global positioning system (GPS) device, or a simple reporting device with no GPS technology; or, they could opt out of using technology altogether by paying a flat rate in lieu of a per-mile fee (ODT, Oregon Department of Transportation "Oregon's Road Usage Charge Program" 2014)(Oregon's Road Usage Charge Program 2014). But even those drivers who chose to report their miles using a smartphone or a GPS were not releasing their exact location coordinates and times of travel. These on-board units were programmed to contain just enough intelligence and knowledge of map boundaries to accumulate and transfer to the billing entity the miles per region or zone, as opposed to exact location (ODT, Oregon Department of Transportation "Legislative Report, Road Usage Charge Pilot Program Preliminary Findings" Oregon Department of Transportation. 2014). Likewise, the on-board units were only programmed to aggregate travel during particular periods, as opposed to during exact times (ODT, Oregon Department of Transportation "Legislative Report, Road Usage Charge Pilot Program Preliminary Findings" Oregon Department of Transportation. 2014). Negotiations with the American Civil Liberties Union shaped the privacy provisions of Oregon's recent 2013 legislation, which set up the 2015 VMT program. Section 9 of the bill limits who has access to the data and requires those who have access—including private sector vendors—to protect it (ODT, Oregon Department of Transportation, "Legislative Report, Road Usage Charge Pilot Program Preliminary Findings" 2014). Furthermore, the data are destroyed 30 days after they are required for payment processing or dispute resolution.

To protect privacy in its VMT study, the University of Iowa pilot study used an on-board "smart card" data recording system that separated the driver's personal information from the vehicle miles recorded (Forkenbrock and Kuhl 2002). To accomplish this, researchers developed a two-stage data entry system whereby participants upload basic miles traveled, but then must separately log into a different system to upload more extensive personal data. The miles driven are thus unlinked from other personal data to insure anonymity.

The (PSRC 2014) encountered similar privacy concerns as a result of its 2002 road tolling demand response study, which monitored participants' mileage on certain types of roadways in a similar manner to the above VMT pilot programs. During the study, an on-board meter used a GPS receiver to match the vehicle's location to a map of the toll-road network embedded in the meter. The meter stored location and toll information, and periodically communicated it to a central computer using cellular wireless communications (PSRC 2014)

The Regional Council suggested that future tolling programs (and by extension, VMT programs) could better protect participants' privacy through a choice of two different mileage recording approaches. Participants could either enroll in the "thin client" or "thick client" operating approaches. The thin client would use an on-board toll meter where all raw data was transferred to the tolling office. In the office, the data would be processed, and the road segments would be recognized and matched with toll rates. In the thick client approach, the tolling process would take place in the on- board unit. After the road section was recognized, the toll rate would be processed in the on-board unit according to the type of the road, time period, and vehicle class. The road information could then be sent to the tolling office in aggregate form. Under this approach, specific road details would never be stored in the toll system office. If the fee was calculated in the on-board unit, it would also be possible to integrate a card slot into the on-board unit for usage of stored value card to pay the fee. Because the prepaid card would have no identifying information (as opposed to a credit card), this method would achieve maximum privacy for the participants.

Another option is to track and collect the fee not through a government-issued device, but through a multi-purpose, private-sector application or tool that records the mileage and then transmits it to a private entity for billing (Baker and G 2011).

A 2017 study in the *Journal of Public Economics* found that "a VMT tax designed to increase highway spending by $55 billion per year increases annual welfare by $10.5 billion or nearly 20% more than a gasoline tax does because: (1) the differentiated VMT tax is better than the gasoline tax at targeting its tax to and affecting the behavior of those drivers who create the greatest externalities, and (2) the greater fuel economy that results from a higher CAFE standard effectively reduces a gasoline tax and its benefits, but has less effect on a VMT tax and its benefits. Therefore, the empirical findings indicate that implementing a VMT tax is a more efficient policy than raising the gasoline tax to improve the financial and economic condition of the highway system. Importantly, we also identify considerations that suggest that a VMT tax is likely to be more politically attractive to policymakers than raising the gasoline tax (Langer, Maheshri and Winston 2017). The VMT tax rate should be related to the current state and local motor fuel taxes, motor license taxes, and highway fees:

VMT tax revenue = current state and local motor fuel taxes + motor license taxes + highway fees.

The VMT tax rate for combination trucks weighing above 60,000 pounds is 2.9 cents in Kentucky. New Mexico, New York, and Oregon also have a gradually increasing VMT tax based on vehicle weight. Buses and commercial traffic charge higher VMT taxes because these vehicles cause more road damage. The speed of motorcycles is low, and because they cause the least damage and have the least impact on congestion, they are usually charged a lower VMT tax.

1.5. Public transportation impact and E-commerce

Due to the spread of the coronavirus, people's daily travel habits have undergone significant changes. In order to minimize contact with others, people began avoiding public transportation, such as buses, trains, and carpooling. Private cars and bicycles are the preferred means of transportation during the pandemic.

During the pandemic, Metrorail's ridership has dropped approximately 90%. Almost 70% of bus trips are work trips in DC. According to Metro data, Metrobus mainly serves low-income, essential workers (WMATA, Washington Metropolitan Area Transit Authority, (2021), "Covid-19 public information: WMATA", 2021).

During the pandemic, the reduction in public transportation use was very obvious. In April 2020, the drop in public transportation riders reached 80%. There was a slight increase after April 2020. However, it remains above 60% by early 2021. Slow growth may occur with the introduction and wider use of vaccines (EBP US 2021).

Due to the pandemic, the total retail sales decreased and most of the growth in retail came from e-commerce sales. According to data from the U.S. Department of Commerce (Feb 2021), US e-commerce sales increased by 32.4%, while total retail sales increased by only 6.9% in 2020. E-commerce accounted for nearly three-quarters (74.6%) of all retail growth in 2020.

Many consumers have discovered the convenience of e-commerce and other online activities during the pandemic. According to an analysis by (McKinsey Global Institute 2021), e-commerce has increased by 3.3 times in the United States during the pandemic. It accounted for about 20% of total retail sales in 2020.

1.6. Congestion ranking and travel impact.

According to (INRIX 2021) INRIX (2020), Washington DC was the 12th most congested city in the United States in 2020 and the 89th most congested city in the world. The annual cost of congestion per driver was $427.43.

According to the (DMV 2021), the number of active vehicle registrations is 358,963, and the number of registrations renewed is 196,237 for the entire year of 2019.

(ValuePengium 2020)conducted an online survey of more than 1,200 Americans during the pandemic and found that 48% of them canceled their summer travel plans during the pandemic. In addition, about one-sixth (16%) of people expected to wait more than a year before they can travel again. The epidemic has also caused people to change their overall view of travel. 52% of them are more afraid of future overseas travel.

According to the US Bureau of Transportation Statistics (2021), the number of trips in the District of Columbia before the pandemic peaked at 4,198,510 trips per month in September 2019. At the beginning of the pandemic (January 2020), the number of trips per month was 3,105,844. Then it began to decrease, to 2,054,240 in January 2021, and then began to gradually increase, reaching 4,209,742 in September 2021 and returning to the number of trips per month in September 2019.

Parking

Washington DC has street parking and parking garages. Privately owned garages charge roughly $10 to $30 per day. Parking charges at the Metro stations vary between $4.60 to $5.10 per day. The parking meter rates are $2.30 per hour for commercial and passenger vehicles citywide. There are around 18,000 metered parking spaces in Washington DC (DDT, District Department of Transportation; Parking meters 2021).

The DC parking permit fee for DC residents is $50 for the first vehicle, $75 for the second vehicle, $100 for the third vehicle, and $150 for each vehicle beyond the first three vehicles (DMV 2021).

DC has an 18% tax for parking motor vehicles in commercial lots. This 18% parking tax in commercial lots is dedicated to the Washington Metropolitan Area Transit Authority according to (DeWitt 2020).

According to the DC Motor Vehicle Administration data, DC issued 837,899 parking tickets throughout 2020. It also issued 53,929 citations for moving violations. The number of speeding tickets and red-light tickets issued by cameras in 2020 is 1.3 million (Austernuhle 2012). Therefore, the possible revenue from parking tickets reaches 62 million $/year. Revenue from moving violations is approximately 8.8 million $/year.

2. Utility functions & integration of the mode choice and revenue models:

In this section utility functions for mode choices are developed and they will be used in the revenue model.

2.1. Methodology:

Machine learning is widely used in many fields throughout the world including the healthcare sector, transportation, advertisement, economics, and image recognition. Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed (Mozaffarian 2015). Furthermore, machine learning at its most basic level is the practice of using algorithms to parse data, learn from it, and then make a determination or prediction about something in the world (Das, et al. 2015). There are two major categories of problems often solved by machine learning i.e., regression and classification. The regression algorithms are used for numeric data and classification problems include binary and multi-category problems (Abduljabbar 2019).

Machine learning algorithms are further divided into two categories including supervised learning and unsupervised learning algorithms (Strecht, et al. 2015).The supervised learning algorithm is performed by using prior knowledge in output values whereas the unsupervised learning algorithm does not have predefined labels; hence, its goal is to infer the natural structures within the dataset (Sathya and Abraham 2013). In this section, the supervised machine learning algorithm, namely logistic regression is used to define the utility functions for the mode choices.

2.2. Data processing:

Most of the data sets contain missing values. Therefore, in the data processing step, the missing values are replaced with meaningful values without changing the structure of the data sets. Data Analysis is carried out using JupyterLab of Python.

2.3. Potential influencing factors and exploratory data analysis (EDA):

The potential influencing factors are identified based on the need for the revenue model, which is called a set of explanatory variables and used as initial inputs of the model. To examine the effect of each of the factors on the bridge condition rating, univariate and bivariate analyses of the variables were conducted.

2.4. Utility theory for discrete choice model:

Utility is an indicator of value to an individual. In a discrete choice experiment, a decision-maker chooses a single alternative from a choice set of finite number of mutually exclusive alternatives where the choice set is exhaustive. An individual is visualized as selecting a mode which maximizes his or her utility (Khan 2007).

The utility of a travelling mode is defined as an attraction associated to by an individual for a specific trip. This hypothesis is known as utility maximization. This can be stated as alternative, 'i', is chosen among a set of alternatives, if and only if the utility of alternative, 'i', is greater than or equal to the utility of all alternatives, ' j', in the choice set, C.

2.5. Model development:

To develop the model, the data set is split into training and test sets. 80% percent of the data set is considered for the training purpose and the other 20%, which was not used in the training process, is used for the evaluation process. The problem is considered as a supervised classification problem, meaning, the data set is labeled, the input variables are known, and the outcome, which is the mode choices, is a multi-class categorial variable. The logistic regression algorithm in machine learning is being used to compute the coefficients of the utility function.

Logistic Regression Model:

Logistic regression is a method for fitting a regression curve, y=f(x), when y consists of binary coded (0, 1- failure, success) data. When the response is a binary (dichotomous) variable and x is numerical, logistic regression fits a logistic curve to the relationship between x and y. The logistic curve is an S-shaped or sigmoid curve, often used to model population growth. A logistic curve starts with slow, linear growth, followed by exponential growth, which then slows again to a stable rate (Park 2013). A simple logistic function is defined by the formula:

$$\boldsymbol{y = \frac{e^x}{e^x + 1}} \tag{1}$$

Logistic regression is one of the machine learning classification algorithms for analyzing a dataset in which there are one or more independent variables (IVs) that determine an outcome and also categorical dependent variable (DV) (Miguel-Hurtado 2016). Linear regression uses output in continuous numeric whereas logistic regression transforms its output using the logistic sigmoid function to return a probability value which can then be mapped to two or more discrete classes (Ng and Jordan 2002).
There are three forms of logistic regression:

a) Binary logistics regression (two possible outcomes in a DV)
b) Multinomial logistics regression (three or more categories in DV without ordering)
c) Ordinal logistics regression (three or more categories in DV with ordering)

Furthermore, the logistic regression model uses a more complex impedance function (known as sigmoid function or logistic function) instead of linear function (Park 2013). Logistic regression limits the cost function values to be between 0 and1. The sigmoid function is defined as follows:

$$\boldsymbol{\sigma(z) = \frac{1}{1 + e^{-z}}} \tag{2}$$

In the formula (2), $\sigma(z)$ represents the output between 0 and 1(probability estimate), z= input to the function and e= base of the natural log (Nishadi 2019).

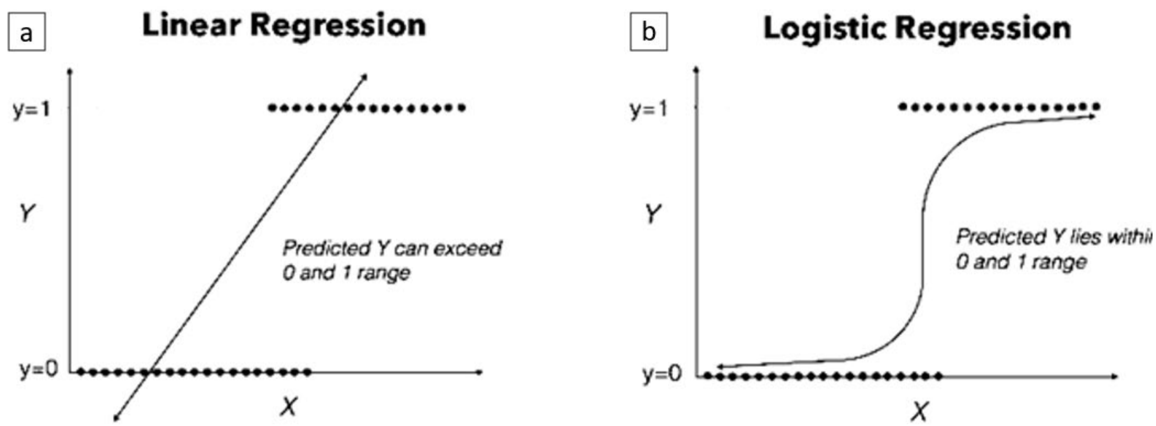


Figure 1. Regression, (a) Linear Regression (b)Logistic regression

According to the data set we are using, 1 indicates a certainty of the individual being in favor of teleworking and 0 indicates that there is no chance of telecommuting for the individual with the given factors.

Input variables:
The list of input variables for the mode choice utility functions are provided in the chart below.

Table 1. Input variables for Mode choice model

Variable name	**Description**
Trip duration minutes	Duration of trip in minute, calculated as the difference between the trip start_time and end_time.
Trip distance miles	Distance in meters measured along the trip route
Household Income	Total income of the household
Household size	Number of persons that makeup the household. Valid values include:1_person, 2_person, 3_person, 4_person, 5_person, 6_person, 7_plus_person.
Trip cost	Is defined as trip duration times average trip cost per mile.

Output variable:

Output variable is the categorical variable "trip_taker_commute_mode" the following are the categories and the percentages of each category based on data observation.

Table 2. Distribution of mode choices in the data

Mode choice	Percentages of each group in the dataset
Private auto	0.530334
Work from home	0.258995
Public transit	0.092108
Carpool	0.073550
Walking	0.024932
Other travel mode choices	0.014311
Biking	0.00577

Data analysis:

In this section data is analyzed, descriptive statistics, univariate and bivariate analysis are presented.

Descriptive statistics

Average trip distance, trip time and household size and household income can be obtained from the following chart.

Table 3. Descriptive statistics- input variables-Mode choice-part 1

	Count	Mean
Trip duration minutes	15971912.0	20.315939
Trip distance miles	15971912.0	9.062855
Individual Income	15971912.0	66.710300713
Household Income	15971912.0	161.307707034
Household size	15971912.0	3.755603
Trip cost	15971912.0	3.936165

Mean of each of the input variables with respect to categories of output variable is presented below:

Table 4. Descriptive statistics- input variables-Mode choice-part 2

	Trip duration minutes	Trip distance miles	Household Income	Trip cost
Private auto	21.901172	10.637543	157.692235084	4.620081
Work from home	19.761217	9.691642	182.490164421	4.209259
Public transit	21.638427	8.603139	153.682520686	3.736502
Carpool	21.942982	10.403210	150.321943452	4.518306
Walking	14.575215	4.447979	139.864240936	1.931840
Other travel mode choices	15.835480	4.073456	152.845627853	1.769177
Biking	22.783554	5.612003	193.501666200	2.437396

Univariate Analysis:

In this section the distribution of each input variable will be analyzed, and the outliers will be deleted. The green line represents the mean of the variable, and the orange line corresponds to the median in the plot.

a) **Trip duration:** The distribution of trip duration is given below.

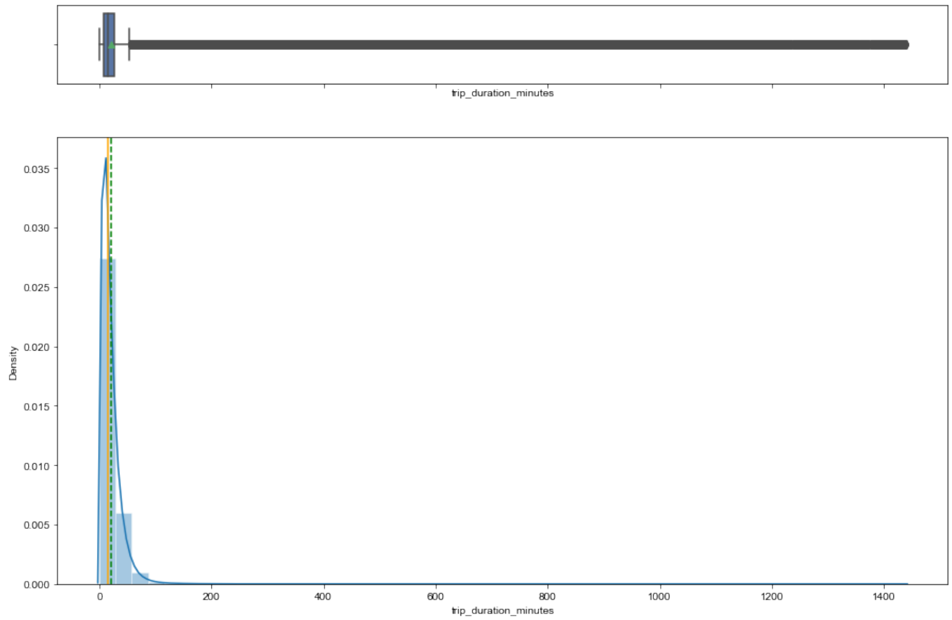


Figure 2. Histogram of Trip Duration with outliers

There are only few trips whose duration is more than 600 minutes, those are considered as outliers and deleted from the dataset. The new plot is shown below.

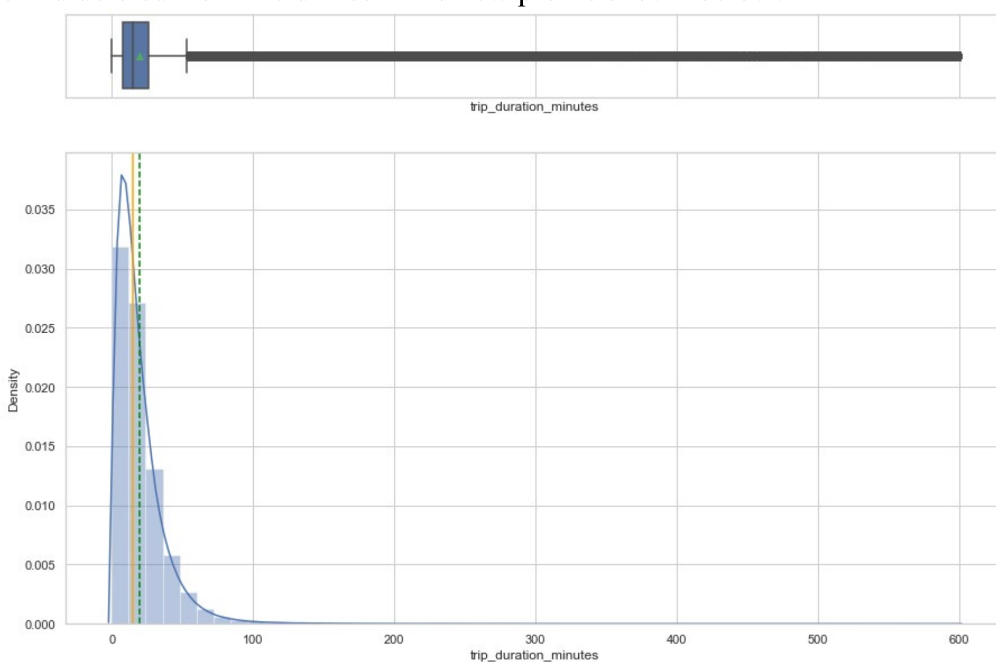


Figure 3. Histogram of Trip Duration without outliers

b) Trip distance: The distribution of trip distance is given below.

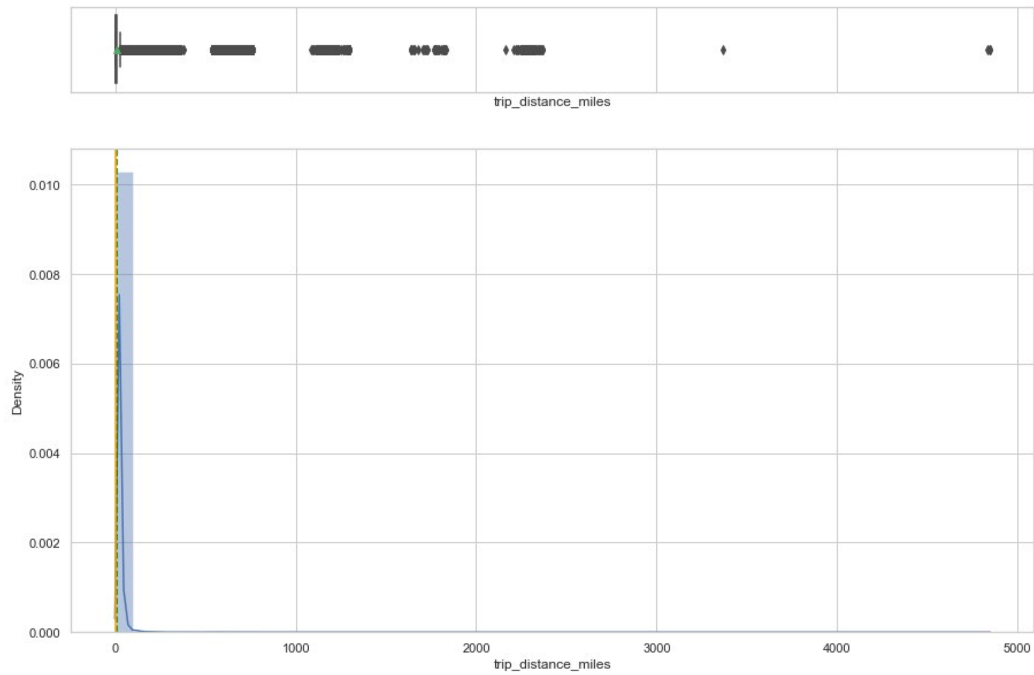


Figure 4. Histogram of Trip Distance with outliers

There are only few trips whose distance is more than 200 miles, those are considered as outliers and deleted from the dataset. The new plot is shown below.

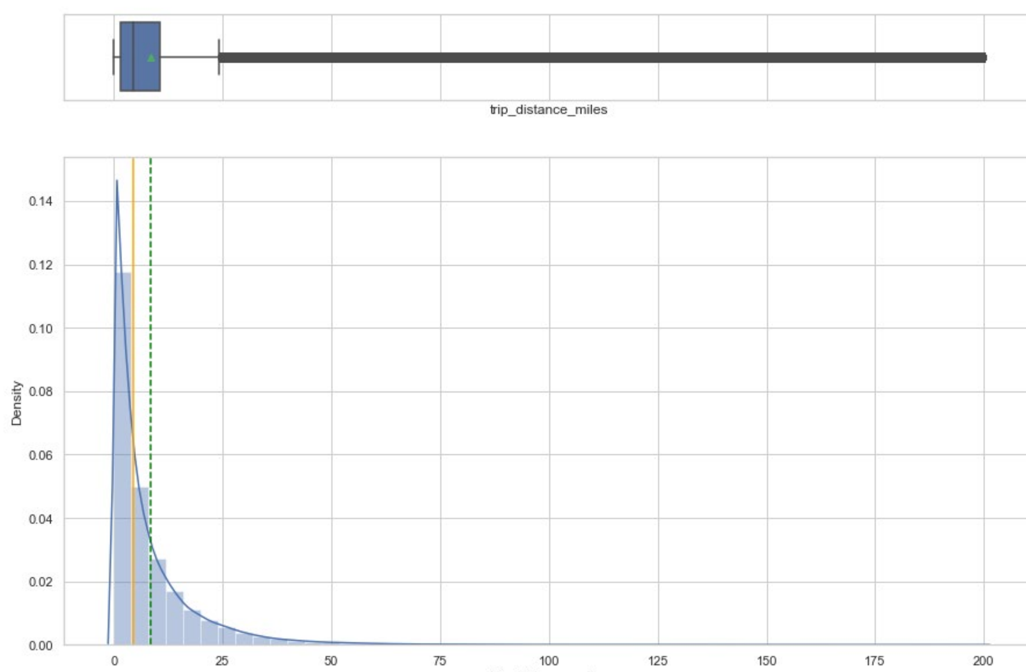


Figure 5. Histogram of Trip Distance without outliers

c) **Household size:** The graph below shows the distribution of the variable household size.

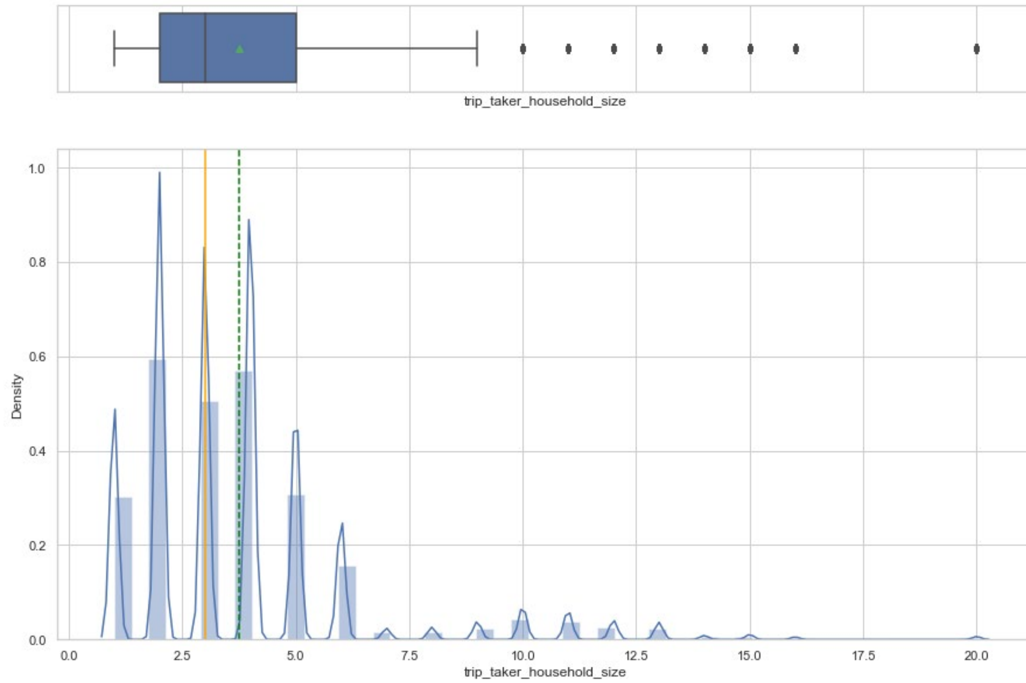


Figure 6. Histogram of Household size with outliers

There are only a few households with size larger than 10, those are considered as outliers and deleted from the dataset, the new plot is shown below.

