sustainability

MDPI

Article

The Relationship between On-Road $FFCO_2$ Emissions and Socio-Economic/Urban Form Factors for Global Cities: Significance, Robustness and Implications

Yang Song * and Kevin R. Gurney

School of Informatics, Computing, and Cyber Systems, Northern Arizona University, Flagstaff, AZ 86011, USA; kevin.gurney@nau.edu

* Correspondence: ys369@nau.edu

Received: 2 July 2020; Accepted: 23 July 2020; Published: 27 July 2020

Abstract: Transportation accounts for 18% of global fossil fuel carbon dioxide ($FFCO_2$) emissions, especially in urban areas. An improved understanding of on-road $FFCO_2$ emissions is essential to both carbon science and mitigation policy. Previous studies have identified the driving factors and quantified their relationship to on-road $FFCO_2$ emissions. However, they have been primarily based on case studies conducted in individual cities, and the research results remain inconclusive due to the considerable heterogeneity of cities and associated outcomes. In order to achieve more general results and to further understand their uncertainties, this study explored the relationships between socio-economic/urban form data and self-reported on-road $FFCO_2$ emissions for a sample of global cities based on the adjusted Stochastic Impacts by Regression on Population, Affluence and Technology (STIRPAT) model. The robustness and sensitivity of these relationships was evaluated by introducing artificial errors, conducting cross-validation, and examining various model specifications. Results indicated that fuel economy (p-value $< 3.1 \times 10^{-8}$), vehicle ownership (p-value $< 3.0 \times 10^{-4}$), road density (p-value $< 4.4 \times 10^{-3}$) and population density (p-value $< 3.1 \times 10^{-3}$) were statistically significant factors that correlate with on-road $FFCO_2$ emissions. Of these four variables, fuel economy and vehicle ownership had the most robust relationships. These results offer potential policy insights into on-road $FFCO_2$ emissions mitigation in cities, in addition to offering a means to generate emissions estimates without detailed bottom-up information.

Keywords: on-road emission; socio-economic/urban form factors; sensitivity analysis; global cities

1. Introduction

On-road transportation provides mobility and accessibility for people and enables the movement of freight, which is an essential part of the economic activity [1]. On-road fossil fuel carbon dioxide ($FFCO_2$) emissions often represent the largest single emitting sector in urban areas and account for 18% of the total $FFCO_2$ emissions worldwide [2]. In the US, on-road $FFCO_2$ emissions constitute 43% of the total $FFCO_2$ in major metropolitan areas [3]. This makes on-road $FFCO_2$ emissions one of the top priorities for greenhouse gas (GHG) mitigation in the urban domain.

Understanding the relationship between on-road $FFCO_2$ emissions and their driving factors is critical to both social science and urban policy. Many researches have investigated and identified relationships between certain socio-economic factors and transportation energy use. For example, Lakshmanan and Han [4] identified growth for travel demand, population, and gross domestic product (GDP) as the three most important factors related to transportation energy use and $FFCO_2$ emissions during 1970–1991 in the US. Cervero and Hansen [5] concluded that travel demand and road supply cause a mutually reinforcing feedback in which road investment induces travel and increasing travel

requires more road investment. Lu et al. [6] reported that economic growth and vehicle ownership were the most important factors responsible for the increase of highway vehicle $FFCO_2$ emissions in Germany, Japan, South Korea and Taiwan during the 1990–2002 time period. Liddle [7] examined the vehicle miles travelled (VMT), income, fuel price and vehicle ownership within the US from 1946 to 2006. He concluded that "US mobility demand has a long-run systemic, mutually causal relationship with gasoline price, income, and vehicle ownership". Timilsina and Shrestha [8] found that per capita GDP, population growth and transportation energy intensity were the major driving factors for the growth of transportation $FFCO_2$ emissions in selected Asian countries during 1980–2005. Barla et al. [9] utilized survey data to explore the impacts of individual/household socio-economic characteristics on transportation greenhouse gas emissions. They found relationships (statistically significant at 1% level) between the transportation GHG emissions and population age (emissions start to decline after the age of 50), gender (females produce 25% less emissions than males), income (high income links to high emissions) and neighborhood (a 10% denser neighborhood would reduce 1.2% emissions on average). Wang et al. [10] found that the growth of per capita GDP and energy consumption intensity were primarily responsible for the growth of $FFCO_2$ emissions in the transportation sector over the 1985–2009 time period in China.

