USDOT Region V Regional University Transportation Center Final Report

NEXTRANS Project No. 035OY02

IMPACT OF PUBLIC TRANSIT MARKET SHARE AND OTHER PASSENGER TRAVEL VARIABLES ON CO_2 EMISSIONS: AMASSING A DATASET AND ESTIMATING A PRELIMINARY STATISTICAL MODEL

By

Rabi G. Mishalani, Principal Investigator
Associate Professor of Civil and Environmental Engineering
and Geodetic Science
The Ohio State University
mishalani@osu.edu

and

Prem Goel, Co-Principal Investigator
Professor of Statistics
The Ohio State University
goel.1@osu.edu

DISCLAIMER

Partial funding for this research was provided by the NEXTRANS Center, Purdue University under Grant No. DTRT07-G-005 of the U.S. Department of Transportation, Research and Innovative Technology Administration (RITA), University Transportation Centers Program. Additional funding was provided by The Ohio State University (OSU) including the College of Engineering and the Department of Statistics. The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

USDOT Region V Regional University Transportation Center Final Report

TECHNICAL SUMMARY

NEXTRANS Project No. 035OY02 | Final Report, December 31, 2011

Impact of Public Transit Market Share and other Passenger Travel Variables on CO_2 Emissions: Amassing a Dataset and Estimating a Preliminary Statistical Model

Introduction

Policies that encourage the use of more efficient transportation modes are considered beneficial in terms of reducing carbon dioxide (CO_2) emissions. In support of developing such policies, the impacts of various transportation demand, supply, and regulation variables on passenger travel related CO_2 – the predominant greenhouse gas (GHG) – emissions are investigated. A methodology for integrating data from multiple sources in a consistent manner is conceived and implemented, producing a dataset consisting of 146 of the largest urbanized areas in the US. A preliminary model for CO_2 emissions per capita in terms of various explanatory variables in this dataset is developed, and future improvements are suggested.

Findings

While the effect of transit share is found to be statistically significant in the preliminary model, other variables exhibit a larger impact on CO_2 emissions, which is understandable in light of the fairly low and narrowly varied values of transit service utilization across most urbanized areas in the US. Additional variables whose coefficients are significant are lane-miles/capita, average travel time, vehicle occupancy, and the reciprocal of population density.

Recommendations

There are some steps that can be taken next to enhance the reliability of both the dataset and model of interest. The data for certain variables need to be further verified and cross-checked to make sure that conclusions made are valid. In addition, incorporating a few more variables in the dataset that could improve the explanatory power of the model is desirable. Further improvement of the model could come from including additional and already available explanatory variables. In addition, there are various interactions that could be investigated. One aspect that has not yet been considered is the influence of government policies and regulations pertaining to CO2 emissions on CO2/capita. It could also be the case that the level of CO2/capita in an urbanized area influences the public policies and regulations that are put in place. Therefore, the inclusion of such explanatory variables would result in simultaneity that would have to be addressed in specifying and estimating the model of interest. All of the above reflect further dataset development and modeling considerations that will be considered following this reporting period.

Contacts

For more information:

Rabi G. Mishalani
Principal Investigator
Dept. of Civil & Environ. Engineering and Geodetic Science
The Ohio State University
mishalani.1@osu.edu

Prem Goel
Co-Principal Investigator
Department of Statistics
The Ohio State University
goel.1@osu.edu

NEXTRANS Center
Purdue University - Discovery Park
2700 Kent B-100
West Lafayette, IN 47906

nextrans@purdue.edu
(765) 496-9729
(765) 807-3123 Fax

www.purdue.edu/dp/nextrans

NEXTRANS Project No. 035OY02

IMPACT OF PUBLIC TRANSIT MARKET SHARE AND OTHER PASSENGER TRAVEL VARIABLES ON CO_2 EMISSIONS: AMASSING A DATASET AND ESTIMATING A PRELIMINARY STATISTICAL MODEL

By

Rabi G. Mishalani, Principal Investigator
Associate Professor of Civil and Environmental Engineering
and Geodetic Science
The Ohio State University
mishalani@osu.edu

and

Prem Goel, Co-Principal Investigator
Professor of Statistics
The Ohio State University
goel.1@osu.edu

Report Submission Date: December 31, 2011

ACKNOWLEDGMENTS AND DISCLAIMER

Partial funding for this research was provided by the NEXTRANS Center, Purdue University under Grant No. DTRT07-G-005 of the U.S. Department of Transportation, Research and Innovative Technology Administration (RITA), University Transportation Centers Program. Additional funding was provided by The Ohio State University (OSU) including the College of Engineering and the Department of Statistics.

The authors thank Fred Salvucci of Massachusetts Institute of Technology for the insightful discussions. The authors also thank Ying Ding for assistance with urbanized area boundary representations using a GIS software and comparisons of these boundaries across data sources, and Promod Chandhok of USDOT/RITA/BTS for his assistance in locating shape files of FAUAs and NTD service areas. The efforts of Graduate Research Assistants Andrew Landgraf, Ashley Westra, and Dunke Zhou are greatly appreciated.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION AND MOTIVATION

Policies that encourage increased use of more efficient transportation modes, such as public transportation and high occupancy private autos, are often considered one of several possible tools in the "sustainable development" toolbox. However, no definitive quantifications are yet available regarding the potential benefits that could be derived from such policies in terms of potential reductions in carbon dioxide (CO_2) – the predominant greenhouse gas (GHG) – emissions.