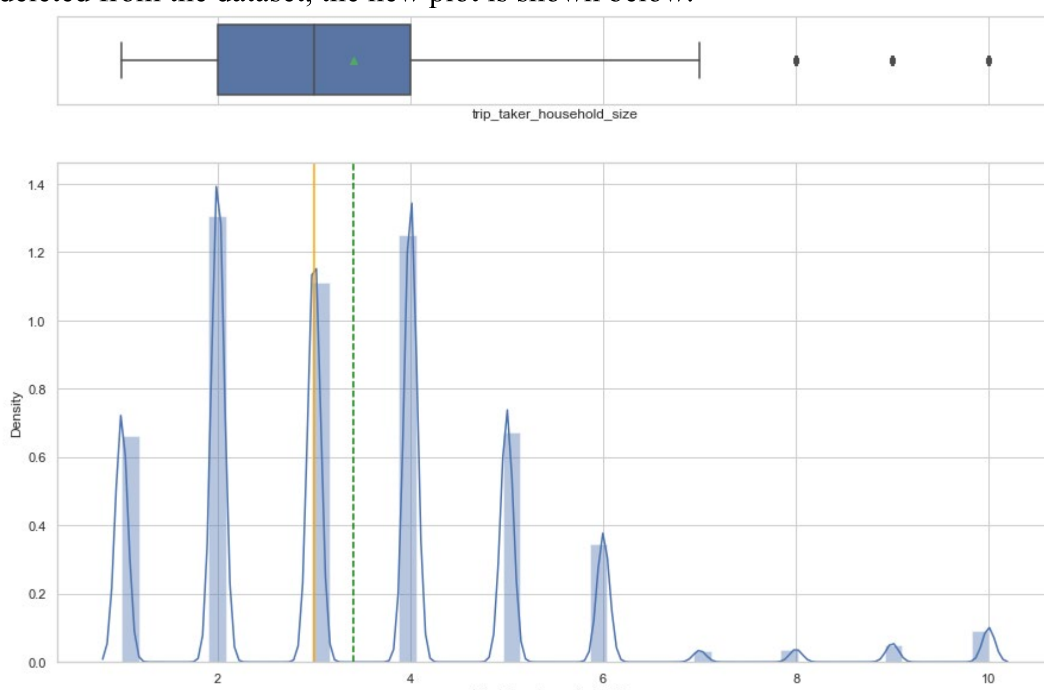


Figure 7. Histogram of Household size without outliers

d) Household income:

The distribution of household income is represented below.

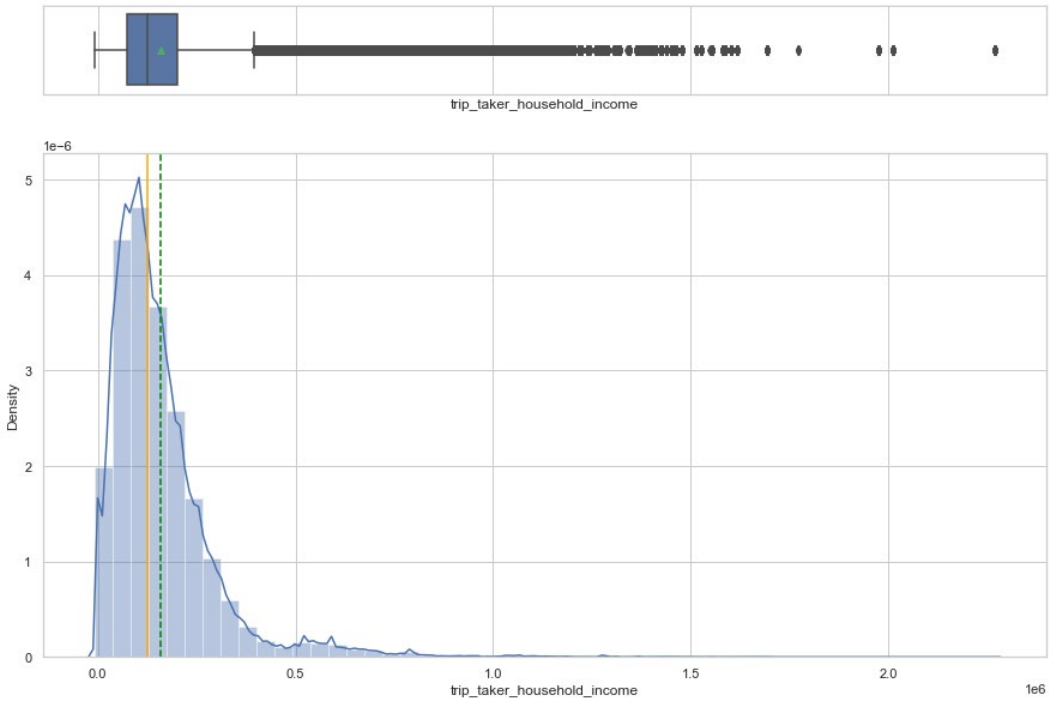


Figure 8. Histogram of Household income

e) Trip cost

The distribution of trip cost is represented below.

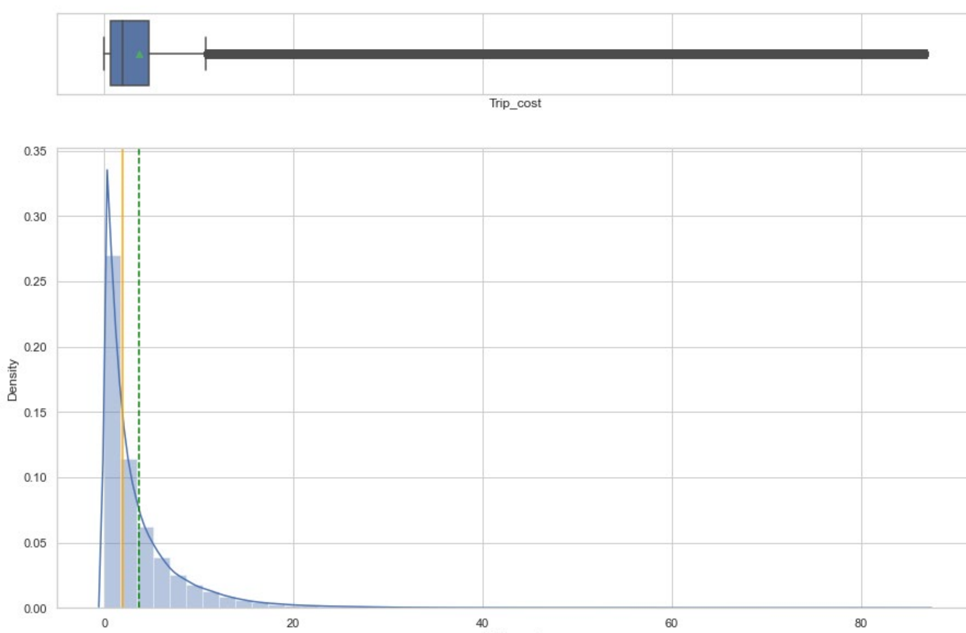


Figure 9. Histogram of Trip cost

Bivariate analysis:
In this section the relation between variables is analyzed. In the following figures household income, Household size, education and age group are plotted with respect to mode choices. The horizontal axis is the mode choices and vertical axis is the count, colors indicate different categories for the plotted variable.

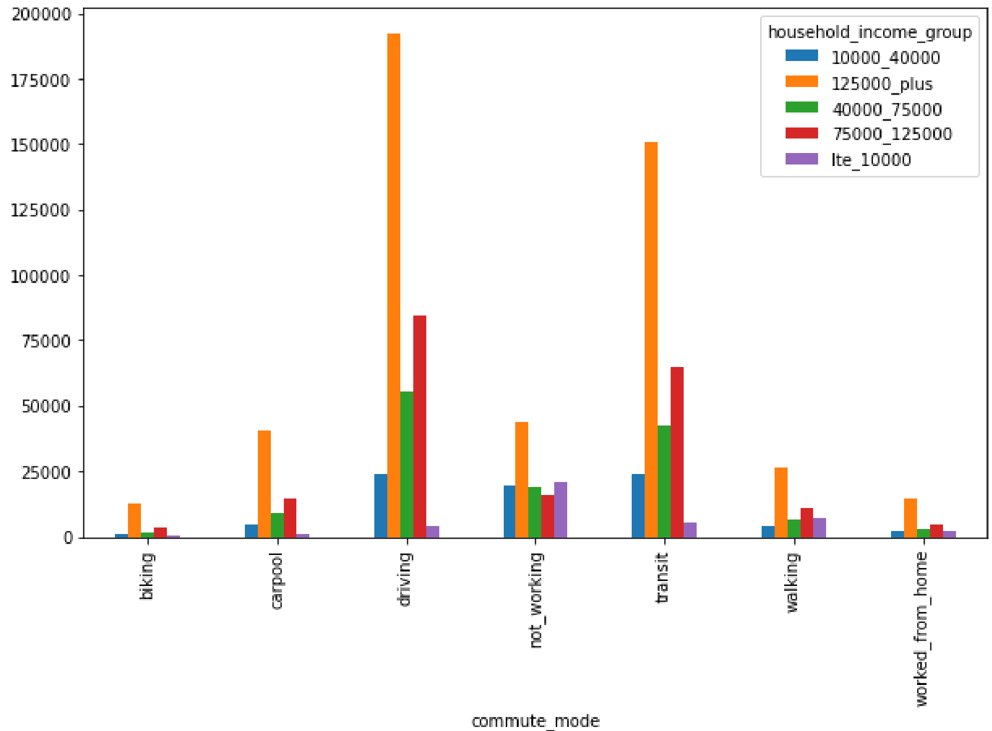


Figure 10. Mode choice & Household income

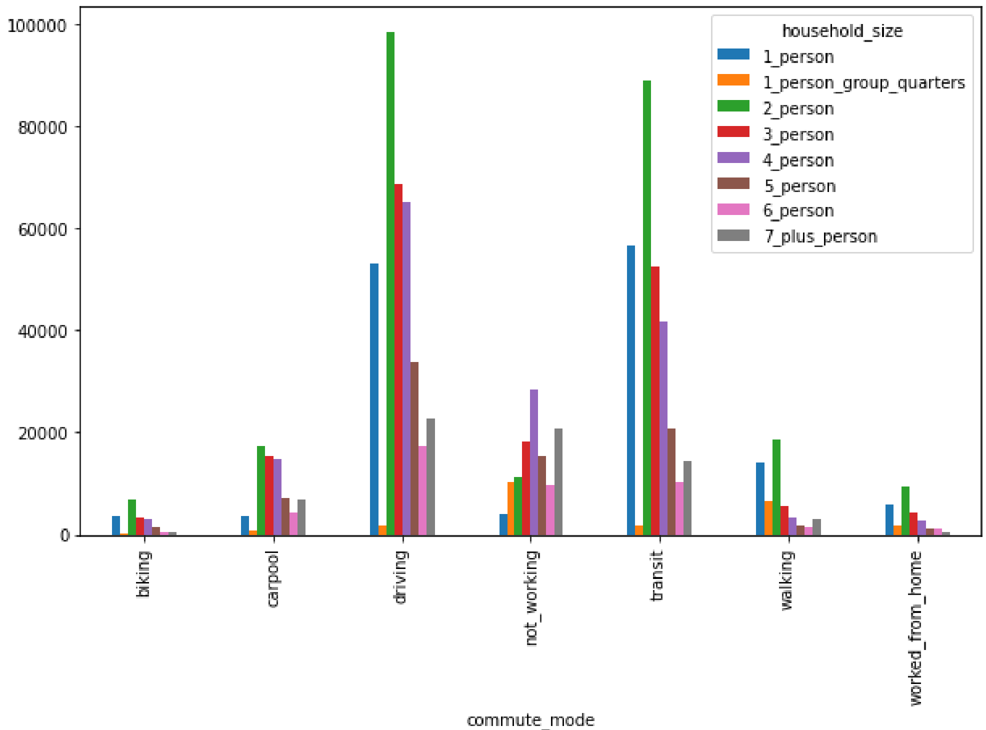


Figure 11. Mode choice & Household size

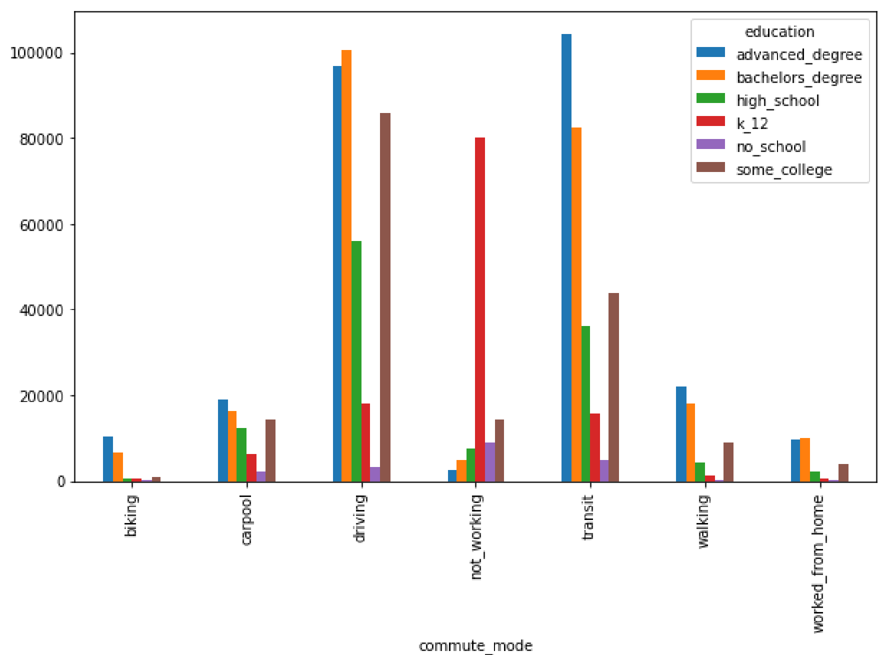


Figure 12. Mode choice & Education

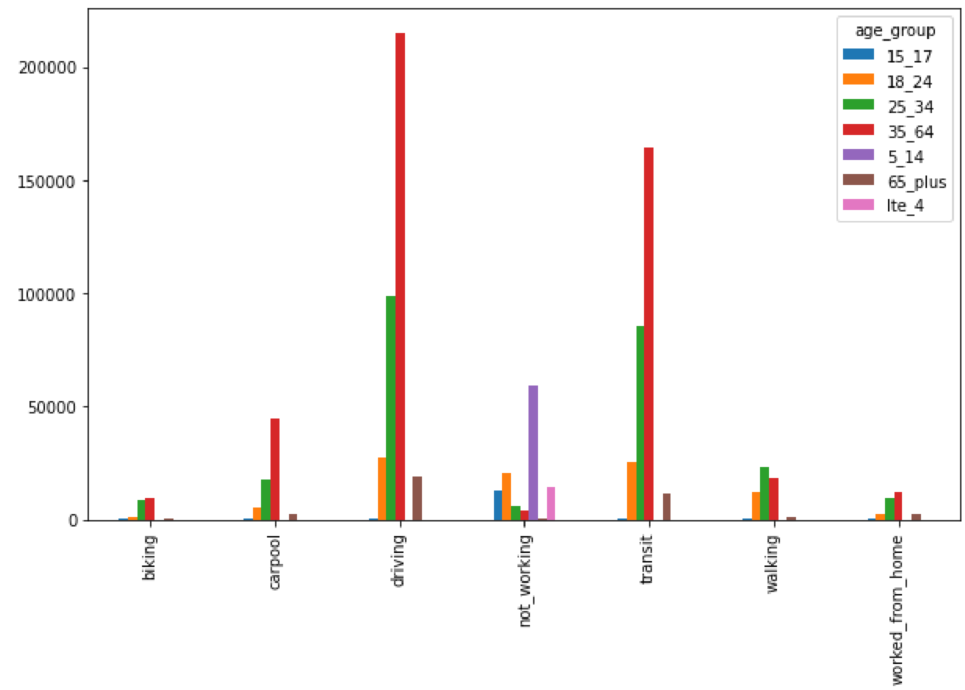


Figure 13. Mode choice & Age group

To check the correlations between the variables a correlation matrix is computed.

Figure 14. Correlation matrix for input variables of mode choice model

2.6. Results: Utility functions for Mode choice:

The categorial variables are encoded, and data scaling is done as a part of data preparation for the development of the model. Then, the data is split into training and test sets. The training set is used for model development and the test set is used for model evaluation.

For all the cases below the dependent variable (output variable) is Mode choice. It is a categorial variable. The chart below represents the distribution of each category.

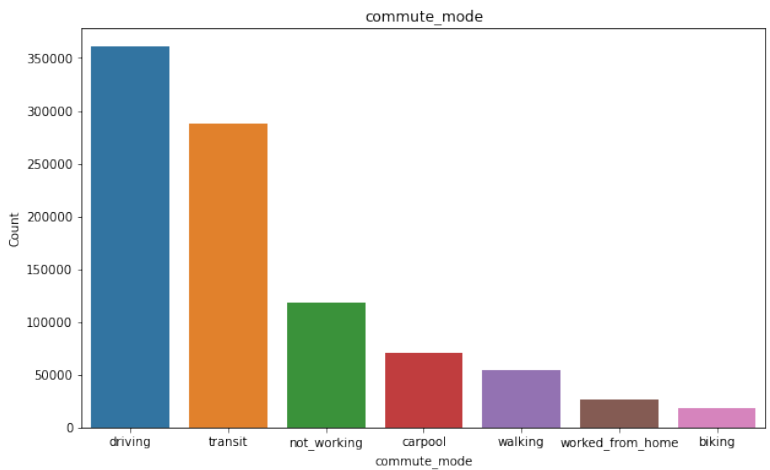


Figure 15. Distribution of Mode choice without combining categories

Table 5. Mode choice frequency table without combining categories

Mode choices	Count
driving	360585
transit	287323
not_working	118128
carpool	70249
walking	54502
worked_from_home	26682
biking	18992

To integrate the mode choice and revenue models, the mode choices for both models must be identical, therefore the driving and carpool categories are combined and called driving, the not working, walking, work from home and biking are combined and called other travel mode choices. The count of each mode choices is given below.

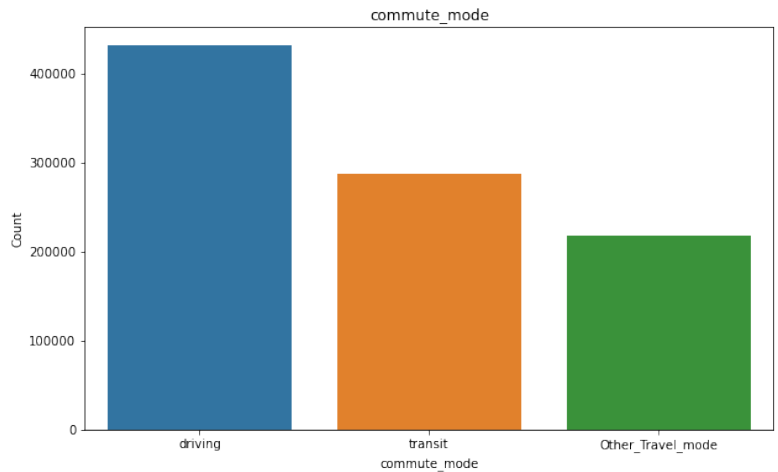


Figure 16. Distribution of Mode choice with combined categories.

Table 6. Mode choice frequency table with combined categories

New mode choices	Count
driving	430834
transit	287323
Other_Travel_mode	218304

Input variables:

X1: Trip duration minute
X2: Trip cost

Utility functions:

$$U_{driving} = -0.023334\ X1 - 0.123639X2 - 0.6257146$$
$$U_{transit} = -0.024045\ X1 - 0.127605X2 - 0.01120879$$
$$U_{\text{other travel mode}} = -0.009863\ X1 - 0.026665X2 - 0.35238241$$

2.7. Integration of the utility functions and revenue model:

To have more accurate results for the revenue model, we must include the wait time and access time for transit mode choice, to do that a total trip time is defined as the trip duration plus a constant that represents the access and wait time. This constant is defined differently in case 1 and 2. The results of the following cases then is used in the development of the revenue model.

Case 1:
X1: Total trip time (For driving and other travel modes is defined as trip duration, for transit it is defined as trip duration+ 18.5 *2)
X2: Trip cost

Utility functions:

$$U_{driving} = -0.043\,X1 - 0.23X2 - 0.55$$
$$U_{transit} = -0.092\,X1 - 0.35X2 - 1.8$$
$$U_{\text{other travel mode}} = -0.003\,X1 - 0.02X2 - 0.37$$

Case 2:
X1: Total trip time (For driving and other travel modes is defined as trip duration, for transit it is defined as trip duration+ 18.5 *1.5)
X2: Trip cost

Utility functions:

$$U_{driving} = -0.047\,X1 - 0.25X2 - 0.55$$
$$U_{transit} = -0.067\,X1 - 0.35X2 - 1.5$$
$$U_{\text{other travel mode}} = -0.007X1 - 0.025X2 - 0.39$$

3. Analysis Methods for Revenues and Other Impacts

3.1. Framework for Transportation Revenue Analysis

3.1.1. Introduction

This chapter presents the model developed for estimating future transportation tax revenues as well some other relevant outcomes (e.g., trips and vehicle miles, mode shares, energy use, expected fatalities and emissions) in the District of Columbia. The basic technical approach used in developing this model is to first classify transportation system users by trip purpose and ability to engage in activities remotely. Then a multinomial logit model is used to estimate the fractions of users who choose various transportation alternatives or modes. Those shares are then used to estimate demand characteristics, including the resulting person trips and vehicle miles. Finally, the demand characteristics, in conjunction with estimated tax mechanisms and rates, are used to estimate transportation revenues, as well as safety and environmental impacts. In this framework, five transportation alternatives are analyzed, with gasoline-powered cars and electric cars treated as competing modes. Those five alternatives are gasoline vehicles (private auto), electric vehicles (private auto), remote activities (e.g., telework and online shopping), public transportation, and commercial trucks. In the overall analysis framework, the commercial trucks, carrying only freight are treated separately from the passenger-serving modes. Q (one-way person trips) are divided into Q_1 and Q_2. Of these, Q_1 refers to trip miles by persons who have the choice of conducting their activities remotely, and served by gasoline vehicles (private auto), electric vehicles (private auto), and public transportation. People can choose their favorite from these alternatives (as well as virtual trips), depending on how they perceive the relative importance of prices, value of time, and travel times. Q_2 refers to trip miles by persons who cannot act remotely. These include gasoline vehicles (private auto), electric vehicles (private auto), and public transportation.

The purposes of travel by persons are distinguished into the working trip purpose, shopping trip purpose, and all other trip purposes. Three types of revenue sources are included in the analysis. These are (1) Vehicle ownership revenues (e.g., from vehicle registration fees), (2) Distance-based tax revenues (e.g., from taxes on fuel or kilowatt hours, and possibly from vehicle miles taxes), and (3) Trip-based tax revenue (e.g., from parking fees or tolls). Other important outputs of the model include total trips and trip miles by mode and purpose, energy use (in gallons and kilowatt hours), expected fatalities, and equivalent carbon dioxide emissions (which translates other pollutants into equivalent carbon dioxide, which is the major greenhouse gas assumed to affect climate). Equivalent carbon dioxide (CO2E) emissions generated in gasoline and diesel vehicles are used in the framework to estimate the environmental impact rate. The following section presents the variables, quantitative relations expressed as equations, and numerical results.

3.1.2. Variables

The variables and parameters that are used in estimating revenues, along with their units and baseline values (if any) are shown in Table 7. Among them, most of the baseline values come from the Replica database or are calculated from the data provided by Replica. The tax on gasoline and electricity for the District of Columbia is from the Council of District of Columbia and Bob Donnellan (Columbia 2023, Donnellan 2023).