Certain socio-economic factors were identified in previous studies to have a significant relationship to transportation energy use. Their resulting elasticity values (the proportional change of one variable in response to the proportional change of another variable) exhibited a wide range in the different locations, mix of variables considered and different interpretations of the key variables (Table 1).

Table 1. Elasticities between socio-economic factors and transportation demand reported in previous research.

Previous Studies.	Elasticity of	Elasticity to	Elasticity Value	Study Area
Lakshmanan and Han [4]	Population (Unit: people)	Passenger travel demand (Unit: passenger-miles)	2.8	United States
		Freight travel demand(Unit: ton-miles)	1.9	
	GDP (Unit: dollars)	Passenger travel demand (Unit: passenger-miles)	1.12	
		Freight travel demand (Unit: ton-miles)	1.47	
Cervero and Hansen [5]	Population (Unit: people)	Vehicle Miles Traveled (Unit: vehicle-miles)	0.69	California, USA
	Income (Unit: dollars per capita)		0.294	
Lu et al. [6]	GDP (Unit: billion US dollars)	Transportation CO_2 emission (Unit: million metric tons)	0.87	Taiwan, Germany, Japan, South Korea
Newman and Kenworthy [11]	Short-term income (Unit: dollars in PPP)	Gasoline Consumption (Unit: gallons per capita)	0.11	N/A [a]
	Long-term income (Unit: dollars in PPP)		0.6	
Eltony and Al-Mutairi [12]	Short-term income (Unit: N/A [b])	Fuel demand (Unit: N/A [b])	0.47	Kuwait
	Long-term income (Unit: N/A [b])		0.93	
Espey [13]	Income (Unit: N/A [c])	Fuel demand (Unit: N/A [c])	0.39–0.81	N/A[c]

[a] Newman and Kenworthy [11] cited these elasticities from previous studies at various places; [b] No units were provided for income and fuel demand by Eltony and Al-Mutairi [12]; [c] Espey [13] summarized this range of elasticities from previous studies at various places.

There is also an extensive literature on the relationship between on-road $FFCO_2$ emissions (or closely related variables) and aspects of urban form or "urban form factors". For example, whether a compact city form is more or less eco-efficient has been actively discussed in sustainability science. Newman and Kenworthy [11] found an inverse relationship between gasoline consumption and population density based on a database of global cities. Cervero & Landis [14] argued that a decentralized urban pattern undermines the public transit service and causes a mismatch in the job-housing balance, which indirectly encourages private driving and results in more VMT. Wang et al. [15] analyzed urbanization and on-road $FFCO_2$ emissions data in Beijing from 2000 to 2009 and found that the decentralization of Beijing had a strong aggravating impact on on-road $FFCO_2$ emissions due to the increasing commute distances and shifting travel mode to private cars instead of public transit.

There are other studies that suggest that compact cities may not be necessarily energy efficient. For instance, compact cities can induce traffic congestion, which lowers the fuel economy, resulting in added fuel consumption [16]. Compact cities may not necessarily reduce VMT because the lack of access to nearby open spaces induces leisure travel demand [17]. Many case studies also did not observe a significant relationship between increasing urban compactness and decreased transportation energy consumption. A case study in the Netherlands by Bouwman and Moll [18] found a 5% difference in the average personal transportation energy use across different urban spatial structures. They concluded that, at least in the Netherlands, the expected negative relationship between urban compactness and personal transportation energy consumption was not observed. In Britain, Echenique et al. [19] conducted scenario analyses on three regions based on travel demand models and found that the impact of the urban spatial structure on transportation energy consumption is negligible compared to other socio-economic factors and population growth.

As for the relationships between socio-economic factors and on-road $FFCO_2$ emissions, previous research also reported a wide range of elasticities between urban form factors and various on-road transportation measures (Table 2).

Table 2. Elasticities between urban form factors and transportation demand reported in previous research.