The infrastructure supporting urban passenger transportation encompasses complementary and competing modes of travel, including private vehicle, urban street bus transit, bus rapid transit, light rail, and heavy rail (in addition to walking and biking, which are viable on a large scale in some urban areas but are not accounted for in this study). The different modes involve multiple characteristics in terms of cost, service, energy consumption, and environmental impacts. Some "discretionary" travelers can choose among two or more of the modes when making many of their trips, whereas other "captive" travelers do not have options. Moreover, the urban form and the corresponding origin-destination flow patterns have a direct bearing on the modes offered in terms of the spatial and temporal nature of the various services and, consequently, on the choices made by travelers. Given the varying supply and demand characteristics of the multiple modes across urban areas, passenger transportation related energy consumption and CO_2 emissions per traveler in an urban area are expected to be highly correlated with these characteristics. Given the efficient nature of public transportation and the greater flexibility in relying on multiple sources of energy, it is expected that, in general, an increased use of public transportation has potential advantages in reducing CO_2 emissions. Similarly, a high private vehicle occupancy is expected to mitigate the negative impacts of the single occupancy vehicle mode.

Transportation mode choices are made at the individual level, while transportation and land-use policies are made at the government level. Clearly, policies have the potential to influence choices, and at the same time the actual choices made under new policies directly determine the impact of such policies. Therefore, in support of evidence based policy making, it is important to establish a good understanding of the impact various passenger travel related variables may have on CO_2 emissions and the magnitudes of their impacts in urban areas, if any.

Several studies have investigated the effects of urban form on auto and energy use (Newman & Kenworthy 1989, Lomax et al. 1994, Holtzclaw et al. 2002, Bento et al. 2005, Ewing et al. 2007, Glaesar & Kahn 2008, Hankey & Marshall 2010, Parshall et al. 2010). For example, Hankey and Marshal (2010) investigated the impact of urban form on GHG emissions. In addition, some studies have investigated the relationship between GHG emissions and transportation (Karathodorou et al. 2009, Kockelman et al. 2009, Cambridge Systematics, Inc. 2009). Kockelman et al. (2009) examined opportunities for reducing such emissions and the "Moving Cooler" report (Cambridge Systematics, Inc. 2009) evaluated certain strategies aimed at such reductions. These studies found clear relationships linking urban form and transportation related variables to GHG emissions. However, these studies did not directly model and quantify the effect of high-occupancy vehicle (HOV) use, either transit or high-occupancy private automobiles, on GHG emissions.

Nevertheless, certain studies and workshops have recognized the potential importance of such variables and quantified their impacts in specific cases. The Urban Land Institute pointed to transit as an important element of "well-planned communities," as urbanized areas aim to mitigate climate change (ULI 2008). The Urban Public Transportation Roundtable at MIT (2009) dedicated a session to discuss the role that public transportation could play in addressing the GHG emissions problem. A fairly recent symposium at The City College of New York (University Transportation Research Center 2010) discussed some of these developments and highlighted the critical role of transit in achieving sustainable transportation. Additionally, the Brookings Institution recently commissioned a report (Brown et al. 2008) that in part recommended transit-oriented development to reduce the metropolitan carbon footprint. Schipper et al. (2010) discussed the reduction in CO_2 emissions resulting from the introduction of Bus Rapid Transit service (Metrobús) in Mexico City. Therefore, an explicit modeling and general quantification of the impact of transit and the extent of HOV use on GHG emissions, along with transportation supply, demand, and policy and regulation characteristics, is needed in order to develop further insights on the impact of passenger transportation on sustainable development in support of effective policy making.

In addition to carbon dioxide (CO_2), GHG includes Hydrofluorocarbons (HFCs) and nitrous oxide (N_2O), among others. In this study, only CO_2 emissions are examined since these emissions constitute 93.4% of the GHG produced in the transportation sector (EIA 2008). The CO_2 emissions focused on are those resulting from passenger travel and the impacts of travelers' choices within the context of available infrastructure and existing urban form. More specifically, various urban travel choices, travel characteristics, and travel and land-use related policies and regulations are related to annual CO_2 emissions produced as a direct result of passenger transportation. Therefore, unlike other studies, freight transportation is not considered. Moreover, CO_2 emissions resulting from the construction of transportation infrastructure and the manufacturing of passenger vehicles (private and public) are outside the scope of this study. That is, the focus is on the marginal impacts related to the passenger travel use-phase rather than the total life-cycle impacts. The rationale motivating the marginal nature of the scope of this study is to quantify relative changes in CO_2 emissions resulting from policies and regulations that might produce changes in existing conditions, a common scenario that policy-makers face.

The focus of this report is twofold. First, it describes the development of a comprehensive dataset compiled by integrating data from multiple sources on the largest 146 urban areas in the US where transit service is available. Second, preliminary modeling results – relating various urban travel characteristics to passenger transportation CO_2 emissions – are presented, with possible improvements to the model suggested.