Table 7. Variables and Parameters Used in Revenue Model

Variable	Definition	Baseline value
b	Subscript indicating base value	-
c_z	Fraction of users who can act remotely (e.g., Virtual) c_1: Working trip purpose c_2: Shopping trip purpose c_3: All other trip purposes	$c_1 = 0.54$ (Working trip) $c_2 = 0.3$ (Shopping trip) $c_3 = 0.4$(All other trips)
E	Elasticity (overall impedance)	-0.3
E'	Environmental impact rates (kg)	0.4 kg per vehicle mile of equivalent carbon dioxide emissions
e	$\ln^{-1}1 = 2.718282$	2.718282
f	Fatalities(deaths)	1.19 per 100 million vehicle miles in DC
F	Energy consumption (gallons or kwh/vehicle mile)	F_1 =0.0413 (Private auto--gas). F_2 = 0.346 kw*hr/mile (Private auto--electric) F_3= 0 (Virtual) F_4 =0.1652(Public transportation) F_5 =0.1239 (Commercial trucks)
g	Average vehicle usage (miles/year).	g=7,500 miles/year (Private auto—gas, Private auto—electric)

		$g_{transit} = 50{,}000\ miles/year$ (Public transportation, Commercial trucks)
k	Estimated coefficient k_p(Price in \$/person trip) k_t(Trip time) k Coefficient	$k_{t-private\ auto}$=0.043, $k_{t-transit}$=0.092, $k_{t-other\ travel\ mode}$=0.003, $k_{p-private\ auto}$=0.23, $k_{p-transit}$=0.35, $k_{p-other\ travel\ mode}$=0.02, $k_{private\ auto}$=0.55, $k_{transit}$=1.8, $k_{other\ travel\ mode}$=0.37,
L_z	Average trip length (miles) L_1: Working trip purpose L_2: Shopping trip purpose L_3: All other trip purposes L_5 : Working trip purpose for commercial trucks	$L_1 = 11.2\ miles$ (Working trip) $L_2 = 5.51\ miles$ (Shopping trip) $L_3 = 8.97\ miles$ (All other trips) $L_5 = 5\ miles$ (Working trip for commercial trucks)
m	Subscript for mode	1: Private auto--gas, 2: Private auto--electric, 3: Virtual, 4: Public transportation, 5: Commercial trucks
n	Subscript for persons who can or cannot act virtually	1: persons who can act virtually. 2: persons who cannot act virtually.
o_m	Avg. occupancy (travelers/veh)	$o_1 = o_2 = 1.5$ (private auto—gas and electric) o_4 = 8.98 (public transportation) o_5 = 1.1 (commercial trucks)
p	Price in \$/person trip, including all out-of-pocket costs. (Price in \$/person trip).	Working trip purpose: $\mathrm{p}_{\mathrm{g-working}} = 3.03$ \$/person trip (Private auto-gas) $\mathrm{p}_{\mathrm{e-working}} = 4.48$ \$/person trip (Private auto-electric) $\mathrm{p}_{\mathrm{t-working}} = 0$ \$/person trip (Virtual) $\mathrm{p}_{p-working} = 3.49$ \$/person trip (Public transportation) $\mathrm{p}_{truck-working} = 4$ \$/person trip (Commercial truck)

		Shopping trip purpose: $p_{\mathrm{g-shopping}} = 1.49$ \$/person trip (Private auto-gas) $p_{\mathrm{e-shopping}} = 2.20$ \$/person trip (Private auto-electric) $p_{\mathrm{t-shopping}} = 0$ \$/person trip (Virtual) $p_{p-shopping} = 1.71$ \$/person trip (Public transportation) All other trip purpose: $p_{\mathrm{g-all\ other}} = 2.43$ \$/person trip (Private auto-gas) $p_{\mathrm{e-all\ other}} = 3.59$ \$/person trip (Private auto-electric) $p_{\mathrm{t-all\ other}} = 0$ \$/person trip (Virtual) $p_{p-all\ other} = 2.79$ \$/person trip (Public transportation)
p’	Impedance ($p'_{yn} = -\bar{u}_{myzn}$) (\$/one-way trip) = negative utility	-$\bar{u}$
Q_{yzn}	Actual person miles per year	-

Q'_{nz}	Potential one-way person miles/year Q'_{11}: Work trip miles by persons who can act virtually Q'_{21}: Work trip miles by persons who cannot act virtually Q'_{12} : Shopping trip miles by persons who can act virtually Q'_{22} : Shopping trip miles by persons who cannot act virtually Q'_{13} : All other trips miles by persons who can act virtually Q'_{23} : All other trips miles by persons who cannot act virtually Q'_5 : Work trip miles for the commercial truck mode	$Q'_{11} = c_1 \times 2{,}000{,}000 * 11.2 miles * 365\ days$ $= 0.54 \times 2{,}000{,}000 * 11.2 miles * 365\ days$ $= 4.4150 \times 10^9\ person\ miles/year$ $Q'_{21} = (1 - c_1) \times 2{,}000{,}000 * 11.2 miles * 365\ days$ $= (1 - 0.54) \times 2{,}000{,}000 * 11.2 miles * 365\ days$ $= 3.7610 \times 10^9\ person\ miles/year$ $Q'_{12} = c_2 \times 800{,}000 * 5.51\ miles * 365\ days$ $= 0.3 \times 800{,}000 * 5.51\ miles * 365\ days$ $= 4.8268 \times 10^8\ person\ miles/year$ $Q'_{22} = (1 - c_2) \times 800{,}000 * 5.51\ miles * 365\ day$ $= (1 - 0.3) \times 800{,}000 * 5.51\ miles * 365\ days$ $= 1.1262 \times 10^9\ person\ miles/year$ $Q'_{13} = c_3 \times 600{,}000 * 8.97 miles * 365\ days$ $= 0.4 \times 600{,}000 * 8.97 miles * 365\ days$ $= 7.8577 \times 10^8\ person\ miles/year$ $Q'_{23} = (1 - c_3) \times 600{,}000 * 8.97 miles * 365\ days$ $= (1 - 0.4) \times 600{,}000 * 8.97 miles * 365\ days$ $= 1.1787 \times 10^9\ person\ miles/year$ $Q'_5 = 19{,}500 * 5 miles * 365\ days$ $= 19{,}500 * 5 miles * 365\ days$ $= 3.5588 \times 10^7\ person\ miles/year$
q	Vehicle miles/year	-
R	Revenue in $/yr	-
S	Mode share	-

T_{mz}	Trip time (hours, including in-vehicle, wait, access) $T_{11}, T_{21}, T_{31}, T_{41}, T_{51}$: Working trip time for Private auto, Virtual, Public transportation, and Commercial truck. (hours) $T_{12}, T_{22}, T_{32}, T_{42}$: Shopping trip time for Private auto, Virtual, Public transportation. (hours) $T_{13}, T_{23}, T_{33}, T_{43}$: Trip time for all other trip purposes for Private auto, Virtual, and Public transportation. (hours)	$T_{11} = T_{21} = 21.51/60$ hrs (Work trip: Private auto—gas and electric) $T_{31} =$ 0 hrs (Work trip: Virtual) $T_{41} = \frac{21.03}{60} + \left(\frac{18.5}{60} * 2\right) = 0.9672$ hrs (Work trip: Public transportation) $T_{51} =$0.545 hrs (Work trip: Commercial truck) $T_{12} = T_{22} =$ 0.27 hrs (Shopping trip: Private auto—gas and electric) $T_{32} =$ 0 hrs (Shopping trip: Virtual) $T_{42} = 0.60 + (\frac{18.5}{60} * 2) = 1.2167$ hrs (Shopping trip: Public transportation) $T_{13} = T_{23} =$ 0.29 hrs (All other trips: Private auto—gas and electric) $T_{33} =$ 0 hrs (All other trips: Virtual) $T_{43} = 0.46 + \left(\frac{18.5}{60} * 2\right) = 1.0767$hrs (All other trips: Public transportation)
u	Utility in $/one-way person trip	$U_{driving} = -0.043\, T_{mz} - 0.23p - 0.55$ (Private auto) $U_{transit} = -0.092\, T_{mz} - 0.35p - 1.8$ (Public transportation) $U_{\text{other travel mode}} = -0.003\, T_{mz} - 0.02p - 0.37$ (Other travel modes)
X	Tax per gallon Distance-based taxes X_{dm} ($/gallon or kwh) Trip-based taxes X_{xmz} : X_{xm1}: Working trip-based taxes X_{xm2}: Shopping trip-based taxes X_{xm3}: All other trips-based taxes ($ one way trip)	$X_{d1} = X_{d5}$ =0.342$/gallon (Private auto—gas and commercial truck, since tax of gasoline and diesel are same in DC); X_{d2} =0.007$/kwh (Private auto -- electric) $X_{d3} = X_{d4} = 0$ (Virtual and public transportation) Working trip-based taxes $X_{x11} = X_{x21} = X_{x51} =$ 0.03 $ one way trip (0.06$ per round trip); (Working trip) $X_{x31} = X_{x41} = 0$ (Working trip) (Virtual and public transportation) Shopping trip-based taxes $X_{x12} = X_{x22} = 0.01$ $ one way trip (0.02$ per round trip); (Shopping trip)

	Ownership taxes (Vehicle Registration fee) X_{rm} ($/vehicle)	$X_{x32} = X_{x42} = 0$ (Shopping trip) (Virtual and public transportation) All other trips-based taxes $X_{x13} = X_{x23} = 0.02$ \$ one way trip (0.04\$ per round trip); (All other trips) $X_{x33} = X_{x43} = 0$ (All other trips) (Virtual and public transportation) $X_{r1} = X_{r2} = X_{r5} = 5\ \$/vehicle$ (Private auto—gas and electric, Commercial truck) $X_{r3} = X_{r4} = 0\ \$/vehicle$ (Virtual and public transportation)
y	Subscript for socio-economic group	Only one socio-economic group
z	Subscript for trip purpose	1: work, 2: shopping, 3: All others

3.1.3. Methodology – Revenue Equation

The equations used in the revenue model are presented below. Three types of revenues are considered, i.e., distance-based revenues (such as from fuel or vehicle miles taxes, trip-based revenues (such as from parking taxes or tolls), and vehicle ownership revenues (e.g., registration taxes).

The general utility function is formulated based on the trip out-of-pocket-cost p and trip time T_{mz}, with their corresponding coefficients:

$$u_{myz} = -k_p p - k_t T_{mz} - k \qquad (1)$$

This function is used to estimate the utility (actually disutility, since it is negative) of these key factors (trip cost and trip time) on subsequent numbers of trips, mode choices, revenues, and other outputs.

The following three equations are the utility functions (case 1) used by the revenue model, where T_{mz} refers to the trip duration, and p is the trip cost. These three utility functions are used in private auto mode, public transit mode, and other travel modes, respectively.

$$U_{driving} = -0.043\, T_{mz} - 0.23p - 0.55 \qquad (2)$$

$$U_{transit} = -0.092\, T_{mz} - 0.35p - 1.8 \qquad (3)$$

$$U_{\text{other travel mode}} = -0.003\, T_{mz} - 0.02p - 0.37 \qquad (4)$$

The mode share S_{myzn} function is formulated as:

$$S_{myzn} = \frac{e^{u_{myz}}}{\sum_m e^{u_{myz}}} \quad (5)$$

For the mode share S_{myzn}, the logit model is used to compute each mode's share for different socio-economic groups and trip purposes; m refers to the travel modes; y refers to the socio-economic groups; z refers to trip purposes; n refers to persons who can or cannot act virtually. The formula is e raised to the power of utility function u_{myz} and divided by the sum of such terms for all five alternatives.

The average utility $\bar{u}_{myzn}$ is obtained when the utility of different modes u_{myz} is weighted by their share of the trips S_{myzn}:

$$\bar{u}_{myzn} = \sum_m \sum_y \sum_z ((S_{myzn})(u_{myz})) \quad (6)$$

The impedance function, reflecting the difficulty or "generalized cost" of making trips is the negative average utility $\bar{u}_{myzn}$ function:

$$p'_{yn} = -\bar{u}_{myzn} \quad (7)$$

The actual person miles Q_{yzn} are estimated based on the potential one-way person miles per year Q'_{nz}, impedance p' and elasticity with respect to overall impedance E. The actual person miles per year Q_{yzn} are formulated as:

$$Q_{yzn} = Q'_{nz} p'^{E}_{yn} \quad (8)$$

The actual person miles Q_{yzn} is multiplied by the mode share S_{myzn} to obtain the person miles by mode Q_{myzn}. The person miles by mode Q_{myzn} are formulated as:

$$Q_{myzn} = Q_{yzn} S_{myzn} \quad (9)$$

The vehicle miles per year by mode q_{myzn} are formulated as:

$$q_{myzn} = \frac{Q_{myzn}}{o_m} \quad (10)$$

According to the obtained actual person miles Q_{myzn} and the average occupancy o_m of different modes, the vehicle miles q_{myzn} by mode m can be obtained. The vehicle miles by mode q_{myzn} will be used in the following revenue function.

The distance-base revenue R_{dmyzn} function is based on the energy consumption F_m, distance-based taxes X_{dm} and vehicle miles q_{myzn} by mode m:

$$R_{dmyzn} = q_{myzn} F_m X_{dm} \tag{11}$$

The energy consumption rate F_m of different modes m (specified in gallons or kilowatt hours per vehicle mile) is multiplied by the distance-based taxes X_{dm} and vehicle miles by mode q_{myzn} to estimate the revenue according to the driving distance R_{dmyzn}.

The trip-based revenue R_{xmyzn} is formulated as the vehicle miles by mode q_{myzn} divided by the average trip length L_z for different trip purposes and multiplied by the trip-based taxes X_{xmz}:

$$R_{xmyzn} = (\frac{q_{myzn}}{L_z}) X_{xmz} \tag{12}$$

The vehicle ownership revenue R_{rmyzn} is based on the vehicle registration fee for each vehicle. Therefore, vehicle miles by mode q_{myzn} are divided by average vehicle usage g and multiplied by ownership taxes (vehicle registration fee) X_{rm} to obtain tax ownership revenue R_{rmyzn}:

$$R_{rmyzn} = (\frac{q_{myzn}}{g}) X_{rm} \tag{13}$$

The total revenue R_n is the sum of distance-based revenue R_{dmyzn}, trip-based revenue R_{xmyzn} and vehicle ownership revenue R_{rmyzn} corresponding to the different socio-economic groups, trip purposes and modes. The total revenue R_n is formulated as:

$$R_n = \sum_m \sum_y \sum_z (R_{dmyzn} + R_{xmyzn} + R_{rmyzn}) \tag{14}$$

For commercial trucks, which carry freight rather than persons, revenues are estimated completely separately from other modes and their revenues. The initial framework equations used for other modes are also used for trucks. Because the commercial trucks can only be used to carry freight rather than persons, only one trip purpose (work) is needed to analyze the three types of revenues (distance-based revenues, trip-based revenues, and vehicle ownership revenues).

The results obtained with the baseline values of input parameters and the sensitivities of those results to changes in the input parameters are presented in Chapter 4.

4. Results Obtained with the Revenue Model

4.1. Results Estimated from the Baseline Values

This framework currently distinguishes the purposes of travel into work trips, shopping trips, and all other trips. The travelers select among four transportation alternatives for persons, i.e., gas vehicles (private auto), electric vehicles (private auto), telework, and public transportation. Therefore, the revenue is first classified by working trip purpose, shopping trip purpose, and all other trip purposes in the following results table. Because commercial trucks provide freight rather than person transportation, they do not compete with the other modes, and their revenues are estimated separately. The following table also shows the equivalent carbon dioxide (CO2E) emissions generated in gasoline and diesel vehicles used to estimate the environmental impact rate. The framework divides persons into two categories, depending on their ability to engage in activities remotely. They are (1) Persons who can act remotely and (2) Persons who cannot act remotely. Those in the first group can choose to act remotely as well as choose among available transportation modes. Finally, the total revenue is obtained by summing up the revenues obtained from these different groups, trip purposes, and modes, as shown in Tables 8 to 13, thus summarizing the results for the baseline values provided in Table 7 of Chapter 3.

Table 8. Baseline outputs by modes

	Persons who can act remotely				Persons who cannot act remotely			Commercial trucks	Total value for each row
	Private auto -- gas	Private auto -- electric	Virtual	Public transportation	Private auto -- gas	Private auto -- electric	Public transportation		
One-way trips/yr	$4.6209 * 10^{7}$	$3.4562 * 10^{7}$	/	$2.2684 * 10^{6}$	$1.3916 * 10^{8}$	$1.0632 * 10^{8}$	$7.1139 * 10^{6}$	$8.2129 * 10^{6}$	$3.4385 * 10^{8}$
Vehicle miles/yr	$4.5500 * 10^{8}$	$3.3536 * 10^{8}$	/	$2.1708 * 10^{7}$	$1.2732 * 10^{9}$	$9.5357 * 10^{8}$	$6.2646 * 10^{7}$	$4.1065 * 10^{7}$	$3.1426 * 10^{9}$
Gallons/ yr	$1.8791 * 10^{7}$	/	/	$3.5861 * 10^{6}$	$5.2584 * 10^{7}$	/	$1.0349 * 10^{7}$	$5.0879 * 10^{6}$	$9.0398 * 10^{7}$
Kwh/yr	/	$1.1603 * 10^{8}$	/	/	/	$3.2994 * 10^{8}$	/	/	$4.4597 * 10^{8}$
Number of Vehicles	$6.0666 * 10^{4}$	$4.4715 * 10^{4}$	/	434.1574	$1.6976 * 10^{5}$	$1.2714 * 10^{5}$	$1.2529 * 10^{3}$	821.2937	$4.0479 * 10^{5}$

The variables corresponding to the values in each row are shown on the left side of the table. The first and second rows of the table correspond to different travel modes and persons who can or can't act remotely in each column. The last column of the table shows the sum of the corresponding variables in each row. Among them, the Virtual mode does not contribute to these revenues.

Three types of revenue are shown Table 9. These are distance-based revenues, trip-based revenues, and vehicle ownership revenues. The revenue generated by persons who cannot act remotely accounts for the largest fraction since those persons may be unable to avoid some transportation taxes. However, virtual and public transportation alternatives do not generate any transportation revenues in the present model.

Table 9. Different Types of Revenue

	Persons who can act remotely				Persons who cannot act remotely			Commercial truck	Total value for each row
	Private auto -- gas	Private auto -- electric	Virtual	Public transportation	Private auto -- gas	Private auto -- electric	Public transportation		
Distance-based revenues ($/yr)	$6.4266 * 10^6$	$8.1224 * 10^5$	0	0	$1.7984 * 10^7$	$2.3096 * 10^6$	0	$1.7401 * 10^6$	$2.9272 * 10^7$
Trip-based revenues ($/yr)	$1.1505 * 10^6$	$8.4284 * 10^5$	0	0	$3.1129 * 10^6$	$2.3107 * 10^6$	0	$2.4639 * 10^5$	$7.6634 * 10^6$
Vehicle ownership revenues ($/yr)	$3.0333 * 10^5$	$2.2357 * 10^5$	0	0	$8.4881 * 10^5$	$6.3572 * 10^5$	0	$4.1065 * 10^3$	$2.0155 * 10^6$
R_n	$9.7592 * 10^6$				$2.7201 * 10^7$			$1.9906 * 10^6$	$3.8951 * 10^7$

Gasoline (private auto) and electric vehicle (private auto) travel modes are used to analyze and compute fatalities, as shown in Table 10. When people cannot act remotely, they travel more by vehicles which generate accidents.

Table 10. Fatalities (deaths)

	Persons who can act remotely	Persons who cannot act remotely
Fatalities per year	9.4052	26.4987
Total fatalities per year	35.9039	

Emissions of various pollutants are combined and transformed into equivalent carbon dioxide (CO2E) emissions are generated by gasoline and diesel vehicles impact the environment. These various emissions are combined and transformed here into equivalent carbon dioxide (CO2E) emissions as shown in (Liu 2016). Therefore, gasoline vehicles (private autos), public transportation, and commercial trucks are used to estimate environmental impact rates. As shown in Table 11, the total amount of equivalent Carbon dioxide emissions produced per year is $7.4145 * 10^8$ kg.

Table 11. Environmental impact rates

	Persons who can act remotely	Persons who cannot act remotely	Commercial truck
Equivalent Carbon dioxide emissions per year (kg)	$1.9068 * 10^8$	$5.3434 * 10^8$	$1.6426 * 10^7$

Total Equivalent Carbon dioxide emissions per year (kg)	$7.4145 * 10^8$

The total revenue obtained by summing up the revenues obtained from these different socio-economic groups, trip purposes, and modes is $\$3.8951 * 10^7$, as shown in Table 12. Among them, public transportation is assumed to generate 0 revenue here.

Table 12. Total Revenue

	Persons who can act remotely				Persons who cannot act remotely			Commercial Truck
	Private auto -- gas	Private auto -- electric	Virtual	Public transportation	Private auto -- gas	Private auto -- electric	Public transportation	
	$7.8805 * 10^6$	$1.8787 * 10^6$	0	0	$2.1945 * 10^7$	$5.1560 * 10^6$	0	$1.9906 * 10^6$
R_n	$9.7592 * 10^6$				$2.7201 * 10^7$			
R_n (Total Revenue)	$3.8951 * 10^7$							

The variables corresponding to the values in each column are displayed on the first row of Table 13. Among them, the revenue generated from the distance-based tax is the highest. Therefore, the vehicle miles traveled (VMT) generated by the passenger each year significantly affect the total revenue.

Table 13. Summary of major outputs

Distance-based Revenues	Trip-based Revenues	Vehicle ownership Revenues	Total Revenues	Vehicle miles/yr	One-way trips/yr	Number of Vehicles	Equivalent Carbon dioxide emissions per year (kg)	Total fatalities per year
$2.9272 * 10^7$	$7.6634 * 10^6$	$2.0155 * 10^6$	$3.8951 * 10^7$	$3.1426 * 10^9$	$3.4385 * 10^8$	$4.0479 * 10^5$	$7.4145 * 10^8$	35.9039

4.2. Parametric Sensitivity Analyses

In this section, the effects of several influential input parameters on revenues are explored. Those parameters include the distance-based tax, trip-based tax, the utility function coefficient, fraction of users who can act remotely, Potential demand, price in $/person trip, unit costs for electric vehicles, tax rates for electric vehicles, average trip lengths, average vehicle usage, energy efficiency/energy consumption for electric or non-electric vehicles, and E (elasticity to overall weighted impedance) are analyzed.

4.2.1. Sensitivity of total annual revenues w.r.t. multiple input parameters

The sensitivity of the total annual revenue to ten selected sets of input parameters is compared in the Figure 17, clearly showing differences in directions and magnitudes of sensitivity. The total annual revenue is most sensitive to the potential demand level – when the potential numbers of annual one-way trips change by a given percentage, the annual revenue changes in the same proportion. This sensitivity pattern is also observed in the changes of other examined outputs in response to the change in the potential demand level. For this reason, the corresponding graph (green solid line) in this and the following figures can serve as a reference line. Among the other examined input parameter sets, the unit distance-based tax rates, the energy consumption rates, the average trip lengths, and the fractions of people able to work remotely are the ones to which the total annual revenue is most sensitive. This output value is slightly sensitive to the unit trip prices and the "betas" for trip prices. In response to changes of these two sets of input parameters, all examined output values (including this) show overlapping sensitivity curves, because these two parameters sets only appear in utility functions in which they multiply.

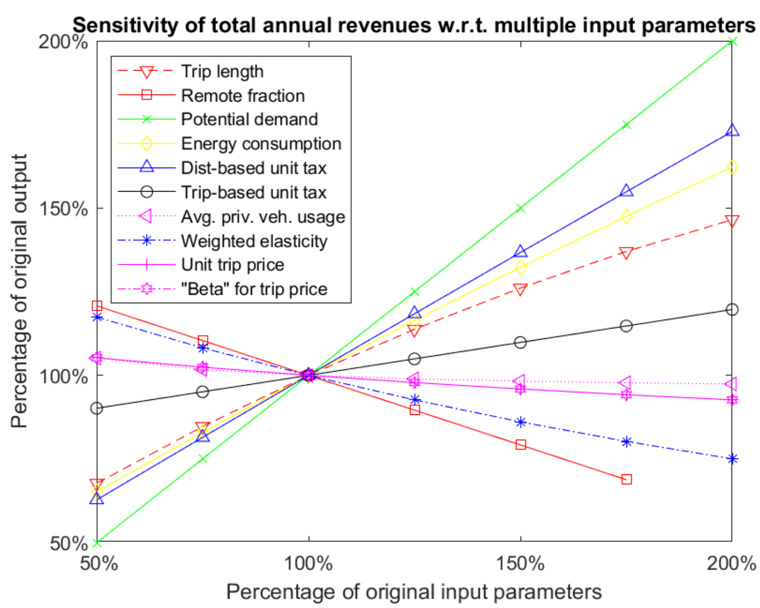


Figure 17 Sensitivity of total annual revenues

(Weighted elasticity: Elasticity to overall weighted impedance, labelled "E")
("Beta" for trip price: Estimated coefficients for trip price in utility functions)

4.2.2. Sensitivity of distance-based/trip based annual revenues w.r.t. multiple input parameters

For the components of total annual revenue in Figures 18 to 20, the total distance-based revenue shows even more pronounced sensitivity to the unit distance-based tax rates, the energy consumption rates, and the average trip lengths. The same changes in these parameters, however, lead to much smaller changes in the total trip-based revenue. As expected, changes in the unit distance-based (trip-based) tax rates have no impact on the total trip-based (distance-based) revenue. The total trip-based revenue varies proportionally with the unit trip-based tax rates, while the total distance-based revenue varies “nearly” proportionally with the unit distance-based tax rates, with some slight bias attributable to the almost negligible impacts of the unit distance-based tax rates on mode shares (as will be shown in Figure 34).

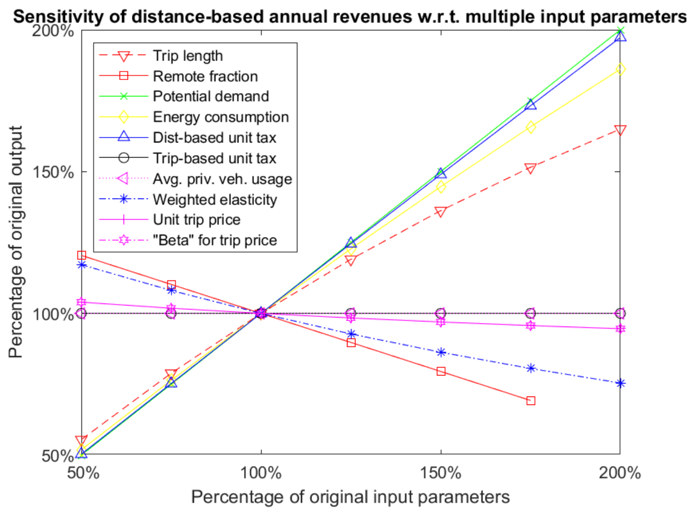


Figure 18 Sensitivity of distance-based annual revenues

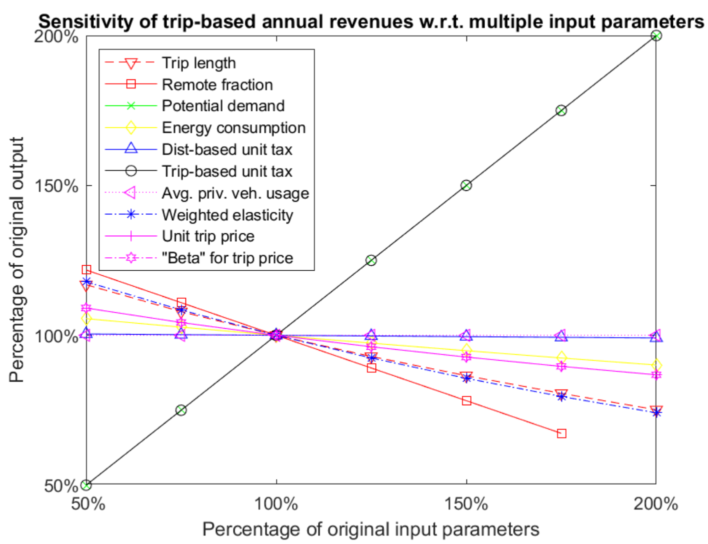


Figure 19 Sensitivity of trip-based annual revenues

4.2.3. Sensitivity of vehicle-ownership-based annual revenues w.r.t. multiple input parameters

For the total vehicle-ownership-based annual revenue, both unit distance-based and trip-based unit tax rates become irrelevant, while changes in the average annual usage per private vehicle result in inversely proportional changes. This stems from the usage of this parameter as a denominator in the model to determine the total number of vehicles. With a very small proportion of vehicle-ownership-based annual revenue in the overall total, this parameter becomes a much weaker decider to the overall total. The ownership-based revenue is also relatively sensitive to the average trip lengths in a nonlinear manner. Many other outputs show non-linear sensitivity to this parameter set because it serves as a multiplier in computing unit trip prices, which are important components of utility values that determine mode shares in a logit model.

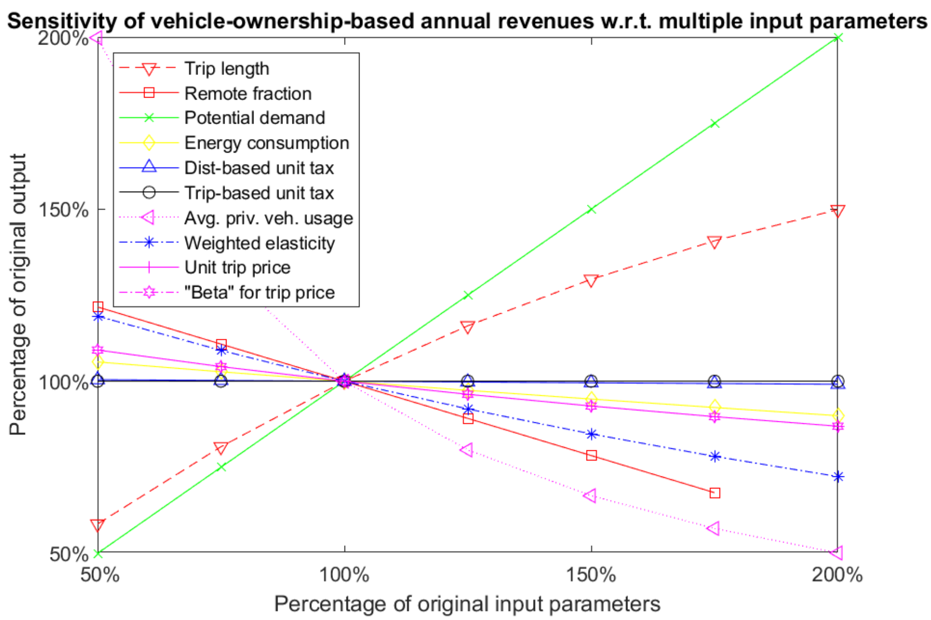


Figure 20 Sensitivity of vehicle-ownership-based annual revenues

The above three components of the total annual revenue are similarly sensitive to the overall weighted elasticity "E" and the fractions of people able to work remotely.