Given that the focus of this study is on CO_2 emissions, it is important to recognize that in addition to travel supply and demand characteristics, other factors could markedly influence emissions and, consequently, the nature of the impacts travel choices might have. One such notable factor is government policy aimed at CO_2 emission regulation and mitigation. As discussed in the last section of this report, such a factor will be considered as part of next steps following this reporting period.

Finally, it is important to note that even if rich statistical models are able to capture the complexities involved with data pertaining to a large number of urban areas, such models would still be too gross in nature to support policy making for a specific urban area. Nevertheless, such models offer value in at least two regards. First, they allow for the quantification of the general effects that specified changes in important variables might have on CO_2 emissions considered jointly. This provides a tool to policy makers to explore the relative impacts of broad policies. Doing so would then guide the customization of specific policies tailored to specific urban areas, taking into account contextual considerations through detailed urban land-use and transportation modeling exercises. Second, such models may be able to offer gross predictions for certain urban areas that could bracket the magnitudes of the impacts that might be feasible to achieve, thus, either motivating and justifying further detailed analyses and policy developments, or discouraging such efforts in light of limited potential.

2. INTEGRATING DATA FROM MULTIPLE SOURCES

2.1 Sources of Data

Given that mode choice takes place in a complex environment with both competing and complementary services, the developed empirical models must effectively capture the range of transportation modes available in urban areas. Therefore, it is necessary to integrate multiple data sources in a consistent manner that could enable the estimation of robust and rich empirical models.

The focus of this study is on urban areas in the United States. The data of interest fall into three general categories relating to transit supply and use, roadway supply and private automobile use, and urban geography. As a result, the main sources of data used in this study included the following:

- 2000 US Census and Public Use Micro-data Samples (PUMS) from that decennial census (US Census Bureau 2010). These data correspond to the US Census Bureau defined "urbanized areas" (UZAs) and are collected on a time-cycle of 10 years.
- 2003 National Transit Data (NTD) (2010), a database produced by the US DOT Federal Transit Administration (FTA). These data correspond to the US Census Bureau defined UZAs and are collected annually.
- 2003 Federal-Aid Urbanized Area (FAUA) database (FHWA 2010b) produced by the USDOT Federal Highway Administration (FHWA). These data correspond to the FHWA defined "federal-aid urbanized areas" (FAUAs) and are collected annually.
- 2001 National Household Travel Survey (NHTS) (US Census Bureau 2010). These data correspond to the US Census Bureau defined UZAs and are collected on a time-cycle of 5 to 8 years.

Additional sources of data are used for certain calculations discussed subsequently. These sources are pointed out at the appropriate points in the presentation below.

Given certain discrepancies across the above sources – for example in terms of how certain related variables are defined, most critically urban area boundaries – the pertinent variables from the different sources must be integrated in a manner that avoids inconsistency as much as possible. Nevertheless, a certain degree of inconsistency will not be possible to address given certain data limitations. One important source of inconsistency is due to the differences in time cycles of the various data collection efforts. The choice of years from the different sources is 2000 or beyond, given that the most recent US Census data available at the inception of this study was 2000. Data corresponding to the year closest to 2000 from other sources are considered taking into account the time-cycle of each source. Regarding the NDT and FHWA sources above, 2003 data are considered for reasons discussed in more detail subsequently.

2.2 Variables of interest

The variables of interest in this study, which are either available in or computable from the above sources of data, are the following: CO_2/capita, transit share, transit use efficiency, average vehicle occupancy, average household vehicle ownership, lane-miles/capita, median household income, average travel time, standard deviation and coefficient of variation of travel time, and population density. The rationale behind these variables is discussed in detail in the next section. The purpose of this discussion is to emphasize and motivate the need for data integration and point out the challenges associated with such an effort. Nevertheless, it is important to note at this stage that given the focus of this study on passenger transportation, rather than all forms of transportation including freight, only data related to passenger travel is considered. Table 1 relates the variables that are of interest or are needed to calculate the variables of interest to the sources of data discussed above.

TABLE 1 Variables of Interest and their Sources

Variables	US Census	FTA (NTD)	FHWA	NHTS
Transit Energy Consumption		x		
Transit Passenger-Miles		x		
Transit Space-Miles		x		
Lane-Miles			x	
Total vehicle miles traveled (VMT)			x	
Average Vehicle Occupancy	x			
Distribution of Travel Time	x			
Median Income	x			
Area	x	x	x	x
Population	x	x	x	x

Notice that different variables are available from different sources. And, as already discussed, the various sources assume certain urban boundary definitions that are not consistent across certain sources. Therefore, it is critical to transform the data in a manner that reduces the spatial inconsistencies as much as possible. Fortunately, certain variables, notably population and area, are available across most data sources. This proves instrumental in developing meaningful transformations. Doing so is discussed in detail subsequently.

2.3 Integration of data and treatment of spatial inconsistencies

As already discussed above, three of the main sources of data are the US Census (2010), NTD (2010), and FHWA (2010b). The Census and NTD sources both have their data available at the UZA level. Prior to 2002, NTD followed the UZA boundaries of the 1990 Census. The US

Census Bureau redefined the UZA boundaries in 2000 and applied them to the data collected in 2000. However, FHWA's data do not correspond to UZA boundaries. The closest spatial boundaries to those of the UZAs the FHWA data are available in are the "federally aided urbanized areas" (FAUAs), which are modified versions of UZAs. Prior to 2003, the FAUA boundaries are modifications of UZA boundaries of 1990. Since 2003, most FAUA boundaries started to adopt modifications of the UZA boundaries of 2000, while some were still modifications of 1990 UZAs. To ensure that the urban area boundaries across the data sources are more comparable to one another, FAUA and NTD 2003 data instead of 2000 data are considered. Among all 464 UZAs, 344 have reported data in all three data sources. This is because some UZAs do not have NTD data. Given the focus on fairly large urban areas in this study, among the 344 UZAs only the ones with populations greater than 150,000 are considered further, resulting in 177 UZAs.

Although for the most part the FAUA 2003 boundaries reflect minor modification of the UZA 2000 boundaries, some exploratory analysis shows that there are still some appreciable differences between certain FAUA and UZA boundaries. To overcome this inconsistency, the following investigations and transformations are performed.

Screening potential inconsistent areas based on population and area

As indicated in Table 1, both FHWA and the US Census Bureau sources have information regarding population and area. Therefore, the FAUAs and UZAs are first compared on the basis of both population and area. The cases where the ratio of the two population values falls outside the 0.9 to 1.1 range or the ratio of the two area values falls outside the 0.6 to 1.7 range are identified as reflecting potential inconsistencies. These bounds of the ranges were determined by examining the empirical cumulative density functions (ECDFs) of the population and area ratios. The reason the acceptable range for the area ratio is larger than that of the population ratio is that consistency in population is more important than consistency in area because most variables of interest are considered on a per capita basis. Among the 177 UZAs, 74 are identified as potentially inconsistent based on the above criterion. The treatment of these 74 cases is described next. As for the complement of these cases, the variable values associated with FAUAs are transformed by factoring the variables by the ratio of the UZA population to the FAUA population to account for the small differences between the two.

Geographical visualization of potentially inconsistent UZAs

Reading UZA and FAUA shape files (National Transportation Atlas Database 2003) in a Geographic Information System (GIS) software, the 74 identified potentially inconsistent areas are visualized. The following four categories of overall patterns describe the nature of the inconsistencies found:

- Pattern 1: The FAUA and corresponding UZA predominantly overlap with one another.
- Pattern 2: The FAUA includes several UZAs, which as a group predominantly overlap with the FAUA.
- Pattern 3: The UZA is included in the FAUA and the FAUA is appreciably larger than the UZA.

- Pattern 4: The FAUA and UZA intersect with non-trivial areas belonging to only one or the other.

Different integration treatments are applied for the various identified patterns. For pattern 1, since the difference between the FAUA and UZA boundaries is limited, the variable values associated with FAUAs are transformed by factoring the variables by the ratio of the UZA population to the FAUA population. As for pattern 2, since the FAUA includes 2 or more UZAs, data related to those UZAs are aggregated to represent a new area corresponding to the FAUA. The resulting aggregated variables are then transformed by factoring the FAUA variables by the ratio of the aggregated UZA populations to the FAUA population. Regarding pattern 3, the areas within the FAUA boundaries that fall outside the corresponding UZA boundaries in five cases constitute mostly bodies of water (*e.g.*, Palm Bay, FL, and Virginia Beach, VA) and, therefore, for these cases the boundaries are considered predominantly overlapping and, therefore, the UZA's area is considered and the variable values associated with FAUAs are transformed in the same manner as described above.

After all of the above transformations are applied, 13 inconsistent cases remain. Six of those have UZA population values less than 166,000. Given that these population values fall just above the previously applied threshold of 150,000 representing fairly large urbanized areas, these six inconsistent cases are not considered further on the basis of their relatively small populations. There are an additional 6 consistent cases where the population values are between 150,000 and 166,000. To maintain consistency, these cases are also not considered further, effectively redefining the threshold for fairly large urbanized areas at a population value of 166,000 (instead of 150,000). At this stage, seven inconsistent cases remain with population values ranging from 190,000 to 225,000. No reasonable transformations for these cases were feasible and, therefore, the corresponding urbanized areas are not considered further. As a result, a total of 158 urbanized areas remain in the data set where the variables are transformed to account for spatial inconsistencies.

For these 158 areas, some of them do not have transit energy consumption data or transit service data. Since these are important for the analysis, these urban areas were excluded. So the final number of areas is 146.

2.4 Calculation of variables not directly observed

Several variables of interest are not readily available in the data sources discussed above in a manner consistent with the UZA boundaries. These variables are: population density, CO_2 emissions due to transportation, transit share, average travel time, and the standard deviation (and coefficient of variation) of travel time. The density is readily calculated as the ratio of UZA population to UZA area. Transit market share is represented by the ratio of transit passenger-miles to transit and vehicle passenger-miles traveled for each urbanized area. This calculation is straightforward based on the variables available in the dataset. The average travel time can be calculated directly from the US Census data (US Census Bureau 2010), as the travel time related data include the total travel time of all commuters and the number of commuters. The calculations of CO_2 emissions and the standard deviation (and coefficient of variation) of travel time are more involved. In what follows, these calculations from data available from the sources discussed above supplemented by data from additional sources are explained.