4.2.4. Sensitivity of annual revenues from people able/unable to remote-work w.r.t. multiple input parameters

The two components of the total revenue presented in Figures 21 and 22 are referred to as "Tr1" and "Tr2", respectively. Both Tr1 and Tr2 vary proportionally with the fraction of remote workers, but in different directions. Both components are rather sensitive to the unit distance-based tax rates and the energy consumption rates, while Tr2 is much more sensitive to the trip lengths and the weighted elasticity. The non-linear change of Tr1 with the trip lengths is more noticeable than that of Tr2.

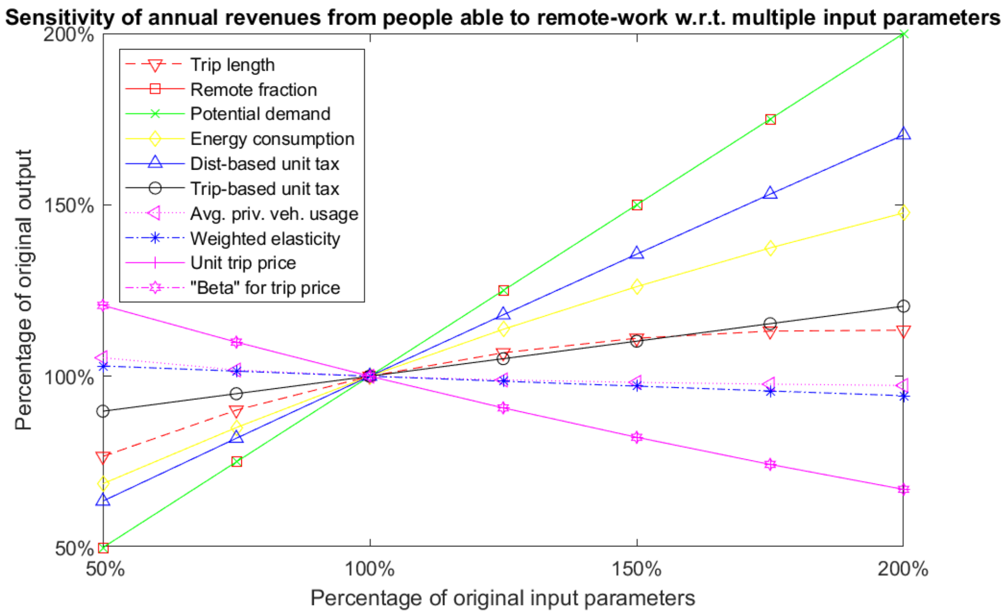


Figure 21 Sensitivity of annual revenues from people able to remote-work

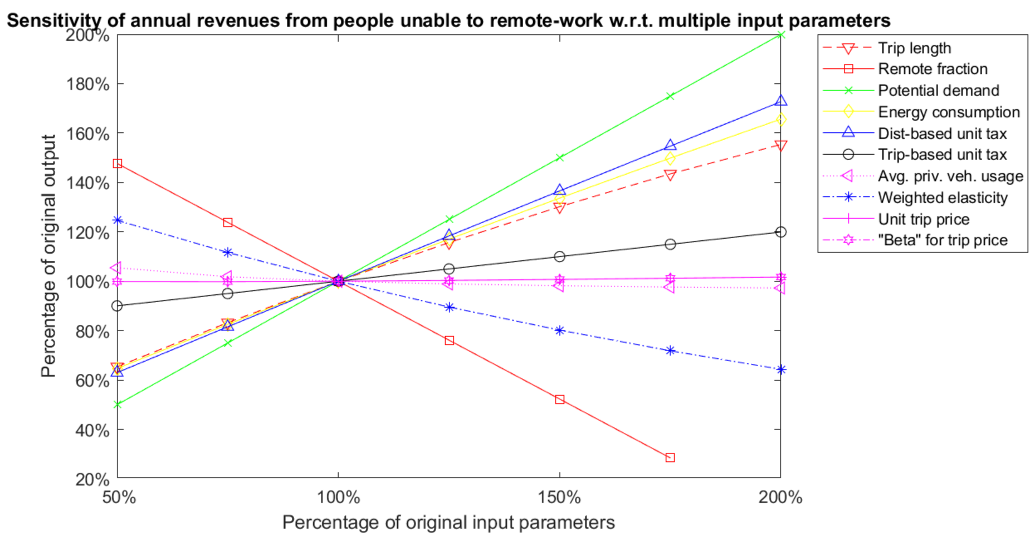


Figure 22 Sensitivity of annual revenues from people not able to remote-work

4.2.5. Sensitivity of total annual one-way trips w.r.t. multiple input parameters

Sensitivity of the total annual one-way trips to all examined input parameters (except the potential demand level) is mostly weak. When these input parameters increase, this output decreases.

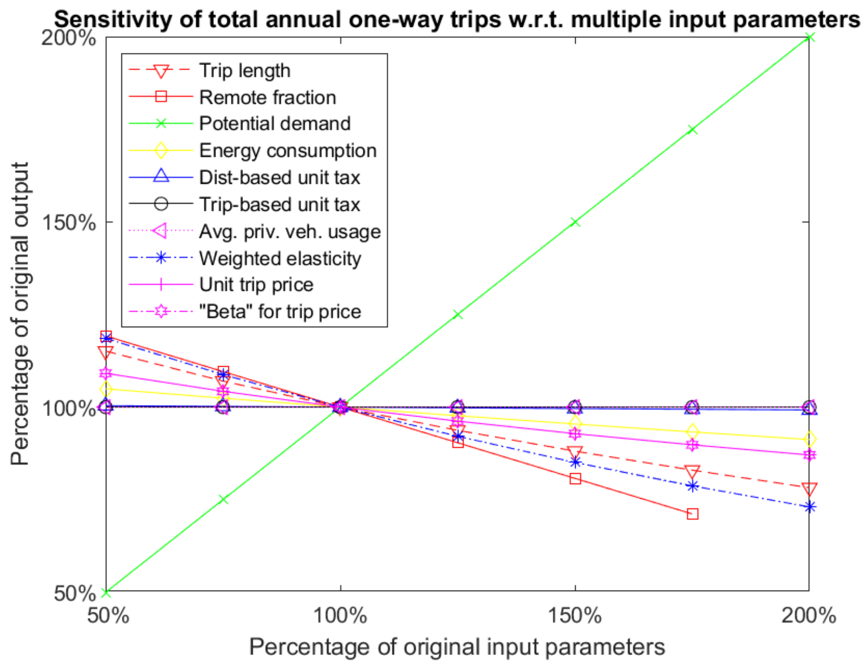


Figure 23 Sensitivity of total annual one-way trips

4.2.6. Sensitivity of annual VMT/number of vehicles/annual fatalities/annual kWh/annual gas consumption/annual CO2 emission w.r.t. multiple input parameters

Figures 24 to 26 are very similar to Figure 20, as the corresponding outputs are similarly sensitive to the examined inputs. Also, the outputs in Figures 27 to 29 are as sensitive to the fractions of remote workers and the weighted elasticity as the outputs in Figures 24 to 26.

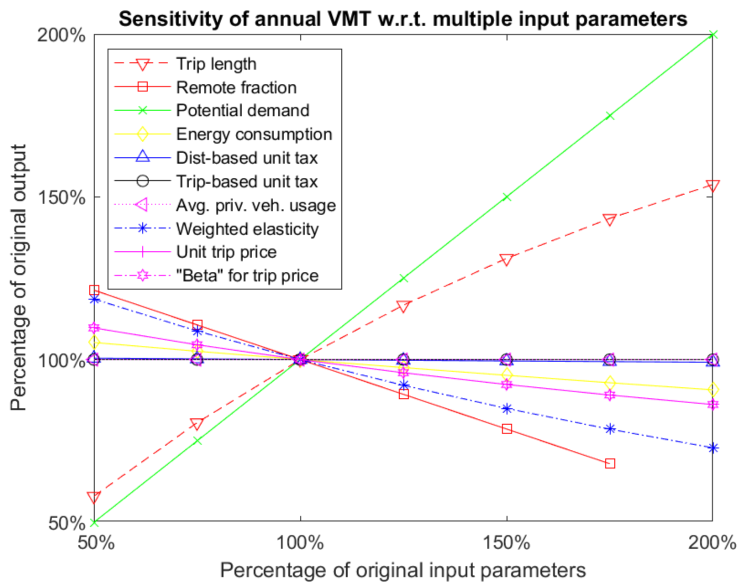


Figure 24 Sensitivity of annual VMT

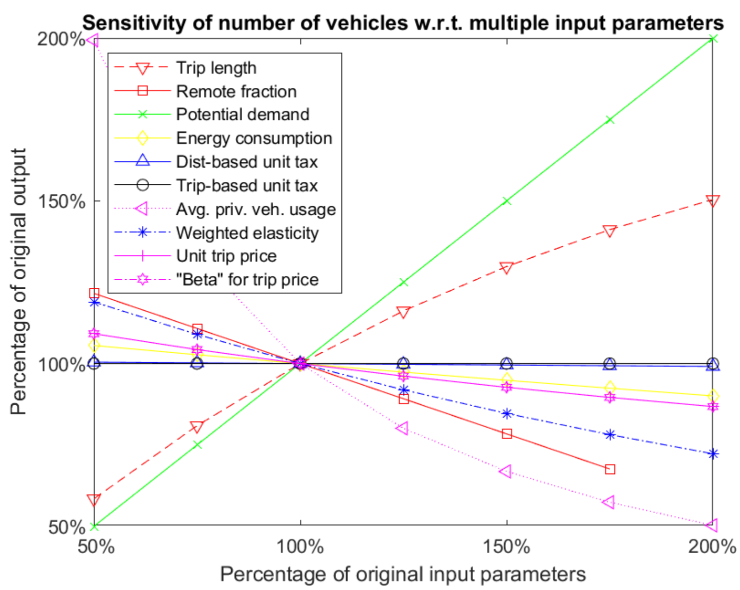


Figure 25 Sensitivity of number of vehicles

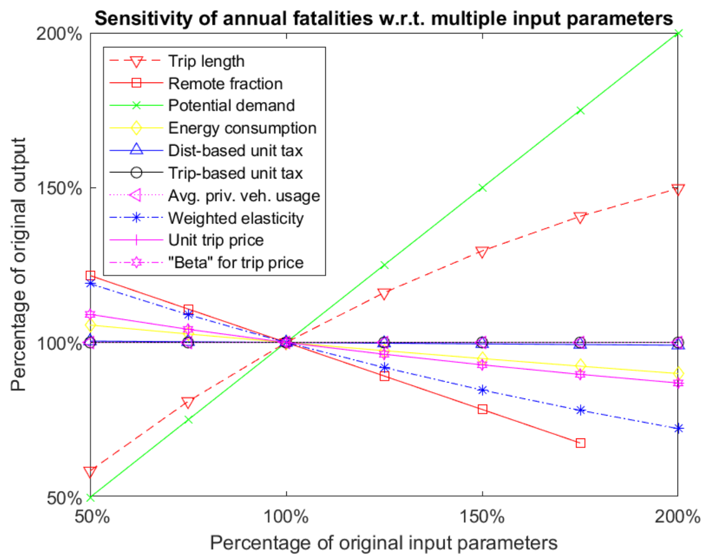


Figure 26 Sensitivity of annual fatalities

The following three energy-consumption-related outputs are rather sensitive to the energy consumption rates. However, they increase less than proportionally with the energy consumption rate due to users' mode shift from EVs and gas vehicles to virtual trips and public transit (as will be shown in Figure 36). The annual carbon dioxide emission and the annual gas consumption, which account for non-EV modes, are more sensitive to the trip lengths than the annual electricity consumption is.

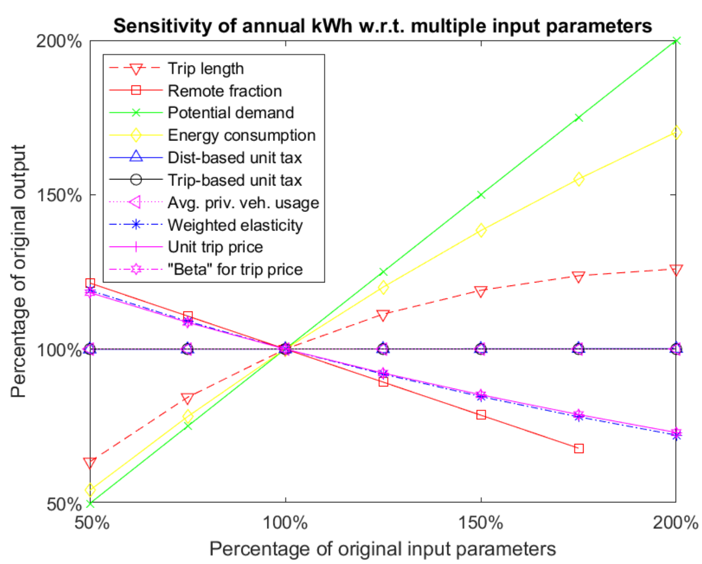


Figure 27 Sensitivity of annual electricity consumption by EVs

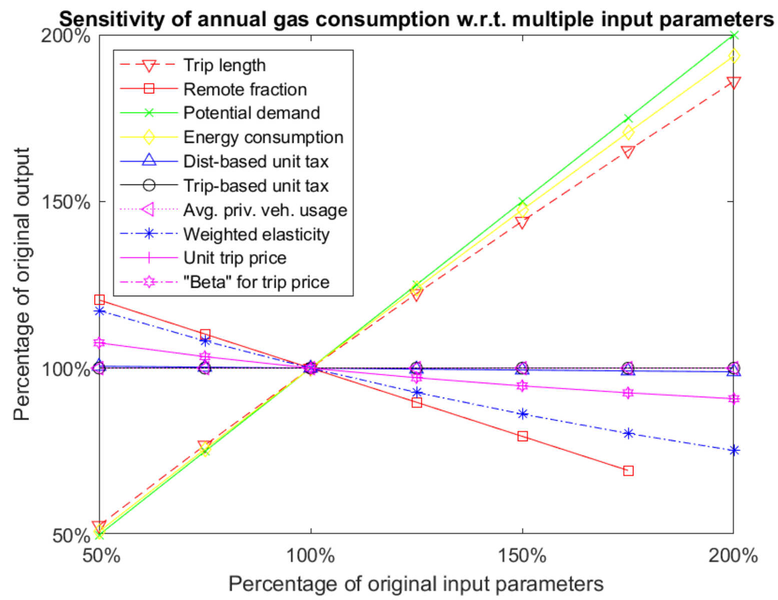


Figure 28. Sensitivity of annual gas consumption

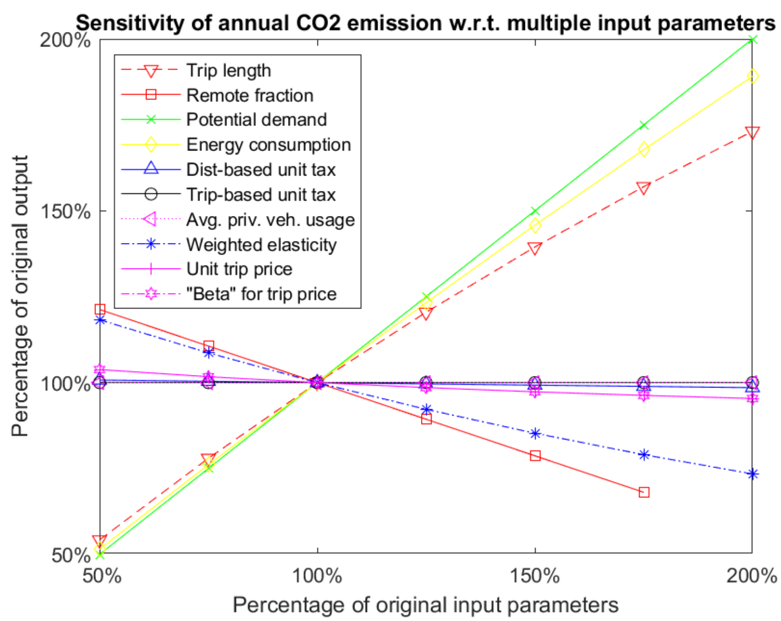


Figure 29. Sensitivity of annual CO2 emission

When the unit energy consumption of all modes changes, the unit CO_2 emission changes proportionally.

4.2.7. Sensitivity of multiple outputs w.r.t. unit costs of EVs/unit tax rates of EVs

The legends for Figures 30 to 33 are explained in Table 14:

Table 14. Explanation of legends

Ddtotal	Total distance-based revenue
Tbtotal	Total trip-based revenue
Tr1	Total revenue from people able to work remotely
Tr2	Total revenue from people unable to work remotely
Tr	Total annual revenue
Vmtotal	Total annual vehicle miles
Nttotal	Total annual number of one-way trips
Nvtotal	Total annual number of vehicles
TotalCO2	Total annual emission of carbon dioxide in kilograms
Kytotal	Total annual electricity consumption by EVs in kWh

The increase in the unit costs (non-fuel cost and energy price) of EVs strongly impacts the mode share of EV and encourages the mode shift to non-EVs, making the annual carbon dioxide emission and the annual electricity consumption highly sensitive to this input parameter. In contrast, the unit tax rates of EVs have little impact on these two output values, which results from the much smaller proportion of energy tax than that of energy price in the composition of EV's trip price.

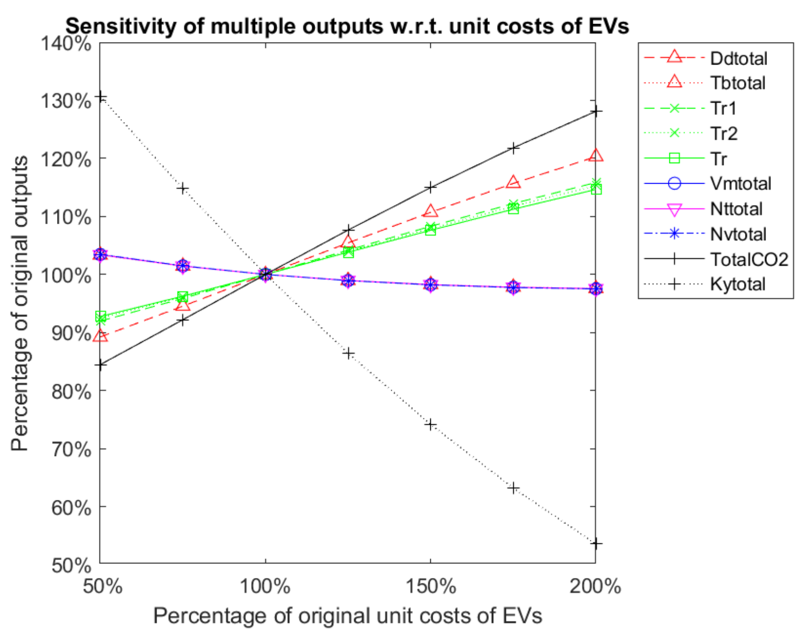


Figure 30. Sensitivity of multiple outputs to unit costs of EVs

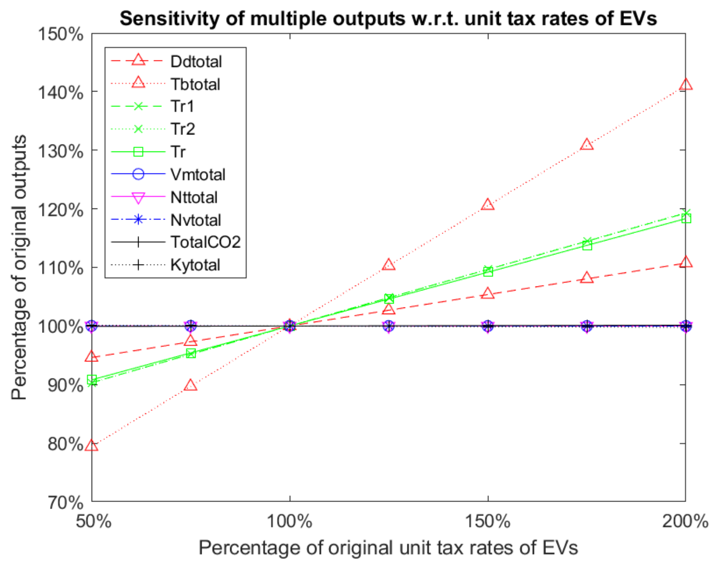


Figure 31. Sensitivity of multiple outputs to unit tax rates of EVs

4.2.8. Sensitivity of multiple outputs w.r.t. unit energy consumption of EVs/non-EVs

The mode shares of EVs and non-EVs are reflected in the annual amounts of electricity consumption and CO_2 emission, respectively. The increase in the energy consumption rate of EVs (non-EVs) slightly discourages the mode choice of EVs (non-EVs) and favors choosing non-EVs (EVs). In Figure 32 (33), the curve corresponding to the annual electricity consumption (annual CO_2 emission) bends downward, which indicates a decreasing mode share of EVs (non-EVs) with the increasing energy consumption rate.

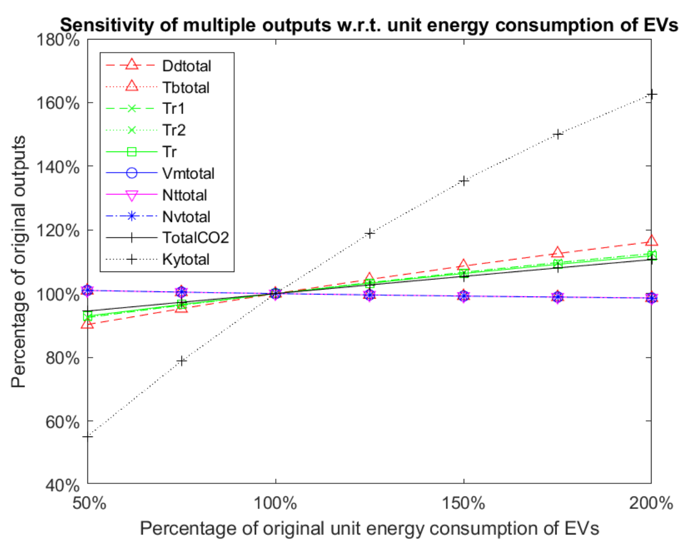


Figure 32. Sensitivity of multiple outputs to unit energy consumption of EVs

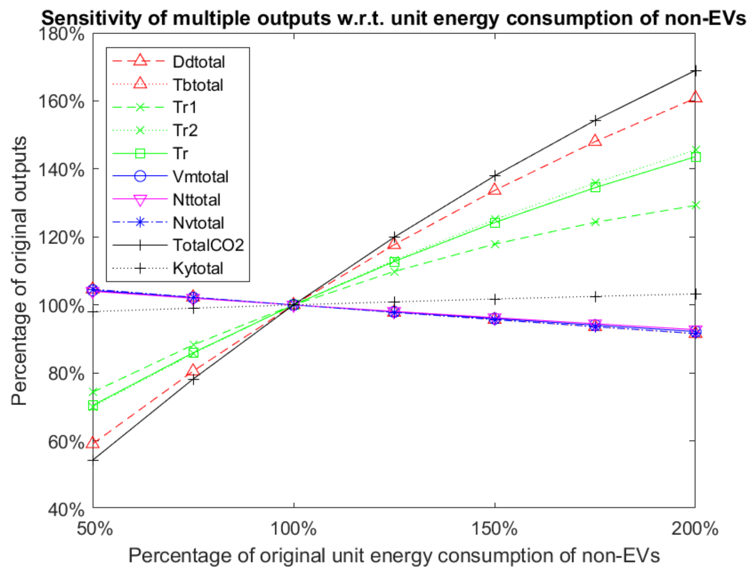


Figure 33. Sensitivity of multiple outputs to unit energy consumption of non-EVs

When the unit energy consumption of non-EVs changes, the unit CO_2 emission changes proportionally.

4.2.9. Sensitivity of mode shares by trip purpose w.r.t. unit distance-based tax rates/ "beta" for trip price

The mode shares by trip purpose are insensitive to changes in unit distance-based tax rates, but sensitive to changes in the estimated coefficients for trip prices in the utility functions. In Figure 35, mode shares for working trips tend to be more sensitive, and the shares of public transit and private EV are more sensitive than those of private gas vehicles and remote working.

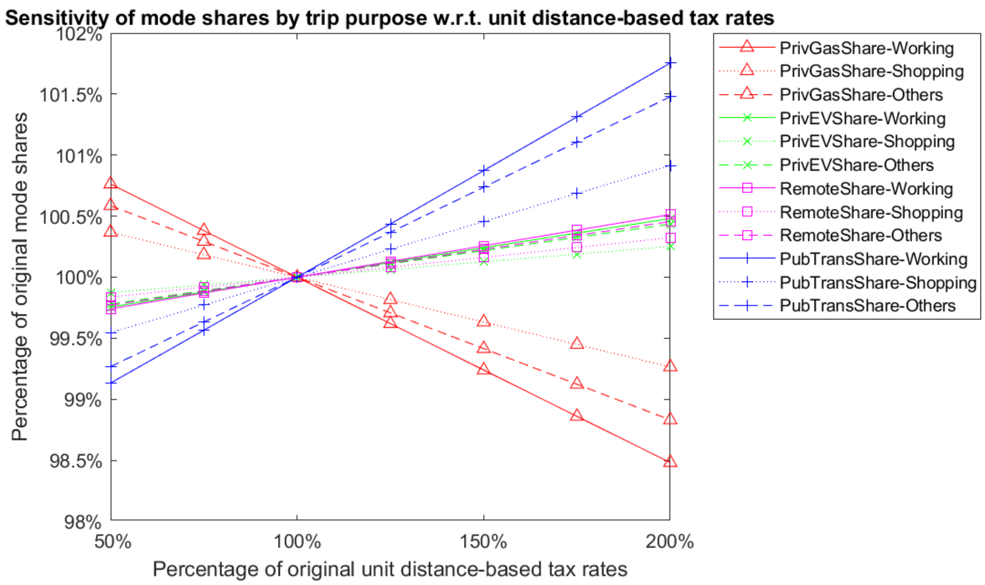


Figure 34. Sensitivity of mode shares by trip purpose to unit distance-based tax rates

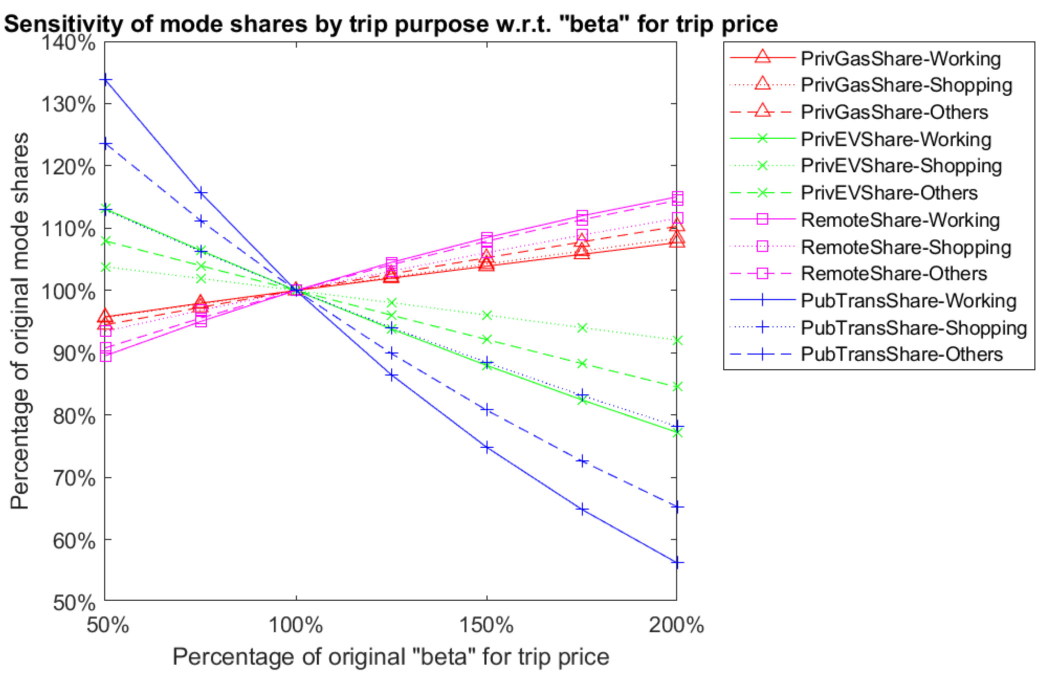


Figure 35. Sensitivity of mode shares by trip purpose to "beta" for trip price

4.2.10. Sensitivity of mode shares by trip purpose w.r.t. energy consumption rates/unit energy costs

It is clearly shown in Figure 36 that an increase in the energy consumption rates, if proportional for all modes, will most strongly discourage users' choice of EV and drive more users to utilize public transit than to work remotely. Mode shares for work trips are also more sensitive than those for other trip purposes. Similar sensitivity trends are shown in Figure 37.