CO_2 emissions

The total CO_2 emissions due to transportation for each urbanized area on the basis of UZA boundaries is derived on the basis of the vehicle miles traveled (VMT) information provided by FHWA and public transportation energy consumption information provided by FTA. Regarding private automobile travel, the basic idea is to estimate the total energy consumption from the VMT information. Then, for both private automobile and public transportation travel, the total CO_2 emissions contributed by transportation are derived from the energy consumption values.

To transform the VMT to energy consumption, two issues need to be addressed. First, the total VMT available includes the VMT contributed by all vehicle types including private and public vehicles and trucks. As already discussed, the VMT contributed by the latter two types is not of interest in this study and, therefore, their contributions to VMT must be subtracted from the total VMT. Second, different types of vehicles have different miles per gallon (MPG) and use different fuel types. Therefore, the private auto VMT must be broken down by type of vehicle in terms of MPG and fuel type they use.

Regarding the first issue, each US state provides the breakdown of total registered vehicles into private vehicles, public vehicles, and trucks (FHWA 2010a). Therefore, for each urbanized area, the percentage of private vehicles in the state the urbanized area belongs to is applied to the VMT of that urbanized area to calculate its VMT contributed by private vehicles. This calculation assumes that the distance traveled by each type is not dependent on the vehicle type, which is not likely to be the case especially when it comes to commercial trucks. However, no data were readily available from the sources considered to avoid this assumption. Therefore, considering this additional complexity is left for future research.

Regarding the second issue, the breakdown of private vehicles into different types such as cars, vans, and SUVs is not available at a fine spatial resolution from the investigated data sources. Therefore, national data available from NHTS (2010) are used instead. The data consists of the VMT contributed by different types of private vehicles using different types of fuels at the national level. This national breakdown of VMT is applied to all urbanized areas. Thus, the VMT contributed by different vehicle types using different fuel types is derived for each urbanized area. These VMT values for each urbanized area are used to calculate fuel consumption based on MPG values for the various private vehicle and fuel types. These MPG values are available from the Environmental Protection Agency (EPA) (2010b).

By adding the fuel consumption by public transportation to the derived fuel consumption of private vehicles, the total fuel consumption by different fuel types due to private vehicles and public transportation are available at this stage. Calculating CO_2 emissions from these fuel consumption values for each urbanized area is done by applying a set of equations and parameters provided by the EPA (2011, 2010a).

Standard Deviation (and coefficient of variation) of Travel time

The US Census data (US Census Bureau 2010) include the frequency distribution of travel time reported in terms of 12 bins, from less than 5 minutes to 90 or more minutes, giving the number of travelers into each bin. The standard deviation of travel time is determined from this frequency distribution assuming that the travel times within the first 11 bins are uniformly distributed and the top bin is centerd at exactly 90 minutes. The estimated coefficient of variation (CV) (a unit-less measure defined as the ratio of the standard deviation to the mean) of travel time is obtained from the mean and standard deviation as determined from the frequency distribution.

3. EXPLORATORY ANALYSIS AND PRELIMINARY MODEL ESTIMATION

3.1 Variables of interest

The pertinent variables of interest and the nature of their relevance in the context of this study are discussed in what follows. The dependent variable of interest is the CO_2 emissions produced by urbanized areas as a direct result of passenger transportation using all modes of travel. As discussed above, freight transportation is not included given that the focus of this study is on relationships that could support policy making decisions regarding passenger travel. The units used are metric tons CO_2/year and the determined CO_2 emissions are normalized by the total population of the urbanized area.

As already discussed, the transit market share is represented by the ratio of transit passenger-miles to the total passenger-miles. Given that CO_2 emissions are dependent on energy consumption, which in turn is dependent on distance traveled, it is important to include distance traveled in this transit market share variable. Given the efficiencies that transit could offer, an increase in this variable is expected to result in a reduction in CO_2 emissions. Transit service utilization, as measured by the ratio of passenger miles traveled to the total space miles offered, is an important variable because if transit utilization is low, the advantages offered by the transit mode given its "mass" use would be lost. Therefore, an increase in efficiency is expected to reduce CO_2 emissions.

The supply of infrastructure for private vehicles of travel could also have an important effect on CO_2 emissions. This variable is defined as lane-miles per capita. An increase in this variable results in a greater supply of roadways for private auto use, which likely increases the reliance on this mode, producing higher CO_2 emissions as a result. Private auto occupancy, on the other hand, would have the opposite effect because the marginal increase in energy consumption and CO_2 emissions due to additional passengers in a private automobile is very low. Finally, a higher number of vehicles owned per person is expected to increase CO_2 emissions because there will be more people that have cars, the more likely they will drive them.

An increase in average travel time for commuter trips across all modes is pertinent in that an increase in this variable is expected to increase CO_2 emissions due to travelers taking longer trips in an urbanized area. A high standard deviation of travel time means that there are some very long and very short travel times. Since travel time is bounded at zero, a higher standard deviation, likely means that there are more very long trips taking place, leading to an increase in CO_2 emissions. The coefficient of variation of travel time could also be pertinent. The coefficient of variation of travel time accounts for a non-linear effect of travel time on CO_2/capita, and its pair-wise correlation with the response variable is negative. Therefore, an increase in the coefficient of variation of travel time is likely to lead to lower CO_2 emissions.

Population density is expected to be pertinent because an increase in density could relate to more travelers traveling shorter distances, which would lead to reduced CO_2 emissions. On the other hand, a higher median household income could suggest the presence of more travel and an increased likelihood of using the private auto mode, resulting in an increase in CO_2 emissions. An increase in area (square miles) of the urbanized area is expected to increase CO_2 emissions because there will potentially be popular destinations far away from where people live, causing an increase in auto travel.

3.2 Summary statistics

Table 2 provides some summary statistics of variables that are thought to be important in modeling CO_2 emissions. Also included is the correlation coefficient with the variable of interest, CO_2/capita. While the signs of the coefficient are consistent with *a priori* expectations, not all of the variables in the table exhibit high correlations with CO_2/capita in absolute terms, transit service utilization being a notable example. Also notice that the coefficient of variation of travel time and median income have moderate correlations with CO_2/capita.

TABLE 2 Summary Statistics of Pertinent Variables

Variable	**Min**	**Max**	**Mean**	**Median**	**Std. Dev.**	**Corr. with CO_2/cap.**
CO_2/Capita (metric tons CO_2/yr)	0.725	3.191	1.535	1.525	0.365	1
Transit share	0.000065	0.166	0.0155	0.00833	0.0206	– 0.171
Area (sq. mi.)	43	3353	366.3	215	461.3	0.170
Density (persons/ sq. mi.)	852	7068	2490	2321	980	– 0.319
Avg. Travel Time (minutes)	17.86	76.04	32.53	29.57	11.00	0.164
Std. Dev. Travel Time	12.94	29.38	17.25	16.55	2.96	0.076
CV of Travel Time	0.612	0.875	0.717	0.703	0.059	– 0.186
Vehicle Occupancy (persons/veh.)	1.093	1.382	1.182	1.175	0.050	– 0.278
Vehicle Ownership	0.425	0.745	0.634	0.644	0.054	0.146
Transit Service Utilization	0.023	0.273	0.112	0.112	0.038	– 0.091
Lane-miles/Capita	0.053	1.517	0.646	0.635	0.254	0.538
Medn. Household Income ($)	28975	74133	42380	41373	7028.7	0.217

To look further into the type of bivariate relationships the explanatory variables with relatively high correlation coefficients in absolute terms have with CO_2 emissions, bivariate scatter plots are investigated as shown in Figure 1. The graph displays the relationship of select variables with CO_2/capita and also the relationships with each other.

Based on Figure 1, all of the examined variables exhibit relationships with CO_2/capita that are consistent with the *a priori* expectations discussed above. The scatter plots indicate that lane-miles/capita has the strongest relationship with CO_2/capita. Transit share seems to have a somewhat weak negative relationship with CO_2/capita, possibly indicating that transit share may not have as big of an effect as the other variables. Finally, because density has a non-linear, negative relationship with CO_2/capita, density was transformed to 1/density for all further modeling. Karathodorou et al. (2009) also found the same relationship for international cities.