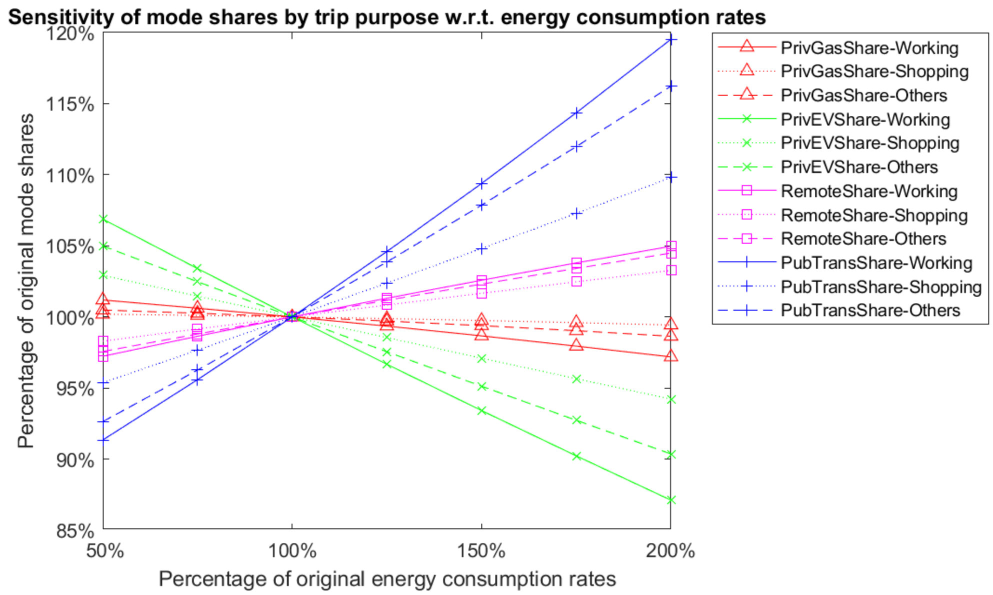


Figure 36. Sensitivity of mode shares by trip purpose to energy consumption rates

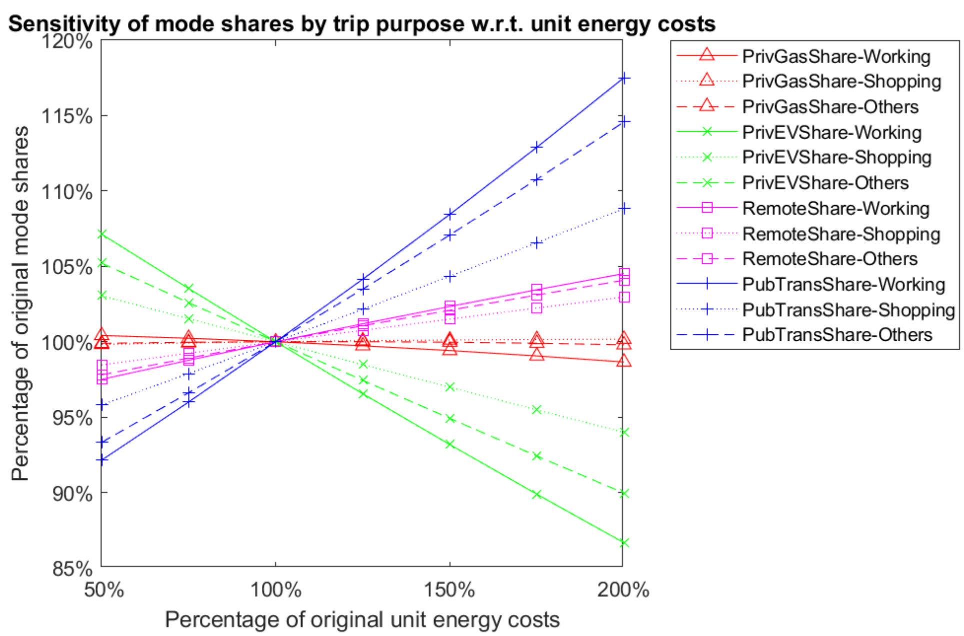


Figure 37. Sensitivity of mode shares by trip purpose to unit energy costs

4.2.11. Sensitivity of multiple outputs w.r.t. available number of EVs

In previous results it is assumed that sufficient EV's are available for the estimated mode shares. In Figure 38 some effects of having fewer EV's in the fleet mix than the equilibrium level are explored. The EV availability is expressed as a percentage of original equilibrium level. Users who prefer EV's but cannot have them must then use gasoline cars. Each 10% decrease of the available EV numbers (from the expected base value) raises the total tax revenue by approximately 3.5% (from the base value).

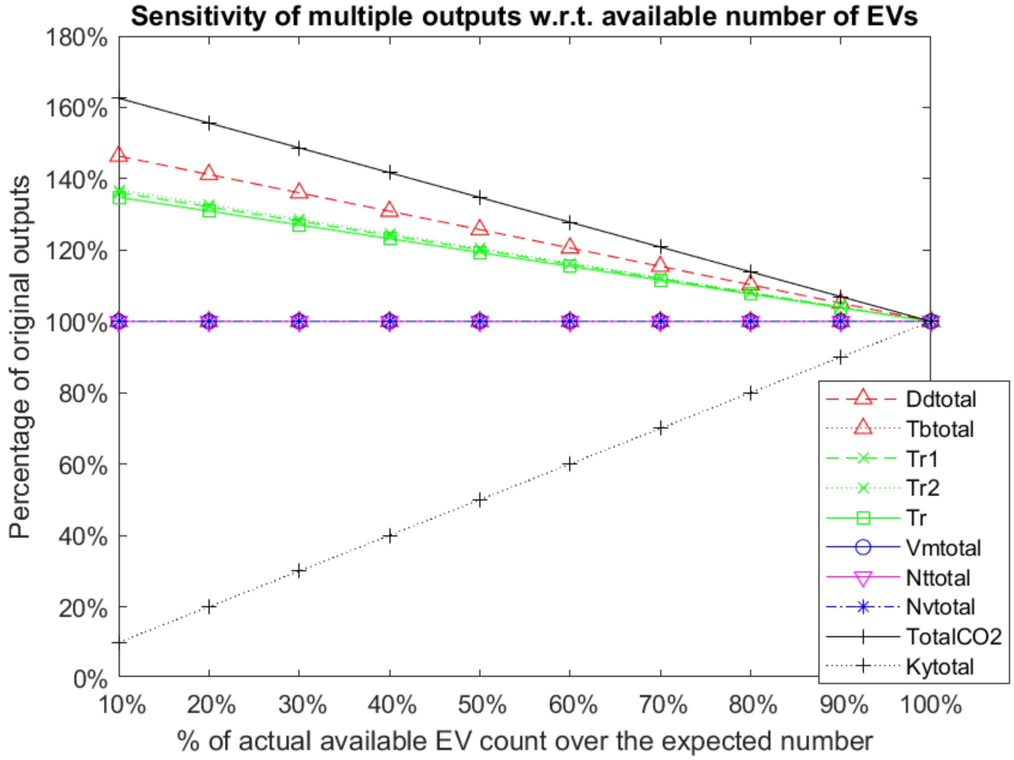


Figure 38. Sensitivity of multiple outputs to actual available EV amount

4.2.12. Change rates of selected outputs under future scenarios

The table below shows change rates of examined outputs under various potential future scenarios where multiple input parameters increase. It can be noted that, with a higher multiple of the original unit energy tax rate, four vehicle-mile-related outputs (total VMT, total one-way trips, total number of vehicles, and total CO_2 emissions) decrease slightly, while the total electricity consumption has a slight increase. However, if the unit energy costs increase all examined outputs in Table 15 (including the revenue and its components) decrease slightly to moderately, reflecting the shift of mode choice towards public transit and remote activity, as shown in Figure 37.

Table 15. Change rates of selected outputs under future scenarios

Potential demand chg. rate	Energy cost chg. rate	Unit tax change rate	**Ddtotal change rate (%)**	**Tbtotal change rate (%)**	**Tr1 change rate (%)**	**Tr2 change rate (%)**	**Tr change rate (%)**	**Vmtotal change rate (%)**	**Nttotal change rate (%)**	**Nvtotal change rate (%)**	**Kytotal change rate (%)**	**TotalCO₂ change rate (%)**
+25%	0	+100%	146.62	147.80	144.72	147.47	146.91	23.98	24.05	23.91	25.16	23.16
		+200%	264.91	268.44	259.32	267.43	265.79	22.97	23.10	22.82	25.31	21.34
		+300%	379.95	386.96	368.93	384.92	381.68	21.97	22.17	21.75	25.45	19.55
		+400%	491.78	503.38	473.72	499.95	494.66	20.98	21.25	20.69	25.58	17.79
	+50%	+100%	139.49	136.12	127.31	141.88	138.64	18.48	18.87	18.06	15.38	20.64
		+200%	254.42	251.08	233.74	259.13	253.58	17.52	17.97	17.03	15.55	18.90
		+300%	366.22	364.00	335.54	373.94	365.65	16.58	17.09	16.01	15.70	17.18
		+400%	474.94	474.93	432.83	486.35	474.92	15.64	16.21	15.00	15.85	15.49
	+100%	+100%	132.66	125.21	111.13	136.45	130.80	13.37	14.05	12.61	6.41	18.21
		+200%	244.37	234.84	209.97	251.05	241.98	12.46	13.20	11.62	6.58	16.54
		+300%	353.05	342.53	304.50	363.26	350.41	11.55	12.36	10.64	6.74	14.90
		+400%	458.78	448.30	394.84	473.09	456.14	10.66	11.52	9.68	6.90	13.27
+50%	0	+100%	195.94	197.36	193.67	196.96	196.29	48.77	48.86	48.69	50.19	47.79
		+200%	337.89	342.13	331.18	340.92	338.94	47.56	47.73	47.39	50.37	45.61
		+300%	475.93	484.35	462.72	481.90	478.02	46.37	46.61	46.10	50.54	43.47
		+400%	610.14	624.05	588.46	619.94	613.59	45.18	45.50	44.83	50.69	41.35
	+50%	+100%	187.38	183.35	172.78	190.25	186.37	42.18	42.64	41.67	38.46	44.76
		+200%	325.30	321.30	300.49	330.95	324.29	41.03	41.57	40.44	38.66	42.68
		+300%	459.46	456.80	422.65	468.73	458.78	39.89	40.50	39.21	38.84	40.62
		+400%	589.93	589.91	539.40	603.62	589.90	38.77	39.45	38.00	39.02	38.59
	+100%	+100%	179.19	170.25	153.36	183.74	176.96	36.04	36.86	35.13	27.69	41.85
		+200%	313.24	301.81	271.96	321.26	310.38	34.95	35.84	33.94	27.90	39.85
		+300%	443.67	431.03	385.40	455.91	440.50	33.86	34.83	32.77	28.09	37.88
		+400%	570.54	557.95	493.81	587.71	567.37	32.79	33.82	31.61	28.28	35.93

5. Study on gas tax and price/ methods for forecasting VMT

Vehicle miles traveled (VMT) is a measure used in transportation planning for a variety of purposes. It measures the amount of travel for all vehicles in a geographic region over a given period, typically a one-year period. VMT is calculated by adding up all the miles driven by all the cars and trucks on all the roadways in a region (Thomas et al., 2016). As shown in figure 6, the conventional forecasting method for VMT should cover: ***Internal/Internal trips.*** These represent the fraction of trips generated by a mixed-use development that both begin and end within the development. The importance of internal trip capture is that those trips satisfy a fraction of the total development's trip generation and they do so without using the external road system (Brian and Benjamin, 2010). ***External/External trips*** -

Trips passing through the study area or city (starts and ends outside the study area). ***Internal/External trips*** -Trips that start from the study area and end outside the study area. ***External/Internal trips*** -Trips that start outside the study area and ends in the study area.

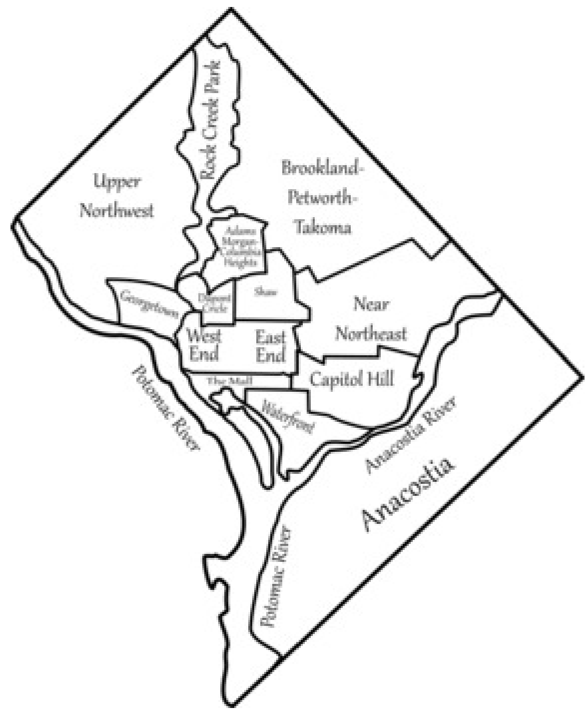


Figure 39. Type of trips in DC

5.1. Factors that influence VMT

Many factors influence travel demand. Since VMT is a measure of travel demand, understanding the factors that influence VMT provides a greater comprehension of VMT and the issues that affect its estimates and forecasts. VMT forecasting can be a difficult and often involves complex processes (Szekeres, Koppula and Frazier 2007). The influencing factors are wide ranging, and their level of influence varies. Factors affecting VMT forecasts include "socio-economic and demographic growth, changes in the cost of travel, urban sprawl, technological innovation, social change, and legislative factors. (Steven, Xuehao and Lavenia 2004) also note that some trends emerging from their work showed that VMT growth may be moderating by increasing trip lengths and travel time budgets. Many of these factors are typically incorporated into the different stages of the travel demand forecasting process and the various statistical models that were developed to estimate and forecast VMT. (Steven, Xuehao and Lavenia 2004)articulated this conceptual relationship, as shown in Figure 39.

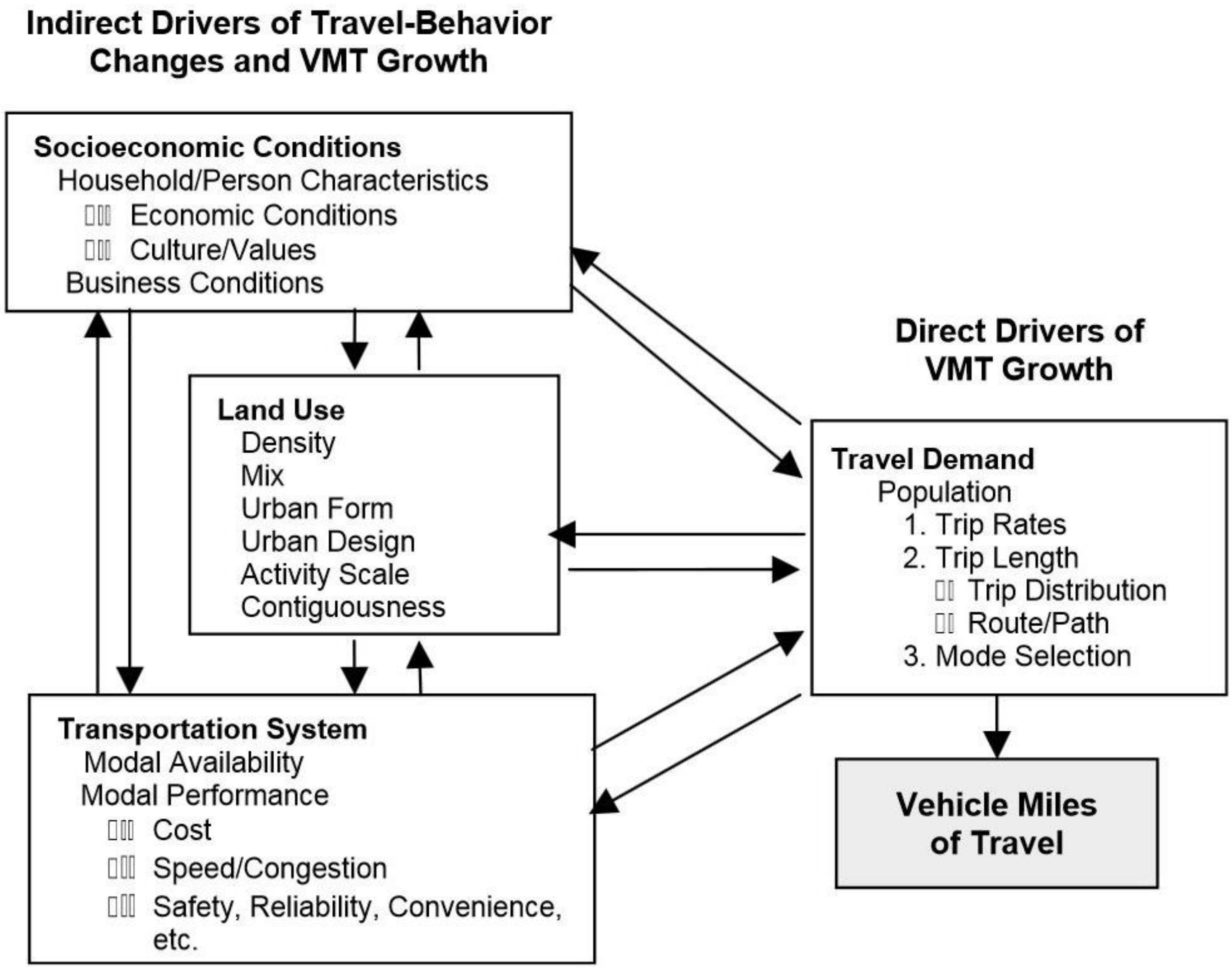


Figure 40. Direct and Indirect drivers of VMT

5.2. Traditional VMT Forecasting Methods

VMT estimating and forecasting methodologies can essentially be placed into one of two categories i.e. Data-based method and Forecasting methods. The data-based comprises:

5.2.1. Traffic count-based Method

Traffic-count based methods are the mostly used forecasting VMT (Feng, et al. 2006). Traffic count methods use both continuous and coverage count data by counting traffic on a particular roadway section for a short period of time. Both types of count data are converted to annual counts using various expansion factors that include seasonal factors and time-of-day factors that may vary by functional class. These methods are typically designed for statewide use, meaning that they may not be statistically valid for smaller units of analysis such as counties or regions. Count-based methods assume that VMT growth will be similar to that of the past. Essentially, the methods do not consider any changes that might occur in land use patterns or any changes in socioeconomic data (Feng, et al. 2006) (Lei and Xiang 2013).

5.2.2. Socioeconomic-Data-Based Methods

Socioeconomic-data-based methods do not rely on any characteristics of the roadway but rather on those factors that affect travel behavior. According to (Feng, et al. 2006), these methods attempt to estimate and forecast VMT growth at a more fundamental level than the growth factor method by using variables that can be projected into the future. Examples of these types of approaches include the use of travel surveys to estimate VMT based on household travel characteristics and licensed driver data, a fuel-sales-based approach used to estimate and forecast VMT, and an approach that relies on odometer readings.

5.2.3. Travel Demand Forecasting Models

The use of travel demand models to forecast VMT combines both traffic count data and socioeconomic data (Thomas, Williams and Brianne 2016). This approach considers travel behavior and other factors that affect VMT growth. Thus, what is created is a network biased toward larger roads, which often limits the number of smaller or local roads in the region. However, calibrating these models can improve the accuracy of the VMT estimates. Usually, a post- processing method to account for non-modeled roadway VMT is added for more localized travel.

This method is suitable for forecasting VMT by functional class, area type, and different jurisdictions such as different counties and regions (Robert and Jon 2007, Feng, et al. 2006). This method may also "provide greater sensitivity to changes in transportation investments or policies compared to many manual calculation procedures" (Szekeres, Koppula and Frazier 2007). However, it is not without concerns. Travel models are not typically statewide in their scope, nor do they account for local travel since local roads are not accounted for on the travel network. The lack of local roads and travel accounted for in these models makes them suspect for statewide VMT forecasting, with biases toward larger roadways and a tendency to overestimate travel demand (Robert and Jon 2007).

One of the major limiting factors in using travel models to forecast VMT is the time and effort needed to produce the forecast. These can be significant and often limit the responsiveness needed to produce VMT forecasts, particularly when comparing scenarios or other alternatives that may require several models to run. However, travel models offer a method that includes the network geographic component—roadway miles—that is fundamental to VMT estimation and forecasting.

5.2.4. Enhanced and Specific Forecasting Methods

Efforts towards improving the accuracy of VMT forecasting have resulted in initiating specialized VMT forecasting methods which include:

Area Roads VMT: Area or community road VMT refers to VMT that occur as a result of localized travel that is not typically accounted for in forecasting models or as a part of traffic-counting programs. Usually, this includes residential streets and commercial parking areas.

Commercial Vehicle VMT: This was done to better estimate commercial-vehicle-related crash rates at the state level. The adjusted state VMT for commercial vehicles supports measurement of program effectiveness and development of countermeasures to promote motor carrier safety." It essentially has a safety purpose that enables state and federal agencies to better focus their safety and enforcement resources (Blake, et al. 2010).

From Table 12, the percentage of gas bought outside DC from 2017 to 2021 remains high ranging from 85.90% to 129.31%. This is regardless of increased consumed gasoline VMT-based in DC. As shown in Figure 8 from 2012 to 2022, gasoline price in DC remained constantly higher than gasoline prices in Virginia and Maryland. Thus, these two states provide gas purchase alternatives to DC, depending on trip origins and destinations.

Table 16. VMT vs Gasoline sales in the District of Columbia

Year	VMT Per Veh.	Active Registered Veh.	Average Gas $/Gal	Tax Rate $/Gal	Av. Mile per gasoline	Consumed Gasoline (VMT-based) (Gal)	Gasoline Sold in DC (Gal)	Gas Tax collected based on VMT	Gas Tax collected based on Gas sold in DC	Gasoline bought outside DC	%Gasoline bought outside DC (Gal)
2017	5300	310,465	2.7	0.279	13	126,574,192	68,088,000	35,314,200	18,996,552	58,486,192	85.90
2018	5261	310,052	2.9	0.235	16	101,948973	72,789,000	23,958,009	17,105,415	29,159,973	40.06
2019	7013	358,963	2.37	0.235	17	157,337970	70,662,000	36,974,423	16,605,570	86,675,970	122.66
2020	6650	358,963	2.5	0.288	17	140,417,879	57,126,000	40,440,349	16,452,288	83,291,879	145.80
2021	7000	376,911	3.4	0.288	18	146,576,500	63,921,000	42,214,032	18,409,248	82,655,500	129.31

Sources: 1) NHTSA National Center 2019, 2) US Department of Energy, Energy Information Administration, State Energy Data System, 2019, 3) American Petroleum Institute, 2018, 2019, 2020

5.3. Effects of Gasoline Price Increase on VMT

Empirical research suggests that total driving, or vehicle miles travelled (VMT), is not very responsive to the price of gasoline in the short-term. CBO's analysis of the influence of gasoline prices on motorists' behaviors is based on four years of data collected from metropolitan freeways in California between 2003 and 2006, and on statewide average gasoline prices and wages over that period. A 10 percent increase in gasoline prices is estimated to reduce VMT by as little as 0.2 percent to 0.3 percent in the short run and by 1.1 percent to 1.5 percent eventually (Kenneth, Small and Van 2007). A 2003 study of corporate average fuel economy (CAFE) standards, published by the National Research Council, cited slightly older estimates of the responsiveness of VMT to gasoline prices that ranged from about 1 percent to 2 percent (NRC 2002).

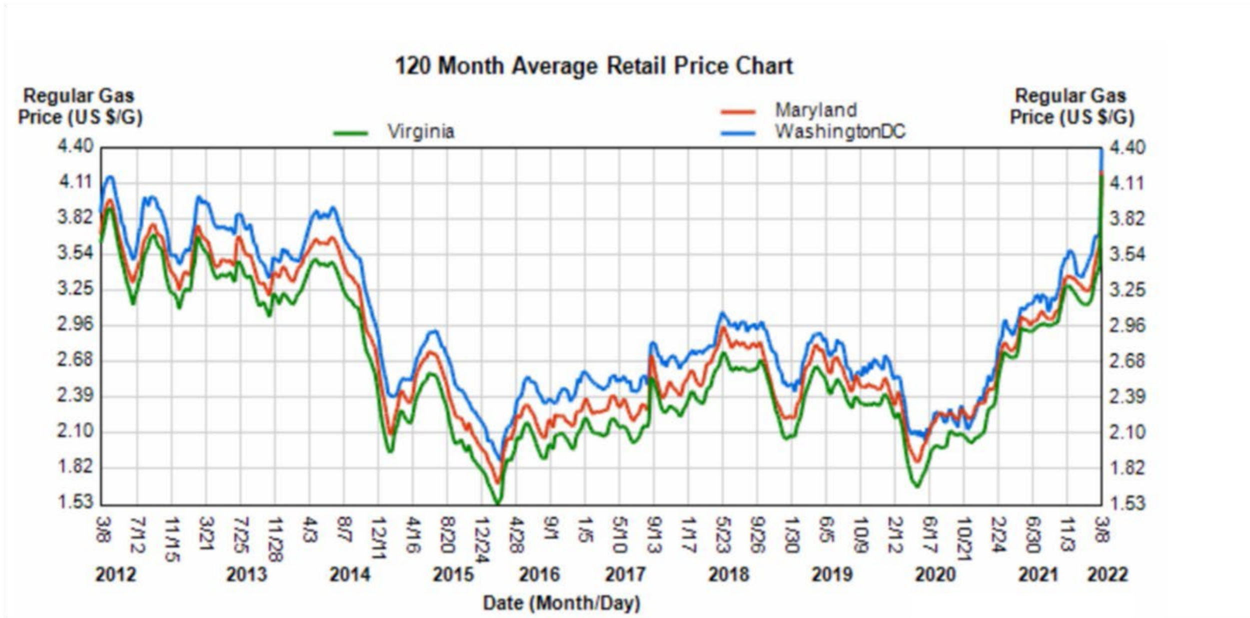


Figure 41. Gas Price Comparison

However, as gasoline prices increased, a study also identified that many transit agencies have claimed that higher gasoline prices have driven ridership growth. It is determined first whether such a correlation is substantiated by the available data and then, if such correlation exists, the nature of this relationship. Five U.S. cities were selected for analysis based on their level of automobile orientation and the extent and variety of transitservices: Atlanta, Georgia; Dallas, Texas; Los Angeles, California; San Francisco, California; and Washington, D.C (Ashley and Randy 2007). Most of the transit systems in the five cities analyzed have experienced ridership growth since early 2004. Exceptions include the Atlanta bus and heavy rail systems and the San Francisco bus systems.