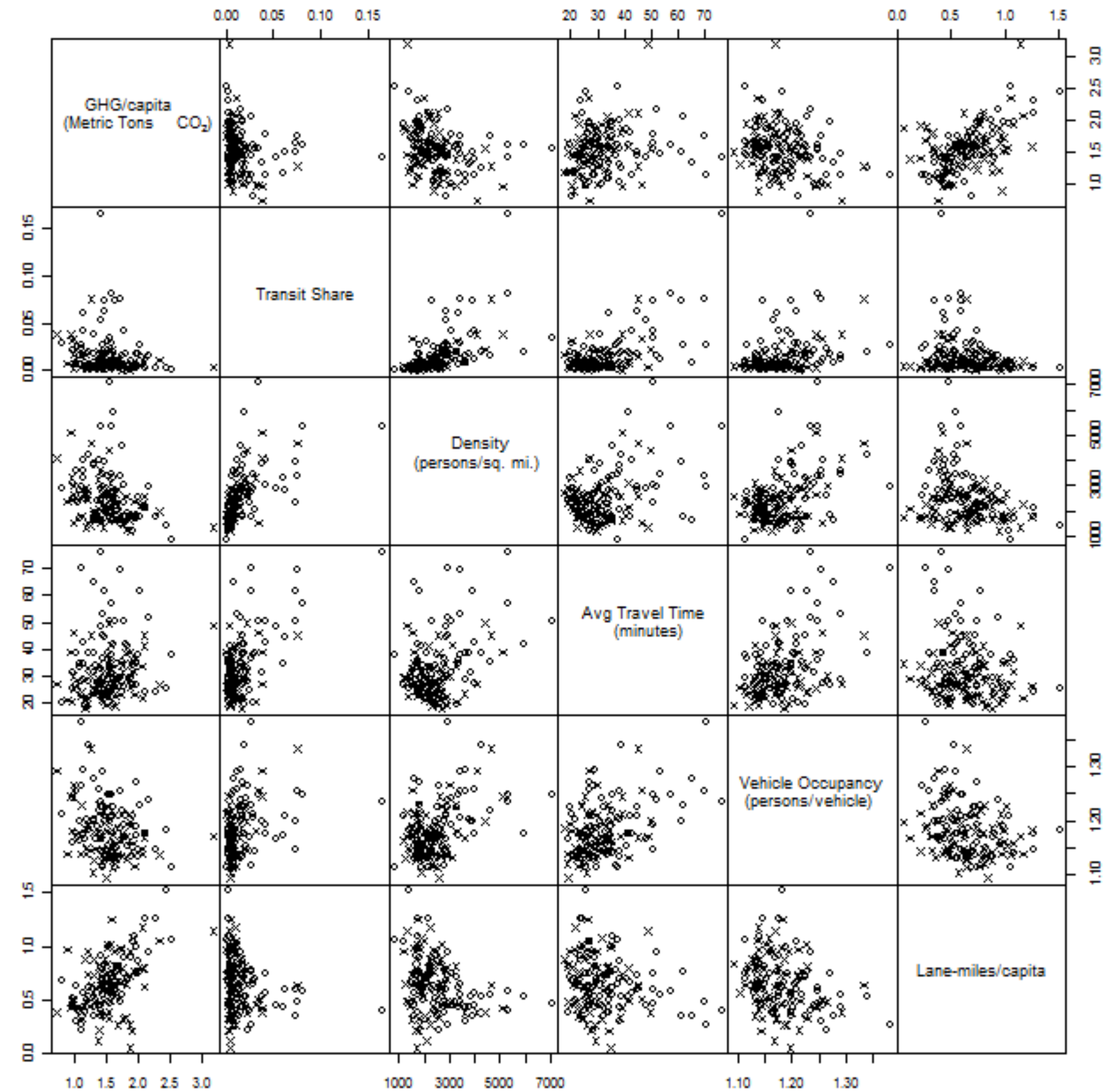

FIGURE 1 Scatter plots of a subset of the pertinent variables.

3.3 Preliminary Model Estimation Results

The impact of transit related variables, such as transit share and transit use efficiency, on CO_2 emissions is of particular interest and, therefore, these variables are considered in estimating a preliminary model. Additional select variables among the ones discussed above are also considered with others (some discussed and some not discussed above) to be explored as part of next steps as discussed in the last section of this report.

A preliminary linear regression model with CO_2/capita as the dependent variable is estimated using Ordinary Least Squares. The explanatory variables included are transit share, 1/density, average travel time, standard deviation of travel time, average vehicle occupancy, transit service utilization, area, and lane-miles/capita. The estimation results are shown in Table 3. This preliminary model fits the data reasonably well and the coefficient of transit share is clearly significant. Additional variables whose coefficients are significant are lane-miles/capita, average travel time, vehicle occupancy, and the reciprocal of population density. Nevertheless, the estimated model exhibits certain concerns the point to substantial room for improvements as discussed in what follows and in the next section.

It is believed that transit service utilization should be significant and the sign of its coefficient should be negative (higher utilization reduces CO_2/capita), but neither of these is the case. Since transit service utilization has such a narrow range, a transformation (such as a log transformation) may be a possible solution to this limitation. Another concern with the preliminary model is that the standard deviation of travel time is somewhat significant but has a counterintuitive sign. This could be a result of the multicollinearity between the average and standard deviation of travel time. To avoid the counterintuitive sign, the standard deviation could be replaced with the coefficient of variation of travel time, which is not as highly correlated with the average travel time. In addition, while the coefficient of the variable area has the expected sign, the estimate is not significant.

TABLE 3 Regression Model for CO_2/capita

Variable	Coefficient	SE of Coeff.	t-statistic	p-value
Constant	2.3575	0.6751	3.49	0.001
Transit share	–5.909	1.698	–3.48	0.001
Lane-miles/Capita	0.62008	0.09019	6.88	0.000
Avg. Travel Time	0.021754	0.005311	4.1	0.000
Std. Dev. Travel Time	–0.02561	0.01506	–1.7	0.091
Vehicle Occupancy	–1.5051	0.5919	–2.54	0.012
Transit Service Utilization	0.9203	0.7092	1.3	0.197
Area	7.84E-05	8.75E-05	0.9	0.372
1/Density	549.4	177.3	3.1	0.002
Sample Size = 146, R-Sq (adj) = 51.9%				

4. NEXT STEPS

There are some steps that can be taken next to enhance the reliability of both the dataset and model of interest. The data for certain variables need to be further verified and cross-checked to make sure that conclusions made are valid. In addition, incorporating a few more variables in the dataset that could improve the explanatory power of the model is desirable. For example, a measure of the variability in population density and the presence of automobile emissions inspection programs are of interest. Further improvement of the model could come from including additional and already available explanatory variables. The variables median household income and average household vehicle ownership are shown to have an influence on CO_2/capita in the exploratory analysis but have not been considered yet.

In addition, there are various interactions that could be investigated. One aspect that has not yet been considered is the influence of government policies and regulations pertaining to CO_2 emissions on CO_2/capita. The presence of automobile emissions inspection programs could be an indicator of such policies and regulations (even though itself is not aimed at mitigating CO_2 emissions). It could also be the case that the level of CO_2/capita in an urbanized area influences the public policies and regulations that are put in place. Therefore, the inclusion of an explanatory variable, such as the presence of automobile emissions inspection programs as an indicator, would result in simultaneity that would have to be addressed in specifying and estimating the model of interest. All of the above reflect further dataset development and modeling considerations that will be considered following this reporting period.

5. REFERENCES

1. Bento, A., M. Cropper, A. Mobarak, and K. Vinha. The Effects of Urban Spatial Structure on Travel Demand in the United States. The Review of Economics and Statistics, 87(3), 2005, pp. 466-478.

2. Brown, M. A., F. Southworth, and A. Sarzynski. Shrinking the carbon footprint of metropolitan America. Brookings Institution, Washington, D.C., 2008.

3. Cambridge Systematics, Inc. Moving Cooler: An Analysis of Transportation Strategies for Reducing Greenhouse Gas Emissions. Prepared by for the Urban Land Institute, Washington, D.C., 2009.

4. Energy Information Administration. Emissions of Greenhouse Gases in the United States 2007. Report No. DOE/EIA-0573, Washington, D.C., 2008.

5. Ewing, R., K. Bartholomew, S. Winkelman, J. Walters, and D. Chen. Growing Cooler: Evidence on Urban Development and Climate Change, Urban Land Institute, Washington, D.C., 2007.

6. Federal Highway Aminisration, US Department of Transportation. Motor Vehicles. URL: http://www.fhwa.dot.gov/policy/ohim/hs03/mv.htm, accessed 2010a.

7. Federal Highway Administration, US Department of Transportation. Roadway Extent, Characteristics and Performance. URL: http://www.fhwa.dot.gov/policy/ohim/hs03/re.htm, accessed 2010b.

8. Glaesar, E., and M. Kahn. The Greenness of Cities. Rappaport Institute for Greater Boston and Taubman Center for State and Local Government Policy Briefs, 2008.

9. Hankey, S., and J. Marshall. Impacts of Urban Form on Future Passenger-vehicle Greenhouse Gas Emissions. Energy Policy 38, 2010, pp. 4880-4887.

10. Holtzclaw, J, R. Clear, H. Dittmar, D. Goldstein, and P. Haas. Location Efficiency: Neighborhood and Socio-Economic Characteristics Determine Auto Ownership and Use – Studies in Chicago, Los Angeles and San Francisco. Transportation Planning and Technology 25(1), 2002, pp. 1-27.

11. Karathodorou, N., D.J. Graham, and R.B. Noland. The Effects of Urban Density on Fuel Demand, Presented at the 88th Annual Meeting of the Transportation Research Board, Washington D.C., 2009.

12. Kockelman, K. M. Bomberg, M. Thompson, and C. Whitehead. CO_2 Emission Control Options: Opportunities for Conservation, Special Report 298 prepared for the National Academy of Science, 2009.

13. Lomax, T., V. Stover, G. Carlin, R. Jensen, D. Landmark, J. Pate, D. Schrank, E. Sulak, and S. Turner. Evaluating Mobility and Energy Efficiency, Technical Report for Texas Transportation Institute, 1994.

14. Massachusetts Institute of Technology. Urban Public Transportation Roundtable: Transit Management and Operations Workshop, held on Sept. 16-17, 2009.

15. National Household Travel Survey, Federal Highway Aminisration, US Department of Transportation. URL: http://nhts.ornl.gov/, accessed 2010.

16. National Transit Data. URL: http://www.ntdprogram.gov/ntdprogram/data.htm, accessed 2010.

17. National Transportation Atlas Database (electronic resources). DC ROM, 2003.

18. Newman, P., and J. Kenworthy. Cities and Automobile Dependence: An international Sourcebook, Gower Publishing Company Limited, VT, 1989.

19. Parshall, L., K. Gurney, S. Hammer, D. Mendoza, Y. Zhou, and S. Geethakumar. Modeling Energy Consumption and $CO2$ Emissions at the Urban Scale: Methodological Challenges and Insights from the United States. Energy Policy 38, 2010, pp. 4765-4782.

20. Schipper, L., E. Deakin, and C. McAndrews. Carbon Dioxide Emissions from Road Transport in Latin America. Chapter 4 in Climate and Transportation Solutions: Findings from the 2009 Asilomar Conference on Transportation and Energy Policy, D. Sperling and J.S. Cannon, Eds. Institute of Transportation Studies, University of California, Davis, 2010, pp. 33-47.

21. ULI – the Urban Land Institute. The City in 2050: Creating Blueprints for Change. Urban Land Institute, Washington, D.C., 2008.

22. University Transportation Research Center – Region II (2010). Sustainable Transit: Developing an Action Agenda. Symposium, The City College of New York. URL: http://www.utrc2.org/events/events.php?viewid=264, May 7, 2010.

23. US Census Bureau. URL: http://www.census.gov/, accessed 2010.

24. US Environmental Protection Agency. Green Power Equivalency Calculator Methodologies. URL: http://www.epa.gov/greenpower/pubs/calcmeth.htm, accessed 2010a.

25. US Environmental Protection Agency. Inventory of U.S. Greenhouse Gas Emissions And Sinks: 1990 – 2009. Report No. EPA 430-R-11-005, Washington, D.C., 2011.

26. US Environmental Protection Agency. Light-Duty Automotive Technology, Carbon Dioxide Emissions, and Fuel Economy Trends: 1975 Through 2009. Report No. EPA-420-R-10-023, Washington, D.C., 2010b.