With the use of time series analysis, seasonal indices, and correlation coefficients, ridership trends are evaluated and compared with corresponding national gasoline prices. This indicates that gasoline price increases have indeed played a role in encouraging transit use in historically automobile-oriented American cities. Finally, the empirical relationships between gasoline price and transit demand are explored. Results indicate that, on average, as gasoline price increases by 1%, transit demand increases on the order of 0.24%; in other words, ridership increases approximately 0.09% for each $0.01 increase in gasoline price.

When faced with an increase in gasoline prices, motorists would most readily curtail their lowest value trips.

If motorists consider weekend trips generally less important than weekday trips, then

weekend traffic volumes should be more sensitive to the price of gasoline.
Urban economists examine the connection between gas prices and driving in the U.S. over the last two decades. Prices matter: increased gas prices result in decreased driving, providing the prices persist for the long-term (Irvin 2019). People make decisions about where to live, how far they're willing to commute to work, whether to own a car (or a second car), and whether to use various other modes (cycling, transit, and walking) on a long-term basis. Over time, prices influence all these decisions.
Historically, the four phases of gas prices and the response of vehicle travelled in America as provided by (Irvin 2019) is shown below:
Phase 1: 2000 through 2004 (4 years): Era of cheap gas and rising VMT: Peak driving' in the U.S. was in June 2005, when Americans drove 27.7 miles per person per day. At the time, gasoline cost an average of about $2.13 per gallon.
Phase 2: 2005 to July 2014: (9+ years): Expensive gas era. "By 2013, the typical American was driving about 25.7 miles, more than 2 miles per person per day less than at the peak."
Phase 3: 2014-2016 (2 years): Gas prices began falling in July 2014. After OPEC met in Vienna on November 27 and decided that maintaining market share preempted stabilizing gas prices, prices plummeted. "In April 2014, gas prices averaged more than $3.70 a gallon, and people drove an average of 25.7 miles per day. Some 22 months later, in February 2016, with prices averaging about $1.75 a gallon, consumers were driving about 26.7 miles per day, about 4 percent more."

Phase 4: 2016-2018 (2 years): A rebound in gas prices: Sustained high prices for gasoline lead to real reductions in vehicle miles travelled, in pollution and in car deaths. If we price travel appropriately, consumers will make different decisions—ones that significantly reduce the social and environmental costs of car travel. Prices matter and should be at the heart of all of our efforts to cope with climate change and build stronger and safer communities. A brief analysis by the State Smart Transportation Initiative found a weak connection between VMT and gas prices, and a rather strong one between vehicle miles and urban density.

5.3.1. Effects of Gas Price Increase on Electric Vehicle Demand

Cars over the next decade are going to flip from mostly gasoline-powered to electrified. Market forecast shows that electric vehicles would grow from 3% of sales in 2021 to 32.3% by 2030 whilst gasoline-powered vehicles will shrink from 87% of sales to 36.5%. In Greater DMV, an interest has surged since the rise in gas price. Search interest in electronic vehicle is up more than 300% since the recent increase in gas price.

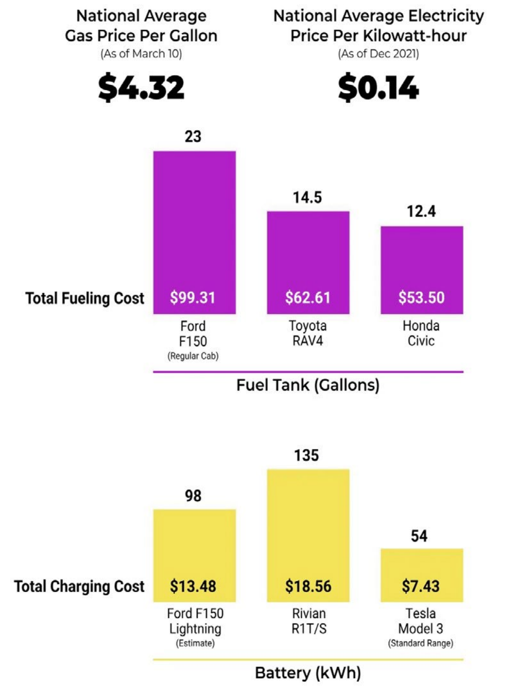


Figure 42. National Average cost of Gasoline and per-kilowatt

The implication of the above on DC is continuous decrease in gas tax providing a strong justification for the implementation of VMT tax. Such decrease in gas tax collections will also result in an irrecoverable gap in the state revenues for transport project funding like the case of Oregon.

5.4. Factors Affecting Vehicle Miles Travelled Implementation

Qualitative methods can be very effective for obtaining a detailed understanding of what peoples' opinions are, as well as nuances about why they hold those opinions (Luntz,1994). Therefore, the factors in Table 17 below summarize qualitatively the factors that affect Vehicle Miles Travelled tax implementation.

Table 17. Factors that affect VMT implementation

Non-financial	**Financial**	**Factors related to System design**
Privacy	Sustainable source of revenue	Data Security
Communication	Cost-effectiveness	Accountability

Fairness in billing		System Flexibility
Complimentary policy objectives		Interoperability and cooperation
Simplicity		Phasing
Enforcement		Users' options

5.5. Privacy

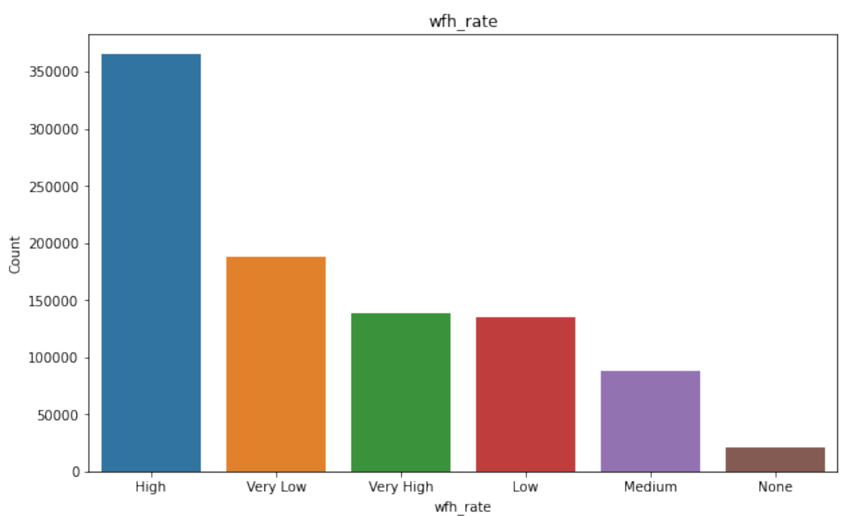
Some motorists are concerned that VMT charging could be an invasion of their privacy if location information is utilized, as any data collection system poses a risk to private information of users (Badger 2011). Oregon's 2012 VMT fee pilot study offered five plans, each with a different technology option and payment method depending on the drivers' privacy preferences. Drivers had the choice to report miles using a smartphone, a global positioning system (GPS) device, or a simple reporting device with no GPS technology; or they could opt out of using technology altogether by paying a flat rate in lieu of a per- mile fee (ODT, Oregon Department of Transportation "Legislative Report, Road Usage Charge Pilot Program Preliminary Findings" Oregon Department of Transportation. 2014). The Puget Sound Regional Council encountered similar privacy concerns as a result of its 2002 road tolling demand response study, which monitored participants' mileage on certain types of roadways in a similar manner (Puget Sound Regional Council 2008).

6. A Comparative Study on Telework for the Washington DC Metropolitan Area Using Machine Learning Algorithms

As a result of the 2020-2022 pandemic interest in teleworking is growing. Employees, employers, company managers, city planners, transportation planners and individuals are affected by the growth of telework. Various factors are involved in an individual decision to telework instead of physically going to a work location. Using the existing data from 2017 and 2022, in the form of two case studies, the most influential factors that affects peoples' decision to telework are identified. A comparison between these data is performed to highlight the impact of Covid-19 pandemic on changing these factors. Machine learning is the practice of using algorithms to parse data, learn from it, and make a determination or prediction which has been employed in various fields. In this study, using Machine Learning, several conceptual models are examined including Logistic Regression, Support Vector Machine (SVM), Decision Tree, Random Forest, K-Nearest Neighbor, and Multinomial Logit. Data analysis is carried out in Python using JupyterLab to examine the correlation matrix, validate the models, and determine the feature importance. Many studies have been conducted on analyzing the level of telecommuting based on the factors that affect people's decision to work from home using conventional statistical methods, however, only few studies carried out on data driven approaches. Therefore, this section focuses on predicting telecommuting by using data driven Machine Learning algorithms that proposes the most influential variables such as socioeconomic, environmental and transit factors. The developed predictive models are adaptive and can be updated using new data sets obtained from different geographical locations to estimate telecommuting behavior in various circumstances.

6.1. Fraction of work trips that can be substituted by telework:

The level of work from home is shown in the Replica data by the variable wfh_rate which is a categorical variable and as 5 categories, None, very low, low, medium, high, very high. To estimate the portion of working trips that can be substituted by telework, we will add the count of the categories in which the work from home rate is high or very high, and to get the percentage we divide this number by the total number of the household in our data set. The chart below represents the distribution of each category.



Level of Telework	Count
High	364977
Very Low	188372
Very High	139072
Low	134601
Medium	87899
None	21540

Figure 43. Distribution of Level of Telework in 2022

To obtain the percentage of the people who have the option to work from home, the number of frequencies when the level of work from home are high and very high are added and being divided by the total number of frequencies. The percentage of the working trips that can be substituted by telework then is estimated as 54%.

6.2. Methodology

Machine learning is widely used in many fields throughout the world including the healthcare sector, transportation, advertisement, economics, and image recognition. Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed (Mozaffarian 2015). Further, machine learning at its most basic is the practice of using algorithms to parse data, learn from it, and then make a determination or prediction about something in the world (Das, et al. 2015). There are two major categories of problems often solved by machine learning i.e., regression and classification. Mainly, the regression algorithms are used for numeric data and classification problems include binary and multi-category problems (Abduljabbar 2019). Machine learning algorithms are further divided into two categories such as supervised learning and unsupervised learning (Strecht, et al. 2015). Basically, supervised learning is performed by using prior knowledge in output values whereas unsupervised learning does not have predefined labels; hence its goal is to infer the natural structures within the dataset (Sathya and Abraham 2013). Therefore, the selection of machine learning algorithms should be carefully evaluated. In this study, the supervised machine learning algorithms are used and the process for the model development is explained as follow. The methodology followed in paper is shown in Figure 43.

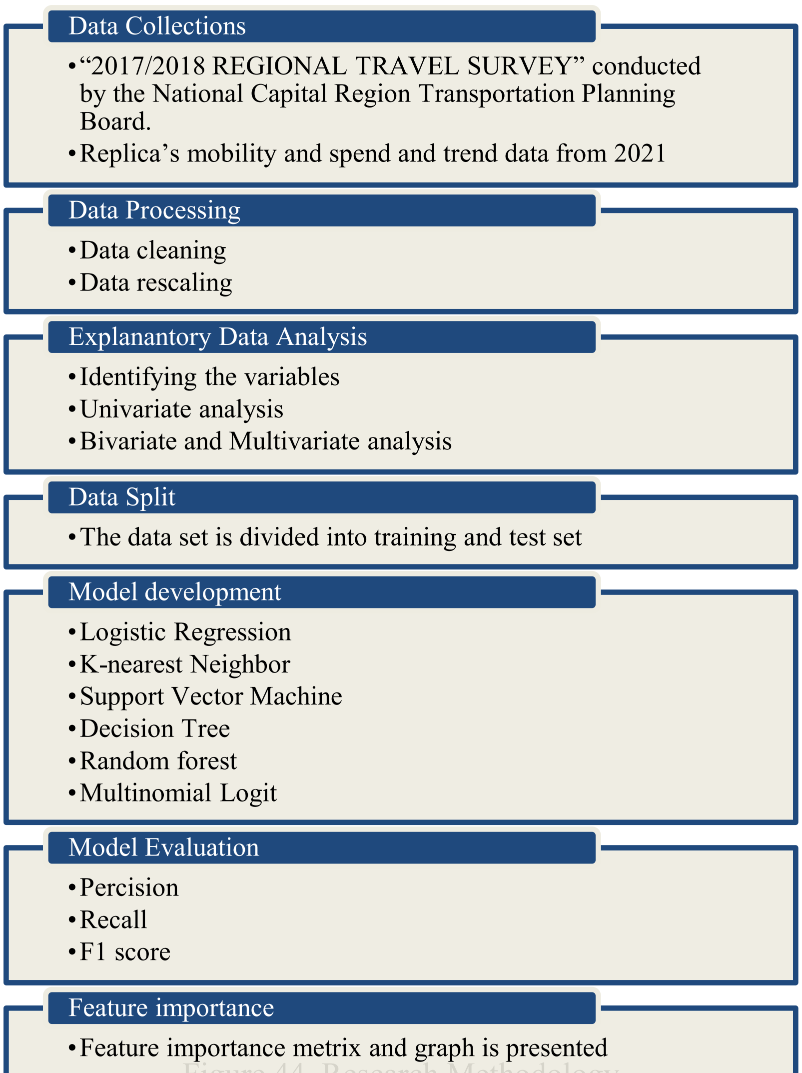


Figure 44: Research Methodology

6.3. Data Collection:

Two compare the level of telework from before and after pandemic, two data sets are collected which are described in below.

First Dataset:

The data set which is used for our first case study is the “2017/2018 REGIONAL TRAVEL SURVEY” conducted by the National Capital Region Transportation Planning Board. This data set is available to the public (TPB 2021). In fulfilling its role as the Metropolitan Planning Organization (MPO) for the Washington, DC region, the National Capital Region Transportation Planning Board (TPB) at the Metropolitan Washington Council of Governments (COG) has conducted a regional household travel survey approximately every ten years since 1968. The survey, which collects demographic and travel information from a randomly selected representative sample of households in the TPB region and adjacent areas, is the primary source of observed data used to estimate, calibrate, and validate the regional travel demand model. The purpose of the survey is to better understand the characteristics of the households and persons in the region and to better understand daily travel and activities: how we travel, why we travel, where we go, how long it takes us, and what we do when we arrive. The survey seeks to obtain a complete picture of travel patterns in the region. As a result, the regional household travel survey is a critical and essential element of the TPB work program. In this study two files were used in the data collection, namely Household and Person files.

Household File: includes characteristics of households, including, among others, household size, income, number of licensed drivers, housing type, and number of vehicles and bicycles. Person File: includes characteristics of individual persons, including, among others, demographic information, employment status, work location, and usual commute mode (TPB 2021).

Second Dataset:

The second dataset is extracted from Replica (Replica 2022). Replica Trends dashboard provides a nationwide estimate of mobility and economic activities, updated weekly at a census-tract level fidelity. Replica's mobility and spend data are available on weekly basis since the beginning of 2019 till 2023.The geographics available in Trends dashboard match the U.S. Census definitions for States, Combined Statistical Areas, Metropolitan Statistical Areas, Cities, Counties, and Census Tracts. Populations are sourced from 2019 U.S. Census ACS data. The jurisdiction used in this research is described as follows:
Geography: Washington-Arlington-Alexandria, DC-VA-MD-WV
Metric: Consumer spend by home location
Metric Option: In Person and Online Spend
Geography Breakdown: county
Date: Fall, 2022
Days of Week: Weekday

Data Processing: Both datasets described above contain missing values. Therefore, in the data processing step, the missing values are replaced with meaningful values without changing the structure of the data sets. Data Analysis is carried out using JupyterLab of Python.

Potential Influencing Factors and Exploratory Data Analysis (EDA) The potential influencing factors are identified based on the literature review, which is called a set of explanatory variables and used as initial inputs of the model. To examine the effect of each of the factors on people's desire to telework, univariate and bivariate analyses of the variables were conducted.

6.4. Model development

To develop the predictive models, the data set is split into training and test sets. 70% percent of the data set is considered for the training purpose and the other 30%, which was not used in the training process, is used for the evaluation process. The problem is considered as a supervised classification problem, meaning, the data set is labeled, the input variables are known, and the outcome, which is people's decision to telework, is a binary variable, meaning if an employee decides or have the option to telework the value of this variable is 1 and otherwise it is 0. The variables are rescaled and some of the well-known machine learning classification methods including Support Vector Machines, Decision Trees, K-Nearest Neighbors, and Random Forest, logistic Regression and multinomial Logit are employed. These algorithms are explained briefly in the following paragraphs.

Scaling the data:

The independent variables in this dataset have different scales. When features have different scales from each other, there is a chance that a higher weightage will be given to features that have a higher magnitude, and they will dominate over other features whose magnitude changes may be smaller but whose percentage changes may be just as significant or even larger. This will impact the performance of our machine learning algorithm, and we do not want our algorithm to be biased towards one feature.

The solution to this issue is Feature Scaling, i.e. scaling the dataset so as to give every transformed variable a comparable scale. Tree based models such as Decision Trees and Random Forest does not require feature scaling to be performed as they are not sensitive to the variance in the data. The data for Logistic Regression and SVM and K-nearest neighbor are scaled. The standard Scaler method is used, which centers and scales the dataset using the Z-Score. It standardizes features by subtracting the mean and scaling it to have unit variance. The standard score of sample x is calculated as:

$$Z = \frac{x - u}{S} \tag{1}$$

Where u is the mean of the training samples (zero) and S is the standard deviation of the training samples.

Support Vector Machine (SVM) is a supervised learning algorithm used for regression, classification, and outlier detection. The main objective of SVM is to create a line or a hyperplane which separates the data into classes.

K-Nearest Neighbors (KNN) is a type of supervised machine learning (ML) algorithm that can be used for both classification and regression predictive problems. However, it is primarily used in the industry for classification and prediction problems.

Decision Trees are tree-based models that help in making decisions in both regression and classification problems. To make a decision, they use a hierarchical structure and split the dataset into smaller subsets.

Random Forest is a commonly used machine learning algorithm that is trademarked, which combines the output of multiple decision trees to reach a single result. Its ease of use and flexibility have fueled its adoption, as it handles both classification and regression problems.

In addition to the classifiers, an ANN algorithm is used which is a parallel information-processing system that has certain performance characteristics similar to biological neural networks. A neural net consists of large numbers of simple processing elements called neurons. Each neuron is connected to other neurons by means of directed links, and each directed link has a weight associated with it. These are used to address problem that are intractable or cumbersome with traditional methods (Agatonovic-Kustrin 2000).

6.5. Validation Methods.

To evaluate the models, several accuracy measurements were examined which are defined as follow.

Accuracy: Accuracy is a metric for classification models that measures the number of predictions that are correct as a percentage of the total number of predictions that are made (Korstanje 2021).

$$\text{Accuracy} = \frac{\text{total \# of correct predictions total}}{\text{total \# of predictions}} \tag{4}$$

Accuracy is a useful metric only when we have an equal distribution of classes in our classification. This means that if we have a use case in which we observe more data points of one class than of another, the accuracy is not a useful metric anymore, and other measures of accuracy will be used such as F1 score, recall and precision which are defined as follows.

Confusion Matrix:

Confusion matrices represent counts from predicted and actual values. The output “TN” stands for True Negative which shows the number of negative examples classified accurately. Similarly, “TP” stands for True Positive which indicates the number of positive examples classified accurately. The term “FP” indicates False Positive values, i.e., the number of actual negative examples classified as positive; and “FN” indicates False Negative values, which is the number of actual positive examples classified as negative (Kulkarni, Chong and Bataresh 2020).

The confusion matrix was utilized for the performance evaluations of the methods used after the classification. For binary classification, the scheme of the confusion matrix is shown in Figure 3.

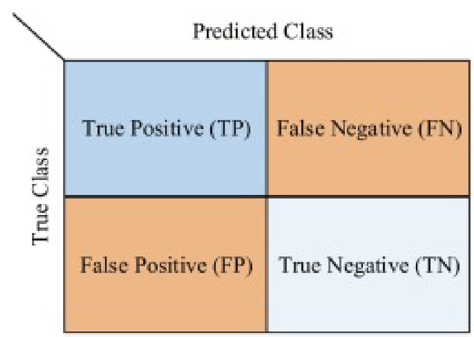


Figure 45. Confusion matrix for binary classification
(Source: Kulkarni, 2020)

ROC Curve:

An ROC curve (receiver operating characteristic curve) is a graph showing the performance of a classification model at all classification thresholds. This curve plots two parameters: True Positive Rate and False Positive Rate. True Positive Rate (TPR) is a synonym for recall and False Positive Rate (FPR) is defined as follows:

$$\text{False positive rate} = \frac{\text{\#of false positives}}{\text{\# of false positives + \# of true Negatives}} \quad (5)$$

An ROC curve plots TPR vs. FPR at different classification thresholds. Lowering the classification threshold classifies more items as positive, thus increasing both False Positives and True Positives. The following figure shows a typical ROC curve.

F1 score:

A simple way to solve class imbalance problems is to use better accuracy metrics such as the F1 score, which takes into account not only the number of prediction errors made by the model, but also the type of errors that are made (Korstanje 2021).

The F1 score is defined as the harmonic mean of precision and recall which is defined in equation (5).

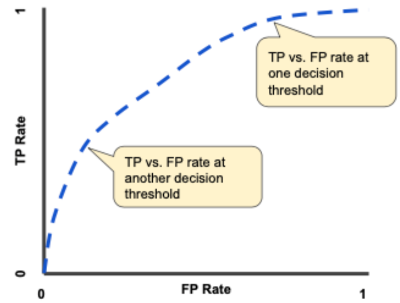


Figure 46. ROC Curve definition (Source: Kulkarni, 2020)

$$\text{F1 score } = 2 * \frac{\text{Precision} * \text{recallprecision}}{\text{precision} + \text{recall}} \quad (6)$$

where Precision is the first part of the F1 Score. It can also be defined as

$$\text{Precision } = 2 * \frac{\text{\# of true positives}}{\text{\#of true positives } + \text{ \# of false positives}} \quad (7)$$

Recall is the second component of the F1 Score, although recall can also be used as an individual machine learning metric. The formula for recall is:

$$\text{Recall } = 2 * \frac{\text{\# of true positives}}{\text{\#of true positives } + \text{ \# of false Negatives}} \quad (8)$$

6.6. Explanatory variables

To develop the predictive models, a list of influential factors is extracted from the literature review, based on availability of data for those factors from the two datasets used in this study *(Replica 2022, TPB 2021)* Table 1 and 2 are formed as input values for the two case studies that are performed in this paper.

Table 18. Input Variables for Case Study 1

Variable Name	**Variable Label**	**Variable Description**

X1:	HOME_OWNERSHIP	Household residence tenure status
X2:	NUMDRIVERS	Number of household members with a license
X3:	NUMWORKERS	Number of workers in household
X4:	NUMVEHICLE	Number of household vehicles
X5:	NUMBICYCLE	Number of household bicycles
X6:	HH_INCOME_DETAILED	Household income
X7:	AGE	Age
X8:	GENDER	Gender
X9:	RACEETHNICITY	Race and Ethnicity
X10:	EMPLOYMENT_STATUS	Employment status
X11:	J1_WORKPLACE_LOC	Employed and has usual work location
X12:	TELECOMMUTE_TIME	Telecommute time
X13	Combined_benefit	Sum of all transit benefit such as parking, carpool, biking and walking benefits

Table 19. Input Variables for Case Study 2

Variable Name	Variable Label	Variable Description
X1:	Age	Age, Numeric variable
X2:	sex	Sex, a categorial variable
X3:	race	Race, a categorial variable
X4:	ethnicity	Ethnicity, a categorial variable
X5:	employment	Employment status, a categorial variable
X6:	education	Education, a categorial variable
X7:	commute_mode	Commute mode choice, a categorial variable
X8:	tenure	Household residence tenure status
X9:	household_size	The size of the household
X10:	household_income_group	Income of the household a categorial variable

X11:	vehicles	Number of vehicles in the household
X12:	office_size	office_size: Very Small: < 10 people Small: 10-24 people Medium: 24-49 people Large: 50-99 people Very Large: >= 100 people
X13:	duration_minutes	Trip duration is calculated as the time it takes for a person to make a trip from their house to the work location as a straight line if they were not teleworking.
X14:	distance_miles	Trip distance is calculated as the distance between from a person's house to their office location, if they were not teleworking.
X15:	Trip_cost	Trip cost is defined as Trip distance times \$0.4331

Dependent Variable:
The dependent variable for both case studies is the telework, if an employee decides to telework it is shown by 1 in the first data set and by yes in our second data set, and if they decide not to telework, it is shown by 0 or No.

Table 3. Dependent variable for telecommuting model

y	TELECOMMUTE	Numeric-Categorical	No 0 Yes 1

6.7. Results and Discussions
In this section the results of the developed models are presented in 2 parts, part one includes the results for the first case study, the second part presents the results for the second case study.

Part 1:

The thirteen (13) influential factors in Table 1 and the dependent variable which represents whether a person teleworks or not are used to construct machine learning models. Optimal configuration for the machine learning models is obtained by trying different methods for scaling the variables, namely standardization and normalization, and by using nonlinear activation functions. Correlation between the variables was examined and the highly correlated variables are eliminated. The correlation matrix is provided following Figure 5.

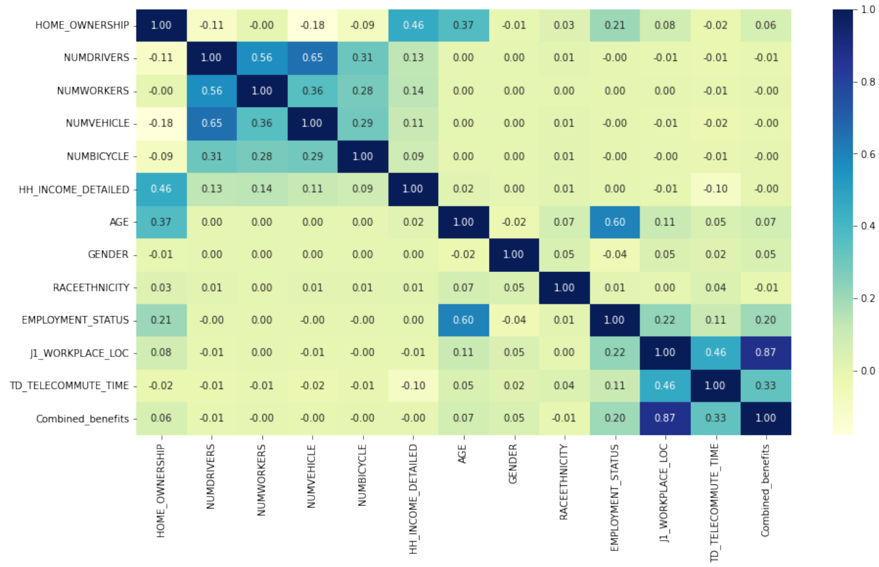


Figure 47. Correlation Matrix for input Variables- Case Study 1

Based on the data distribution shown in the Figure 6., 88% of people in Washington metrop olitan area did not telework in 2017 and only 12 percent did.

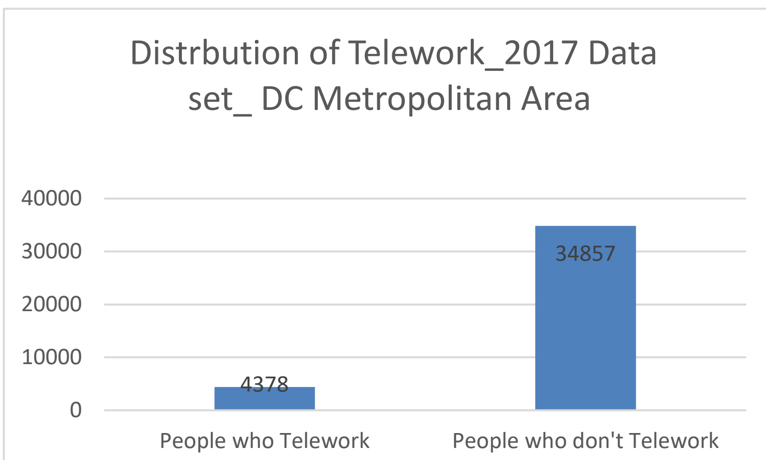


Figure 48. Distribution of people who telecommute- Case Study 1

6.7.1. Model Validation for Case Study 1.

The Machine Learning algorithms discussed in the methodology were implemented on the training set consisting of 27,465 households in the train set, and they have been validated using the proposed methods on the test set including 11,770 households. There are two type of errors these models can make, if a person does not telecommute and the model predicts they would (False positive), or when someone does telecommute, and the model predicts they do not (False negative). A high Recall will reduce the number of false negatives and a high Precision will reduce the number of false positives. Unfortunately, you can’t have both precision and recall high. If you increase precision, it will reduce recall, and vice versa. This is called the precision/recall tradeoff**.** For this matter, along with accuracy we also consider, f1 score which is the harmonic mean of Precision and Recall, for both training and test sets. Also, sometime, the predictive models perform very well on the test set and not so well on the training set, to show the complete report, the accuracy measurement for both training and test sets are provided in the tables below.

As for the logistic regression model, since the dataset is unbalanced, i.e. the percentage of people who telework 12% to those who do not 88% is very significant, several methods are used to increase the accuracy of model. The table below shows the accuracy of logistic regression with no weight applied to the two classes of the output variable and with the balanced weight applied. A threshold of the Precision and Recall is computed and used to improve the model performance, the results are provided in Table 3 and Figure 25.

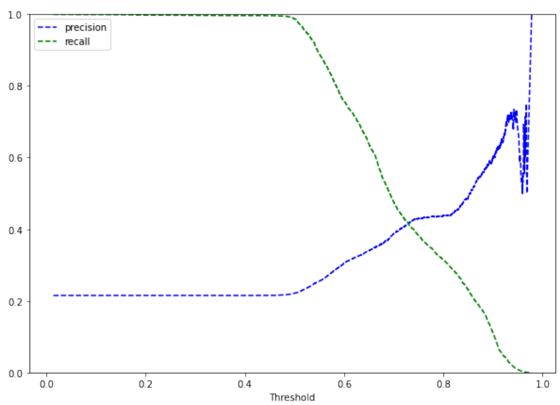


Figure 49. Precision and Recall Threshold for Logistic Regression_ Case Study1

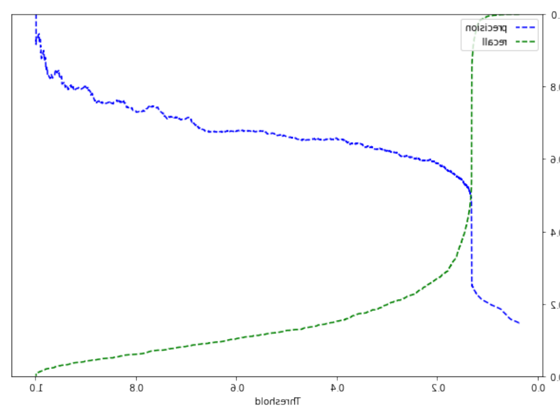


Figure 50. Precision and Recall Threshold for SVM_ Case Study 1

Table 3 shows, the best performance of the Logistic Regression on the 2017 dataset is when the Precision and Recall threshold is considered in the model.

Table 20. Comparison of Logistic Regression models

Model	Accuracy Score Training Set	Accuracy Score on Test Set	F1 Score for the Validation Set for the class of people who telecommute	F1 Score for the Validation Set for the class of people who do not telecommute
Logistic regression with no weight applied to the classes of output variable	89%	89%	22%	94%
Logistic regression with balanced weight applied to the classes of output variable	61%	61%	36%	72%

Logistic Regression with Precision and Recall threshold	87%	87%	40%	93%

The same process has been implemented for SVM model. The performance of SVM model is better than Logistics Regression, the results are shown in Table 4 and Figure 26.

Table 21. Primary results for SVM Model

Model	Accuracy Score Training Set	Accuracy Score on Test Set	F1 Score for the Validation Set for the class of people who telecommute	F1 Score for the Validation Set for the class of people who do not telecommute
SVM without Precision and Recall threshold	89%	89%	0%	94%
SVM with Precision and Recall threshold	89%	86%	48%	94%

Decision Tree, Random Forest and K-Nearest Neighbors Models perform better on training set than test set. Table 5. Represents the results for these models.

Table 22. Model Comparisons_ Case Study 1

Model	Accuracy Score Training Set	Accuracy Score on Test Set	F1 Score for the Training Set_People who Telecommute	F1 Score for the Training Set_People who don't Telecommute	F1 Score for the Validation Set_People who Telecommute	F1 Score for the Validation Set_People who don't Telecommute
K-Nearest Neighbors	91%	86%	44%	95%	26%	92%
Decision Tree	97%	86%	82%	98%	29%	92%
Random Forest	97%	88%	82%	98%	29%	93%

Feature Importance. Using the developed models, the level of importance of all factors including continuous and categorial input variables were quantified for these 39235 households' data point. The estimated numbers indicate the relative importance of each factor affecting a person's decision to telecommute. A larger value shows that the factor has a larger impact in people's decision to telecommute. The results for the first 9 influential factors are presented in Figure 9. Age is the most influential factor on people's decision to telecommute based on 2017 dataset, i.e., before Covid-19 pandemic.

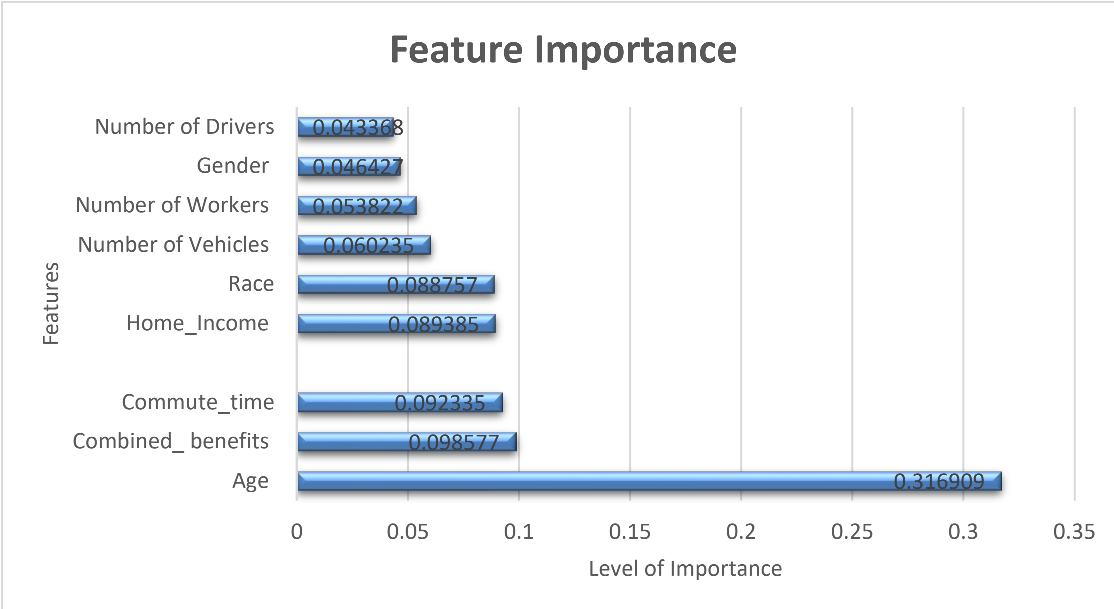


Figure 51. Feature Importance_ Case Study 1

The results from Tables 3,4 and 5 shows, the classification models can offer a very good prediction for the people who do not telecommute, that is because the data set is unbalanced and there are a few data points (18% of the entire dataset) which represent people who do telecommute, in order to increase the performance of these models more data must be collected to represent both group of people, this is covered in our second case study.

Part 2:

The thirteen (15) influential factors in Table 2 and the dependent variable which represents whether a person teleworks or not are used to construct machine learning models. The second case study is performed using the data set collected from Replica for the year 2021, after Covid-19 pandemic. This dataset includes information about 936460 households in Washington metropolitan area. Same processes that were explained for the case study 1 is applied to the data sets for case study 2, and the results are provided as follows.

The percentage of people who telework has drastically increased from 18% to 55% in the new dataset from 2021, as shown in the Figure 10.

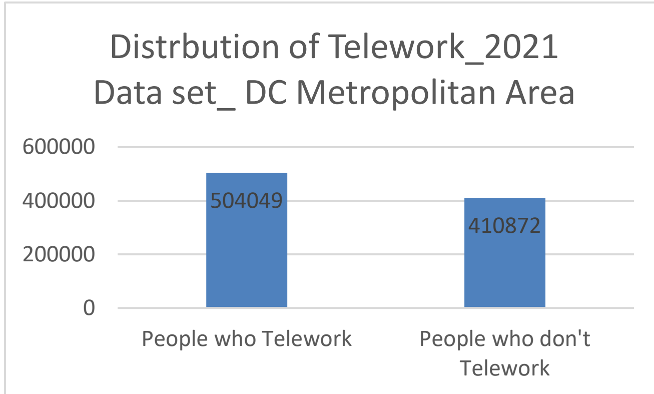


Figure 52. Distribution of people who telecommute- Case Study 2

6.7.2. Model Validation for Case Study 2.

The Machine Learning algorithms discussed in the methodology were implemented on the dataset containing 914921 households' information. After splitting data to training and test set for model development purposes, the training set consists of 274477 households, and the test set includes 274476 households. Since the data set for the second case study is balanced there is no need to add weight to the classes. The precision and Recall threshold was computed and optimal results for logistic regression is presented in the Table 6.

Table 23. Model Comparisons_ Case Study 2

Model	Accuracy Score Training Set	Accuracy Score on Test Set	F1 Score for the Training Set_People who Telecommute	F1 Score for the Training Set_People who don't Telecommute	F1 Score for the Validation Set_People who Telecommute	F1 Score for the Validation Set_People who don't Telecommute
K-Nearest Neighbors	81%	86%	83%	80%	75%	65%
Decision Tree	100%	86%	82%	98%	71%	65%
Random Forest	98%	75%	98%	98%	80%	70%
Logistic Regression	70%	70%	76%	65%	76%	62%

The results show, when the data set is balanced and we have enough data points to train the models, most of the classification models perform almost the same and can be used for building a predictive model and computing feature importance. The predictive model presented here is a statistical technique using machine learning and data mining to predict and forecast likely future outcomes with the aid of existing data. It works by analyzing current data and projecting what it learns on a model generated to forecast likely outcomes. A predictive model is not fixed; it is validated or revised regularly to incorporate changes in the underlying data (Rami 2020). The model developed in this study are predictive models, meaning they can forecast the future outcome of telecommuting estimation using new data sets.

6.8. Conclusions:

In this section, by using two existing data sets from the Transportation Planning Board and Replica, a methodology to develop several Machine Learning models was proposed to predict the people decision to telecommute and to study the comparison between these models. From univariate and bivariate data analyses, and the verification of the results of the proposed models, the following conclusions can be drawn:

- A comprehensive literature review was performed and influential factors impacting telecommuting were identified.
- A correlation matrix was established, and highly correlated factors were identified. A representative of the correlated variables was introduced to be used as a new input factor in developing the model.
- Data analysis was conducted and the growth of telework was highlighted.
- Several computational models were developed based on the most famous Machine Learning classification algorithms, using influential factors as input and telecommuting as output variables.
- The proposed models predict telecommuting with accuracy of up to 98% on the training sets and 86% on the validation sets.
- A comparative study among the developed models was carried out to identify the main factors influencing telework.
- From the computed importance factors, it was observed that, age, transportation benefit, income, commute time, education and office size have the most significant impact on the people’s decision to telework.
- The developed predictive models are adaptive and can be updated using new data sets with new set of input variables obtained from different geographical locations to estimate telecommuting behavior in various circumstances.

7. Methodology for the Survey Questionnaires

Task 4 of the project requires that we carry out a survey of workers’ mode choice to workplaces in Metropolitan DC. The survey will assist to determine workers' mode choice to workplaces, it will add credibility to the project and gauge people's attitude with respect to telecommuting. The data outcome of the survey will complement any other sources of data that will be used for the revenues and telecommuting models. Therefore, the surveys (questionnaires) contain both quantitative and qualitative questions. The quantitative questions take the form of yes/no, or rating scale (1 to 5), whilst the qualitative questions present a box where people can write in their own words. The questionnaires are comprehensible using clear language to ease the cognitive burden for the respondents. Two sets of questionnaires have been designed to be administered, as shown in Appendix I and II, to employers and workers, respectively. The questionnaires are being administered through Institution Research in Morgan State University (Campus Lab) and using information from the DC Chamber of Commerce. It is planned that a total of 1500 respondents comprising 1000 workers and 500 employers will be sampled in Metropolitan DC. It is considered that based on the data required for the models, time and budgetary constraints, the number of sample size is justified. Pretesting of the questionnaires has commenced using select targeted respondents. We are in the process of obtaining IRB approvals to enable us to administer the questionnaires. Whilst we are expecting additional data from replica, the type of data from replica will our further adjustments of the questionnaires.

8. Conclusions

8.1. Main Findings

In this section the main findings of this study are summarized.

- A comprehensive literature review was performed and influential factors impacting telecommuting were identified.
- A correlation matrix was established, and highly correlated factors were identified. A representative of the correlated variables was introduced to be used as a new input factor in developing the model.
- Data analysis was conducted and the growth of telework was highlighted.
- Several computational models were developed based on the most famous Machine Learning classification algorithms, using influential factors as input and telecommuting as output variables.

- The proposed models predict telecommuting with accuracy of up to 98% on the training sets and 86% on the validation sets.
- A comparative study among the developed models was carried out to identify the main factors influencing telework.
- From the computed importance factors, it was observed that, age, transportation benefit, income, commute time, education and office size have the most significant impact on the people's decision to telework.
- The developed predictive models are adaptive and can be updated using new data sets with new set of input variables obtained from different geographical locations to estimate telecommuting behavior in various circumstances.
- The results estimated with the revenue model show the sensitivity of various types of revenues and other relevant measures of effectiveness (such as numbers of trips and vehicle miles, mode choices, equivalent greenhouse gas (GHG) emissions and fatalities) to numerous influencing factors, singly and in some combinations. It should be emphasized that the relative (e.g., percentage) changes are more accurate and reliable than the absolute numerical values of results.
- The results provided in Chapter 4 indicate the relative effectiveness of various tax policies and rates in generating revenues for the DC government.
- The results in Chapter 4, especially in Figures 2, 3, and 4, indicate that the sensitivity of vehicle energy use to vehicle energy taxes is relatively low. The reason is that energy taxes account for a small fraction of the total trip impedance (which includes out-of-pocket costs and value of travel time) experienced by users. Even within the total cost of a gallon of fuel, DC taxes account for a relatively small share. This suggests that DC could increase those fuel taxes quite significantly to greatly increase revenues, provided that adjoining jurisdictions (in Maryland and Virginia) roughly match the DC tax rate increases.
- The figures and Table 22 in Chapter 4 indicate what changes in revenues and other relevant effects can be expected in future years, e.g., by considering the effects of changes in potential demand, electric vehicle penetration in the overall fleet mix, and technological changes such as vehicle energy efficiency improvements.

8.2. Recommendations

A major finding of this study is the relative insensitivity of travel demand to increases in transportation tax rates. This could have been expected qualitatively before the study began, but this study quantified the effect, which is striking. Currently, DC taxes account for a relatively small fraction of the total trip impedance perceived by **travelers** (which includes their out-of-pocket-cost and the value of their travel time). Thus, transportation tax rates can be increased significantly (e.g., even doubled) without greatly reducing the amount of travel.

From current tax rate levels, increases in tax rates yield nearly proportional increases in tax revenue, provided that jurisdictions near DC (i.e., in Maryland and Virginia) roughly match the tax rate increases in DC. If the neighboring jurisdictions do not match DC's increases in fuel tax rates, DC can still consider the option of imposing tolls, parking taxes, and/or vehicle mile taxes, which are not directly affected by the tax rate decisions of neighboring jurisdictions. It should be noted that this study has not modelled some possible long -term effects of increased DC taxes, such as greater shifts to remote activities or relocations of activities and residences away from DC.

In applying the results of this study, it should be remembered that it only estimated revenues from DC transportation taxes. Additional possible sources of revenues, such as from more general taxes (e.g., on incomes, sales and properties) and from the Federal government, should also be considered. Conversely, transportation taxes might in some circumstances be used to cover non-transportation expenses.

8.3. Possible extensions

Due to time and other constraints the methods and results presented here have some limitations which may be improved upon in future studies. The following extensions may be considered through further research:

1. The current model does not distinguish sufficiently among major socio-economic groups. Separating the causal relations and results by socio-economic groups would enable better consideration of equity issues.
2. The current model does not directly consider how decisions (e.g., regarding tax rates or other fees) in one period may affect the future fleet mix, location decisions, mode choices and other variables in future periods.
3. The current model does not sufficiently consider the effects of tax differences between the District of Columbia and adjacent jurisdictions. Although the motorists' sensitivity to vehicle energy taxes is relatively small, and seems to allow relatively large increases in such tax rates, the cross-elasticity to tax rates in Virginia and Maryland may be quite substantial. Thus, the model tends to underestimate the increase in DC revenues due to energy tax rates increases if those rate increases are not matched by neighboring jurisdictions.
4. The current model considers a limited number of transportation alternatives that are available to users. Other modes or mode combinations, including non-motorized modes, may be included in future studies.
5. The current model estimates the effects of various factors on transportation revenues to DC but does not consider the effects of those factors on the costs of providing the infrastructure, which clearly affect the need for revenues.
6. Further research seems worthwhile on technological and socio-economic trends that may affect the demand for transportation. These may include improvements in communication

technologies that facilitate remote activities as well as changes in production technologies (e.g., 3-dimensional printing) and distribution systems (e.g., drone deliveries) that may affect freight transportation.

7. It seems desirable to estimate the possible effects of connected and driverless vehicles on trip demand, parking use, mode shares, trip lengths and frequencies, revenues, costs, environmental impacts and others relevant outcomes.
8. It may be desirable to consider some long-term effects of changes in DC transportation tax rates and policies, such as greater shifts to remote activities or relocations of activities and residences away from DC.

Bibliography

Abduljabbar, Rusul and Dia, Hussein and Liyanage, Sohani and Bagloee, Saeed Asadi. 2019. "Applications of Artificial Intelligence in Transport: An Overview." *Sustainability* 11.

Adobati, F, and A Debernardi. 2022. "The breath of the metropolis: Smart working and new urban geographies. Sustainability, 14(2), 1028."

Agatonovic-Kustrin, S and Beresford, R. 2000. "Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research." *Journal of Pharmaceutical and Biomedical Analysis* 22 (5): 717--727.

ANL. 2021. "Argonne National Laboratory (2021), A look at Argonne's biggest breakthroughs in 202, US, Department of Energy."

Appel-Meulenbroek, R, A Kemperman, A van de Water, M Weijs-Perrée, and J and Verhaegh. 2022. " How to attract employees back to the office? A stated choice study on hybrid working preferences. Journal of Environmental Psychology, 81, 101784."

ARBA. 2020. "American Road and Transportation Builders Association, Bridge Report, 2020.U.S. ."

ASCE. 2021. "Report Card for American Infrastructure ."

Ashley, R. Haire, and B. Machemeh Randy. 2007. " Impact of Rising Fuel Prices on U.S. Transit Ridership Transportation Research Record: Journal of the Transportation Research Board National Academy of Sciences: Transportation Research Board Vol. 1992 Issue1."

Asmussen, K.E, A Mondal, C.R Bhat, and R.M Pendyala. 2022. "On modeling future."

Astroza, S, A Tirachini, R Hurtubia, J.A Carrasco, A Guevara, M Munizaga, M Figueroa, and V Torres. 2020. " Mobility changes, teleworking, and remote communication during the COVID-19 pandemic in Chile. Findings, 13489."

Atkinson, Robert D. 2009. " Paying our way : a new framework for transportation finance, final report, National Transportation Library."

Austernuhle, M. 2012. "Revenue from D.C. parking tickets dropped by half during pandemic, traffic cameras are humming along, NPR, WAMU 88.5."

Badger, E. 2011. "Badger, E. "Are Road use fees just too creepy to work?". The Atlantic Cities.Bloomberg.com. US Edition."

Baker, R, and Goodin G. 2011. " "Exploratory study: vehicle mile fees in Texas" Texas Transportation Institute, Technical report; Mar. 2010-July 2010."

Bjursell, C, I Bergmo-Prvulovic, and J Hedegaard. 2021. "Telework and lifelong learning. Front. Sociol. 6."

Blake, D Ratner, Meinzen-Dick Ruth, K. May Candace, and Haglund, Eric. 2010. " Resource Conflict, Collective Action, and Resilience: An Analytical Framework, International Journals of Commons (7)1."

Brewer, A.M, and D.A Hensher. 2000. "Distributed work and travel behaviour: the dynamics of interactive agency choices between employers and employees. Transportation 27, 117–148."

Caldarola, B, and S Sorrell. 2022. " Do teleworkers travel less? Evidence from the English National Travel Survey. Transportation Research Part A, 159, 282-303."

Caulfield, B. 2015. "“Does it pay to work from home? Examining the factors influencing working from home in the Greater Dublin Area”, Elsevier Volume 3, Issue2, June2015, 206-14."

Caves, R W. 2004. " Encyclopedia of the City. Routledge. p. 663. ISBN 978-0415862875."

CBO. 2019. "Congressional Budget Office, “Issues and options for a tax on vehicle miles traveled by commercial trucks,” 10-23."

CBO. 2020. "Congressional Budget Office, Highway Trust Fund Accounts—CBO’s Baseline as of March 6, 2020."

Choo, S, P.L Mokhtarian, and I Salomon. 2005. " Does telecommuting reduce vehicle-miles traveled? An aggregate time series analysis for the US. Transportation 32,37–64."

Choo, S, P.L Mokhtarian, and I Salomon. 2005. ". Does telecommuting reduce vehicle-miles traveled? An aggregate time series analysis for the US. Transportation 32,37–64."

Columbia, Council of District of. 2023. "“Code of the District of Columbia § 47–2501. Gas, electric lighting, telephone, telecommunications, and heating oil companies”."

Danalet, A, A Justen, and N.A Mathys. 2021. " Working from home in Switzerland, 2015-2050. In 21st Swiss Transport Research Conference, September."

Das, S., A. Dey, A. Pal, and N. Roy. 2015. "Applications of Artificial Intelligence in Machine Learning: Review and Prospect." *International Journal of Computer Applications* (IJCA) 115.

Davis, B. 2021. "‘’What is telecommuting explain in brief’’ Mvorganizing.org."

DDOT, DC. 2017. " Transit Development Plan Update, December 2017."

DDT. 2011. " Report on the condition, planning and revenues options for the support of the transportation trust fund" Transportation Trust Fund Task Force."

—. 2021. *District Department of Transportation; Parking meters.* https://ddot.dc.gov/page/parking- meters.

DeWitt, J. 2020. "“D.C. Tax Facts,” Government of the District of Columbia Office of the Chief Financial Officer Office of Revenue Analysis Part 2, p. 33."

DMV. 2021. "Department of Motor Vehicles Vehicle registration fees, https://dmv.dc.gov/node/155452."

Donnellan, Bob. 2023. "Gas tax by state 2022 - current State Diesel & Motor Fuel Tax Rates." IGEN.

EBP US, Inc. 2021. " The Impact of the COVID-19 Pandemic on Public Transit Funding Needs in the U.S., p. 3-5 Eco-Friendly Vehicles Draining State Road Repair Budgets- US News and World Report."

Eliza, B. 2020. " ‘’D.C. Has Some of The Longest Commutes In The Country. What Help Is Available’’, NPR WAMU 88.5."

Elldér, E. 2020. "Telework and daily travel: New evidence from Sweden. Journal of Transport Geography, 86, 102777."

Ellison. 2004. "Telework and Social Change: how technology is reshaping the boundaries between home and work, Westport Connecticut: Praeger, p. 18, ISBN."

Errichiello, L, and T Pianese. 2021. "The Role of Organizational Support in Effective Remote Work Implementation in the Post-COVID Era. Handbook of Research on Remote Work and Worker Well-Being in the Post-COVID-19 Era IGI Global."

EV, Adotion. 2020. " “EV Sales Forecasts,” accessed https://evadoption.com/ev-sales/ev-sales- forecasts/."

Feng, L, Robert G. K, Michail Z, and A Christopher. 2006. " Growth Forecasting of Vehicle Miles of Travel at County and Statewide Levels. Transportation Research Record pp. 56–65."

Forkenbrock, D, and J G Kuhl. 2002. " "A New Approach to Assessing Road User Charges". Public Policy Center, University of Iowa."

Gajendran, R. S., and D. A. Harrison. 2007. "The good, the bad, and the unknown about telecommuting: Meta-analysis of psychological mediators and individual consequences." *Applied Psychology* 92 (6): 1524–1541.

Gallup. 2020. " COVID-19: A leader ’s guide to developing a work-from-home strategy. Available at: https://www.gallup.com/workplace/310988/covid-19-working-from-home-guide.aspx. Accessed March 11, 2022."

Gareis, K, and N Kordey. 1999. " Telework-an Overview of Likely Impacts on Traffic and Settlement Patterns. NETCOM: Reseaux, communication et territoires/Netw.

Commun. Studies 13, 265–286."

Garret, Kelly, and R T. 2007. "Which Telework? Defining and Testing a Taxonomy of Technology-Mediated Work at a Distance. In Social Science Computer Review (Vol. 25, Issue 1, pp. 27–47). SAGE Publications. https://doi.org/10.1177/08944393062938."

Gordon, G. 1992. " ‘’Trip-reduction legislation signed into law in New Jersey, Telecommuting Review: The Gordon Report’’, 1 August, p. 11."

Gordon, G. 1996. "‘’Clinton calls for `family friendly’ work arrangements’’, Telecommuting Review, 13(8), p. 16."

Gordon, G. 1993. "‘’Georgia Legislature passes commuter efficiency resolution, Telecommuting Review’’ The Gordon Report, May, p. 3."

GSA. 2012. " "Mobile Worker Toolkit: A Notional Guide" GSA Enterprise Transformation."

Gurchiek, K. 2021. "Hybrid Work Model Likely to Be New Norm in 2021 [Online]. Society for Human Resource Management. Available: https://www.shrm.org/hr-today/ news/hr-news/pages/hybrid-work-model-likely-to-be-new-norm-in-2021.aspx [Accessed September 1, 20."

Heiden, M, L Widar, B Wiitavaara, and E Boman. 2021. " Telework in academia: associations with health and well-being among staff. Higher Education, 81(4), 707-722."

Hesse, B. W. 1995. "Curb cuts in the virtual community: telework and persons with disabilities, Proceedings of the 28th Meeting of the Hawaii International Conference on System Sciences, Vol. 4, pp. 418±425."

Hill, J E, B C Miller, S P Weiner, and J Colihan. 1998. " "Influences of the virtual office on aspects of work and work/life balance". Personnel Psychology. 51 (3): 667–683."

Hoffman, C.L. 2021. " The experience of teleworking with dogs and cats in the United States during COVID-19. Animals, 11(2), 268."

Igeltjørn, A, and L Habib. 2020. " Homebased telework as a tool for inclusion? A literature review of telework, disabilities and work-life balance. In: International Conference on Human-Computer Interaction,Springer, 420-436."

INRIX. 2021. " Global Traffic Scorecard Report www.inrix.com/scorecard, Inrix."

Irvin, D. 2019. " Gas Prices and Vehicles Miles Traveled Analyzed since 2000, Planetizen."

Irwin, F. 2004. " Gaining the air quality and climate benefit for telework. World Res. Institute. Retrieved from http://goo. gl/IvdkU."

Jaffe, Eric. 2013. "In Oregon, Pushing Ahead with Road Use Fees. The Atlantic Cities. Retrieved 7 January 2014."

Kelly, B. 2019. "Data: More work from home in US, Virginia, and DC area."

Kenneth, A, Kurt Small, and D Van. 2007. " "Fuel Efficiency and Motor Vehicle Travel: The Declining Rebound Effect," Energy Journal, vol. 28, no. 1, pp. 25–51."

Khan, O. 2007. *Modelling Passenger Mode Choice Behavior Using Computer Aided Stated Preference Data.* Queensland University of Technology.

Korstanje, J. 2021. *The F1 score* . https://towardsdatascience.com/the-f1-score-bec2bbc38aa6.

KPMG. 2022. " Current Trends in Remote Working. Available at:https://assets.kpmg/content/dam/kpmg/xx/pdf/2022/04/current-trends-in-remoteworking- 04-2022.pdf. Accessed on June 20, 2022."

Kulkarni, A., D. Chong, and F. Bataresh. 2020. "Foundations of data imbalance and solutions for a data democracy." In *Foundations of data imbalance and solutions for a data democracy*.

Lachapelle, U, G.A Tanguay, and L Neumark-Gaudet. 2018. "Telecommuting and sustainable travel: reduction of overall travel time, increases in non-motorised travel and congestion relief? Urban Studies, 55(10), 2226-2244."

Langer, A, V Maheshri, and C. Winston. 2017. " From gallons to miles: A disaggregate analysis of automobile travel and externality taxes. In Journal of Public Economics (Vol. 152, pp. 34–46). Elsevier BV. https://doi.org/10.1016/j.jpubeco.2017.05.003."

Lari, A. 2012. " 2012. Telework/Workforce flexibility to reduce congestion and environmental degradation? Procedia-social Behav. Sci. 48, 712–721."

Larson, W, and W Zhao. 2017. "Telework: Urban form, energy consumption, and greenhouse gas implications. Econ. Inq. 55, 714–735."

Lei, Z, and H Xiang. 2013. " "Feasibility and Advantages of Estimating Local Road Vehicle Miles Traveled on Basis of Global Positioning System Travel Data." Transportation Research Record, No. 2399, pp. 94–102."

Lewis, R. 2017. "Telecommuting extends the work week, at little extra pay. Iowa Now [Online]. Available: https://now.uiowa.edu/2017/01/telecommuting-extends- work-week-little-extra-pay [Accessed November 16, 2021]."

Liu, Y. &Cirillo C. 2016. "Evaluating policies to reduce greenhouse gas emissions from private transportation." *Transportation Research Part D* 219-233.

Liz, Malm. 2015. ""When Did Your State Enact Its Gas Tax?" Tax Foundation."

Margaret, W, and S Elena. 2014. " ''A Review of the Literature on Telecommuting and Its Implications for Vehicle Travel and Emissions'' Resources for the Future, Discussion Paper 04- 44."

Martin, L, L Hauret, and C. Fuhrer. 2022. " Digitally transformed home office impacts on job satisfaction, job stress and job productivity. COVID-19 findings. Plos one, 17(3), e0265131."

Matthews, H.S, and E Williams. 2005. " Telework adoption and energy use in building and transport sectors in the United States and Japan. J. Infrastruct. Syst. 11, 21–30."

McKinsey Global Institute. 2021. "McKinsey Global Institute, “The future of work after Covid-19.”."

MCNA. 2008. "Metro Capital Needs Assessment 2011-2020(2008). Riders Advisory Council."

Metro Rail Fleet Plan. 2010. " Customer Services, Operations and Safety Committee, Board Information Item III-D."

Microsoft. 2012. " "Ordinary Or Extraordinary?". My Career. Fairfax Media. Retrieved June 29, 2012."

Miglioretti, M, A Gragnano, S Margheritti, E Picco, and 2021. 2021. " Not all telework is valuable. J. Work Organ. Psychol. 37, 11–19."

Miguel-Hurtado, O., Guest, R., Stevenage, S. V., Neil, G. J., & Black, S. 2016. "Comparing Machine Learning Classifiers and Linear/Logistic Regression to Explore the Relationship between Hand Dimensions and Demographic Characteristics." *PLOS ONE* 11 (11).

Moeckel, R. 2017. " Working from home: Modeling the impact of telework on transportation and land use. Transportation Research Procedia, 26, 207-214."

Mohammadi, M, E Rahimi, A Davatgari, M Javadinasr, A Mohammadian, M.W Bhagat-Conway, D Salon, S Derrible, R.M Pendyala, and S. and Khoeini. 2022. " Examining the persistence of telecommuting after the COVID-19 pandemic. Transportation Letters,."

Mokhtarian, P.L, and I Salomon. 1997. " Modeling the desire to telecommute: The importance of attitudinal factors in behavioral models. Transp. Res. Part A: Policy Practice 31, 35–50."

Mokhtarian, P.L, S.L Handy, and I Salomon. 1995. " Methodological issues in the estimation of the travel, energy, and air quality impacts of telecommuting. Transp. Res. Part A: Policy Pract. 29, 283–302."

Mokhtarian, P.L, S.L Handy, and I Salomon. 1995. "Methodological issues in the estimation of the travel, energy, and air quality impacts of telecommuting. Transp. Res. Part A: Policy Pract. 29, 283–302."

Mozaffarian, D. et al. 2015. "Heart disease and stroke statistics--2015 update: a report from the American Heart Association." *Heart disease and stroke statistics* 131 (4): 29-322.

MWCG. 2017. " WMATA’s Funding Needs, The Magnitude and the Effect, Updated to Reflect WMATA’s Proposed FY 2018 Budget, Presentation to the DC Building Industry Association Microsoft, (2012). "Ordinary Or Extraordi."

MWCG. 2015. "Metropolitan Washington Council of Governmenets, State of the Region Infrastructure Report."

NASEM. 2020. "National Academy of Science, Engineering and Medicine “Telework Transportation Research in the Covid-19 Pandemic “Transport Research Board."

Nayak, S, and D Pandit. 2021. "Potential of telecommuting for different employees in the Indian context beyond COVID-19 lockdown. Transp. Policy 111, 98–110."

Ng, Andrew, and Michael Jordan. 2002. "On Discriminative vs. Generative Classifiers: A comparison of logistic regression and naive Bayes." *Adv. Neural Inf. Process. Sys* 2.

Nguyen, M.H. 2021. " Factors influencing home-based telework in Hanoi (Vietnam) during and after the COVID-19 era. Transportation, 48(6), 3207-3238."

Nilles, J. M. 1989. "‘’Traffic Reduction by Telecommuting: A Status Review and Selected Bibliography’’. Transportation Research A, Vol. 22A, No. 4, pp. 301-317. 13."

Nishadi, A. S. T. 2019. "Predicting Heart Diseases In Logistic Regression Of Machine Learning Algorithms By Python Jupyterlab." *International Journal of Advance Research and Publication* 3 (8): 69-74.

Noonan, M.C, and J.L Glass. 2012. "The hard truth about telecommuting. Monthly Lab. Rev. 135, 38."

NRC. 2002. " ‘’Effectiveness and Impact of Corporate Average Fuel Economy (CAFE) Standards (Washington, D.C.: National Academy Press,."

ODT. 2014. "Oregon Department of Transportation "Oregon's Road Usage Charge Program"."

ODT. 2014. "Oregon Department of Transportation “Legislative Report, Road Usage Charge Pilot Program Preliminary Findings" Oregon Department of Transportation."

ODT. 2014. "Oregon Department of Transportation, “Legislative Report, Road Usage Charge Pilot Program Preliminary Findings" ."

Ollo-Lopez,́ A, S Goni-Legaz, and A Erro-Garć es. 2020. " Home-based telework: usefulness and facilitators. Int. J. Manpower 42 (4), 644–660."

Park, H. A. 2013. "An introduction to logistic regression: from basic concepts to interpretation with particular attention to nursing domain." *Journal of Korean Academy of Nursing* 43 (2): 154-164.

Parker, K, and J., Minkin, R Horowitz. 2020. " How the coronavirus outbreak has–and hasn’t–changed the way Americans work. Pew Res Center. Tavares, A.I., 2017. Telework and health effects review. Int. J. Healthcare 3, 30."

Parker, K, J.M Horowitz, and R Minkin. 2022. ". COVID-19 Pandemic Continues To Reshape Work in America. Pew Research Center. Available at: https://www.pewresearch.org/socialtrends/2022/02/16/covid-19-pandemic-continues-to-reshape-work-in-america/. Acce."

Patricia. 2012. " About one in five workers worldwide telecommute: poll, Reuters."

Pluut, H, and J Wonders. 2020. " Not able to lead a healthy life when you need it the most: Dual role of lifestyle behaviors in the association of blurred work-life boundaries with well-being. Front. Psychol. 11, 3600."

PSRC. 2014. "Puget Sound Regional Council, A Global Position Based Pricing Pilot Project: Evaluating Traveler Response to Variable Road Tolling Through a Sample of Volunteer Participants “Traffic Choices Study, Summary Report."

Puget Sound Regional Council. 2008. " A Global Position Based Pricing Pilot Project: Evaluating Traveler Response to Variable Road Tolling Through a Sample of Volunteer Participants “Traffic Choices Study, Summary Report."

Rami, Ali. 2020. *Predictive Modeling: Types, Benefits, and Algorithms.* 09 23. https://www.netsuite.com/portal/resource/articles/financial-management/predictive-modeling.shtml.

Replica. 2022. https://replicahq.com/about/.

Robert, K. K, and D. F Jon. 2007. " “Review of Methods for Estimating Vehicle Miles Traveled.” Transportation Research Record, No. 1551, pp. 59–66."

Salomon, I. 1986. "‘’Telecommunications and Travel Relationships: A Review. Transportation Research A, Vol. 20A, No. 3, pp. 223-238. 12."

Salon, D, M.W Conway, D.C Da Silva, R.S Chauhan, S Derrible, A.K Mohammadian, S Khoeini, N Parker, L Mirtich, and A Shamshiripour. 2021. "The potential stickiness of pandemic-induced behavior changes in the United States. Proc. National Acad."

San, Diego. 2020. "’’Work from anywhere' is here to stay. How will it change our workplaces?"."

Sathya, R., and A. Abraham. 2013. "Comparison of Supervised and Unsupervised Learning Algorithms for Pattern Classification." *Advanced Research in Artificial Intelligence(IJARAI)* 2 (2).

Shabanpour, R, N Golshani, M Tayarani, J Auld, and A.K Mohammadian. 2018. "Analysis of telecommuting behavior and impacts on travel demand and the environment. Transp. Res. Part D: Transp. Environ. 62, 563–576."

Singh, P, R Paleti, S Jenkins, and C.R Bhat. 2013. " On modeling telecommuting behavior: Option, choice, and frequency. Transportation, 40(2), 373-396."

Song, Y, and J Gao. 2020. " Does telework stress employees out? A study on working at home and subjective well-being for wage/salary workers. J. Happiness Stud. 21, 2649–2668."

Steven, E. P, C Xuehao, and T Lavenia. 2004. "Forecasts of Future Vehicle Miles of Travel in the United States ." *Transportation Research Record* 147–155.

Strecht, P., L. Cruz, C. Soares, J. Mendes-Moreira, and R. Abreu. 2015. "A Comparative Study of Classification and Regression Algorithms for Modelling Students' Academic Performance." *International Conference on Educational Data Mining (EDM).* Madrid, Spain.

Szekeres, D, N Koppula, and J Frazier. 2007. "‘’VMT Forecasting Alternatives for Air Quality Analysis’’ Transportation Land Use, Planning and Air Quality. PP110-117."

Thomas, A, B. C. J. B Williams, and G. J.D Brianne. 2016. " ‘’Methodologies used to estimate and Forecast Vehicles Miles Travelled’’, Texas A & M Transportation Institute."

TPB, Transportation Planning Board. 2021. *COVID-19 Travel Monitoring Snapshot Pre-Pandemic.* National Capital Region.

TPCBB. 2020. "Tax Policy Center Briefing Book, ‘’Key Elements of the US Tax Systems’’ Tax Policy Centre."

Transportation Planning, Board. 2021. "‘’COVID-19 Travel Monitoring Snapshot Pre-Pandemic’’, National Capital Region."

TRIP. 2020. " “Key Facts About the U.S. Surface Transportation System,” A National Transportation Research Non-Profit."

TRIP. 2019. " “Restoring the Interstate Highway System,” Texas A&M Transportation Institute, 2019 Urban Mobility Report."

TRIP. 2018. "“Bumpy Road Ahead: America’s Roughest Rides and Strategies to Make Our Roads Smoother,”."

USBTS. 2019. "“Commute Mode” State Transportation Statistics."

USDOE. 1994. " ‘’Energy, emissions, and social consequences of telecommuting’’ Office of Policy, Planning, and Program Evaluation."

USDOT. 1993. " “Transportation Implications of Telecommuting” US DOT, Washington, DC."

USDOT. 2020. "U.S. Department of Transportation, National Bridge Inventory Management and Preservation."

USDOT. 2021. "US Vehicles-Miles” Bureau of Transportation Statistics."

USDOT, FHA. 2020. " “Status of the Nation’s Highways, Bridges, and Transit Conditions and Performance Report,”."

USDOT, FHA. 2016. " MAP-21 Comprehensive Truck Size and Weight Limits Study, Freight Management and Operations, 21st Century Operations Using 21st Century Technologies."

USEIA. 2020. "U.S. Energy information Administration, (2020). “Annual Energy Outlook 2020 with projections to 2050”."

ValuePengium. 2020. " Travel Up 43% From 2020, and Americans Are More Likely to Fly, Lendingtree."

VDM. 2021. "Vehicle registration historical data, https://dmv.dc.gov/page/vehicle-registration-historical-data."

Walls, M, E Safirova, and Y Jiang. 2007. "What drives telecommuting? Relative impact of worker demographics, employer characteristics, and job types. Transp. Res.Rec. 2010, 111–120."

Washington, DC. 2021. "Washington, D.C.'S Infrastructure Report Card. Report Card for America's Infrastructure."

WMATA. 2017. " Metrobus Fleet Management Plan, Office of Bus Planning, Final Report Version 2.1."

WMATA. 2021. "Washington Metropolitan Area Transit Authority, (2021), "Covid-19 public information: WMATA",."

Woody, L. 1995. " ''The Underground Guide to Telecommuting" Addison-Wesley. ISBN 978-0- 201-48343-7."

Yen, J.-R, and H.S Mahmassani. 1997. " Telecommuting adoption: Conceptual framework and model estimation. Transp. Res. Rec. 1606, 95–102."

Yen, J.-R, H.S Mahmassani, and R Herman. 1994. " Employer attitudes and stated preferences toward telecommuting: An exploratory analysis. Transp. Res. Rec. 1463, 15–25."

Zhang, S, R Moeckel, A.T Moreno, B Shuai, and J Gao. 2020. " A work-life conflict."

Zhu, P, L Wang, Y Jiang, and J Zhou. 2018. " Metropolitan size and the impacts of."

APPENDIX

EMPLOYERS' QUESTIONNAIRES (APPENDIX I)

SECTION 1 EMPLOYERS' INFORMATION

Telecommuting as a mode choice has been in existence over a long period. The outbreak of the Covid-19 pandemic has increased the adoption of telecommuting with its associated effects. You are hereby invited to participate in a survey focused on gathering information needed to conduct parametric study on the effects of telecommuting. There are no risks in participating in this study. In addition, the responses you provide will be kept confidential and used only for statistical analyses. You may withdraw from this survey at any time if you choose to. However, your contribution will help us to develop models for forecasting transportation revenues which will in turn help the DC Department of Transportation in making decisions. For any questions, please do not hesitate to contact me at desmond.amiegbebhor@morgan.edu or mehdi.shokouhian@morgan.edu

SECTION 1: Information about trips in DC

In this section of this survey, we would like to learn more about your organization. This information will only be used to classify your responses and will never be presented in an individually identifiable form. Please, select the applicable response by indicating with the (✓) symbol in the appropriate box.

1. For each commute you make to your destination within the core central business District of Columbia, please select amongst the following the mode of transport you currently use most:

☐ Teleworking
☐ Private car
☐ Carpool/Vanpool
☐ Rail
☐ Bus
☐ non-motorized mode (Cycle/Scooter/Walk)
☐ Multiple modes
☐ Other_______________________________________

2. If multiple mode, state reason(s)____________________

3. (A) If the **telecommuting** is currently your most used mode of transport, about how much are you saving on transportation? _______$/week

(i) For those who **drive alone – ONLY,** please estimate how much you pay for parking: $_______/day **or** $_______/week **or** $_______/month

ii. Approximately how much do you pay for fuel per week for your ***one-way*** commute to downtown ***only***? $_______/week

(B) For those who **carpool/vanpool – ONLY.**

i. How many other persons do you carpool/vanpool with on average? _______Persons.

ii. Approximately how much do you pay individually for carpooling? $______/week

(C) How much time do you spend inside the vehicle (driving alone or carpooling) while commuting from your home to work location on a ***one-way trip***?

________ minute(s) or ______ hour(s)

(D).What is the approximate ***one-way*** driving distance between your home and your work destination? ______miles.

(E) Approximately how many minutes does it take you to walk to your destination from the moment you get off your vehicle at the parking lot ? ______minute(s)

(F) Do you have a monthly public transit pass?

☐ Yes ☐ No

4. If either **rail or bus** (includes users of Park & Ride facilities) is currently your most used mode of transport for commuting to your work location, please answer the following:

(A). How long do you wait at the transit stop/station for your bus or train?

Approximately ______minutes per one-way trip

(B). Approximately how long does it take you to get to your workplace from the moment you board your public transit vehicle on a ***one-way trip***
______minutes or _____hour(s)
(C). Do you make any transfers during your daily commute from your origin (home) to your work location on a ***one-way trip***?

☐ Yes. How many? _____ Transfer(s). ☐ No

(D). Approximately how many minutes does it take you to walk to your destination from the moment you get off your public transit vehicle? ______minute(s)
(E). How much do you pay for transit fare for each **one-way** trip to workplace

☐ _______$/trip ☐ Own transit pass ☐ Other _________ ☐ $7.75 (day pass)

(F) Do you currently own/lease a vehicle?

☐ Yes ☐ No

5. If a **non-motorized** mode (e.g. bicycle/walking/scooting) is currently your most used mode of travel to your workplace, approximately how long does it take you to travel from your home to work? ___________ minutes **or** _________ hour(s)
6. On how many days do you go to your office in a typical week?

☐ 5 days/week ☐ 2 days/week
☐ 4 days/week ☐ 1 day/week
☐ 3 days/week ☐ Other_______

8 Does your organization offer any parking benefits such as free parking, vouchers for parking? ☐ None ☐ Free parking ☐ Parking vouchers ☐ others_______________

9 Please indicate on a scale of 1(low) to 5 (high) the relative importance of service attributes 1 – 6 when you choose a travel mode. Six modes are listed below:

1Rail: **Reliability** 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (2) **Cost** 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

(3) **Distance** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (4) **Convenience** 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5☐ (5) **Safety**: 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (6) **Covid Precaution** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

(7) **Travel Time** 1 ☐ 2☐ 3 ☐ 4☐ 5 ☐

2Bus: **Reliability:** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (2) **Cost** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (3) **Distance** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

(4) **Convenience** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (5) **Safety** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (6) **Covid Precaution** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (7) **Travel Time** 1 ☐ 2☐ 3 ☐ 4☐ 5 ☐

3 Private car: **Reliability:** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (2) **Cost** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (3) **Distance** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

(4) **Convenience** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (5) **Safety** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (6) **Covid Precaution** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (7) **Travel Time** 1 ☐ 2☐ 3 ☐ 4☐ 5 ☐

4 NMT: **Reliability:** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (2) **Cost** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (3) **Distance** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

(4**) Convenience** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (5) **Safety** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (6) **Covid Precaution** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (7) **Travel Time** 1 ☐ 2☐ 3 ☐ 4☐ 5 ☐

5 Telework: **Reliability:** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (2) **Cost** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (3) **Distance** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

(4**) Convenience** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (5) **Safety** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (6) **Covid Precaution** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (7) **Travel Time** 1 ☐ 2☐ 3 ☐ 4☐ 5 ☐

6) Network Transit Companies: **Reliability:** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (2) **Cost** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (3) **Distance** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (4**) Convenience** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (5) **Safety** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ (6) **Covid Precaution** 1☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

(7) **Travel Time** 1 ☐ 2☐ 3 ☐ 4☐ 5 ☐

10) How many days in a week do you commute? Please specify…………………

Definition of Terms:

Reliability: Quality of performance trustworthiness and consistency

Telecommuting: Working from a remote location outside the traditional office

Cost: An amount to be paid or spent to move from origin to destination

Distance: The length of a trip

Flexibility: Ability to move without restriction

Convenience: Use with little or no difficulty

Comfort: Ease and freedom from tension

Safety: Protection from danger, risk etc.

Covid Precaution: Self-precautionary measures to keep safe from covid

Rail: Rail is a means of public transport service operated on rail track. It includes Light or heavy rail

Bus: Means of public transport based on regular operation of transit buses along a route with scheduled timetable

Private car: A passenger car assigned for private use

Non-Motorized Transport: Non-Motorized Transport includes all means of transportation that do not used motorized propeller. It includes walking, scooter, bicycling etc.

Network Transit Companies: These include ridesharing, ride-hailing, dial-a-ride etc.

SECTION 2

In this section of this survey, we would like to know more about your response to increase in the price of gasoline. This information will only be used to classify your responses and will never be presented in an individually identifiable form. Please, select the applicable response by indicating with the (✓) symbol in the appropriate box.

1 How many one-way trips do you make in a week? _______________

2 What is your mode of trip?

Car ☐ Rail ☐ Bus☐ NMT☐ Uber or ride hailing ☐ Walk☐ Other______________

3 If you choose drive, do you drive in an electric ☐or gasoline car☐?

4 How many miles do you cover per one-way trip? _________________

5 How farther will you travel if gasoline price is decreased by 30%? ____________Miles per one-way trip

6 How many trips per week will you make with the above decrease? _____________

7 Where do you buy gas? DC ☐ Maryland ☐ Virginia ☐

SECTION 3

In this section, we would like to know about your demography and response to telecommuting. Please, select the applicable response by indicating with (✓) symbol in the appropriate box.

1. What is the highest level of Education you have attained?

(a)☐ GED (b) ☐ High school Diploma © ☐ Some college (d) ☐ Associates degree (e) ☐ bachelor's degree

(f) ☐ Graduate Degree (g) ☐ Post graduate degree (h) ☐ others

1)What is your gender?

(a) ☐ Male (b) ☐ Female

2)In 2021, what was your household's total annual income (from all sources) before taxes or other deductions from pay? (Anonymous and will be grouped with answers from all other participating households.)
a) Less than 10,000 (b) ☐ $10,000-14,999 (c) ☐ 15000-24999 (d) ☐ 25000-34999 (e) ☐ 35000-49999

☐50000-74999 (g) ☐75000-99999 (h)☐ 100000-149999 (i) ☐150000-199999 (j) ☐200000 or more

3)How old are you?

a) ☐ 5-11 years (b) ☐11- 17 years (c) ☐18-22 years (d) ☐23-30 years e) ☐30-40 years (f) ☐40-55 years (g) ☐56-65 years (h) ☐66-75 years (i) ☐76 years or older

4)Race?

(a) ☐Hispanic or Latino (b) ☐African American (c) ☐ Asian (d) ☐White (e) ☐Middle eastern (f)☐Other/ two or more races (g) ☐Prefer not to answer

5)What is your employment status?

a) ☐Worker, including self-employed (b) ☐Retired (c) ☐Volunteer (d) ☐Homemaker
e) ☐ Unemployed but looking for a work (f) ☐Unemployed, not seeking employment
g) ☐ Student (i) ☐ Disabled non-worker (j) ☐ other_____________

6)What is your employment type?

a) ☐Work for private for-profit firm compony
b) ☐Work for nonprofit firm/organization
c) ☐Work for federal government
d) ☐Work for state or local government, agency or organization
e) ☐Self-employed
f) ☐Other_________________

7)Work location?

a) ☐Usually same location (outside home)
b) ☐Workplace regularly varies (different offices or jobsites)
c) ☐At home (telecommute or self-employed with home office)
d) ☐Drive for living (driver, salesperson)
(f) ☐Do you rent or own your current residence?
9) Home
a) ☐Own/ buying (paying mortgage)
b) ☐Rent
c) ☐Provided by job
d) ☐Other______________

8Does your current employer offer Telecommuting?

a) ☐Yes b) ☐ No

9If not, and you were given the opportunity, would you telecommute?

a) ☐Yes b) ☐No

11How many days per week do you telecommute?

a) ☐0 days b) ☐1 day c) ☐2 days d) ☐3 days e) ☐4 days f) ☐ 5+ days

12How many hours per one-way commuting trip do you spend? __________

13How many motor vehicles (in working order) are there in your household?

0☐ 1☐ 2☐ 3 ☐other__________

14How many bicycles are there in your household primarily used by household members 16 years of age or OLDER?

☐ 0 ☐ 1 ☐ 2 ☐ 3 other__________

15Does your employer offer transit benefits?

☐ No ☐ Yes_________________

16Does your employer offer carpool benefits?

☐No ☐ Yes ________________

17Does your employer offer walking-biking benefits?

☐No ☐ Yes_______________