Previous Studies	Elasticity of	Elasticity to	Elasticity Value	Study Area
Cervero and Hansen [5]	Lane Miles (Unit: lane-miles)	Vehicle Miles Traveled (Unit: vehicle-miles)	0.59	Counties in California, USA
Barla et al. [9]	Residential density (Unit: residences per km^2)	Transportation GHG emissions (Unit: grams of CO_2 equivalent)	−0.12 to −0.07	Quebec, Canada
Hansen and Huang [20]	Lane Miles (short term) (Unit: lane-miles)	Vehicle Miles Traveled (Unit: vehicle-miles)	0.3	California metropolitan areas, USA
	Lane Miles (long term) (Unit: lane-miles)		0.9	
Noland and Lem [21]	Lane Miles (short term) (Unit: lane-miles)	Vehicle Miles Traveled (Unit: vehicle-miles)	0.3–0.5	United Kingdom and United States
	Lane Miles (long term) (Unit: lane-miles)		0.7–1.0	
Fulton et al. [22]	Lane Miles (short term) (Unit: lane-miles)	Vehicle Miles Traveled (Unit: vehicle-miles)	0.1–0.4	Counties in Atlanta, USA
	Lane Miles (long term) (Unit: lane-miles)		0.5–0.8	
Johansson and Schipper [23]	Population density (Unit: people per km^2)	Annual driving distance (Unit: km/year)	−1.75 to −0.3	Organisation for Economic Co-operation and Development (OECD) countries
Ewing et al. [24]	Population density (Unit: unknown [a])	Vehicle miles travelled (Unit: unknown [a])	0.075	N/A [a]
Karathodorou et al. [25]	Population density (Unit: people per hectare)	Fuel Consumption per capita (Unit: unknown [b])	−0.334	World cities
Lee and Lee [26]	Population density (Unit: people per acre)	$FFCO_2$ from private cars (Unit: lbs)	−0.224	Urbanized areas in US

[a] Ewing et al. [24] estimated this elasticity number based on 23 previous studies; [b] Karathodorou et al. [25] did not address the unit of fuel consumption in their paper.

To sum up, previous studies on the relationship between socio-economic/urban form factors and transportation energy consumption or $FFCO_2$-related emissions remain inconclusive due to limitations in the sample size (i.e., single city studies), differences in the precise metric examined and underlying model construction. Moreover, previous results also showed a wide range of elasticities of various determining factors to the environmental impact, possibly due to the heterogeneity of the study areas and data sources. This study, by contrast, aims to examine a large cross-section of cities at the global scale with a multivariate approach, to quantitatively estimate the relationships between $FFCO_2$ emissions and key socio-economic/urban form factors and assess their robustness.

2. Methods and Data

The study area considered in this research consists of a series of global cities. These cities were selected according to four criteria. First, they must cover a wide range of socio-economic and urban form characteristics to reduce the sampling bias. Second, they must have an area that is large enough for the calculation of certain urban form indices (such as population centrality). Third, they must be the industrial, financial or political center of the surrounding regions. Finally, they must have reported $FFCO_2$ emissions available in existing databases. Using data meeting these criteria from the World Bank, the C40 project and the Millennium City Database (MCD), 58 global cities (Appendix A Figure A1 and Table A1) were selected for this study [27–29].

The foundation of the analysis used here is the IPAT model relating environmental impact (I) to a multiplicative product of population (P), affluence (A) and technology (T) [30]. Through this form ($I = P \times A \times T$), the IPAT model describes the relationship among these variables as interdependent [31].

A weakness of the IPAT model is that it is a conceptual framework that cannot be tested or falsified. Dietz and Rosa [32,33] modified the IPAT model into the Stochastic Impacts by Regression on Population, Affluence and Technology (STIRPAT) model. A scaling factor was added, resulting in

$$I = aP^bA^cT^de \tag{1}$$

where a is the scaling factor, e is the error term and b, c and d are the exponents of P, A and T.

This formula can be converted into a linear form with a log transformation.

$$Ln(I) = bLn(P) + cLn(A) + dLn(T) + e \tag{2}$$

This form allows quantitative analyses such as linear multiple regression to be performed and the results to be evaluated. Therefore, it can be used to test hypotheses regarding causal relationships.

Another weakness of the IPAT model is that it assumes that the relationship between I and each of the driving factors is proportional and monotonic, which may not be true in reality [31]. On the other hand, a STIRPAT model with exponential parameters allows non-proportional and non-monotonic relationships to be falsified [31]. However, limitations still exist for the STIRPAT model, as it assumes that there is no residual non-linearity after the log transformation.

Based on previous literature, this study uses on-road $FFCO_2$ emission data as the environmental impact (I); Population density to represent P; Affluence (A) is represented by per capita GDP; and a variety of socioeconomic/urban form factors to represent technology (T) in the STIRPAT model. The socio-economic/urban form factors included road density, road length per capita and population centrality, miles per gallon and vehicles owned per capita. These factors were chosen because of their correlations with the energy consumption in the previous literature. A summary of the regression variables and their sources is provided in Tables 3 and 4. Data quality was examined for all of the input variables. For example, the population density data had an 87.8% accuracy when validated by the United States census data [34]. Road network data, on average, had errors from ground truth amounting to less than 6 m in England [35] and France [36]. The on-road $FFCO_2$ emission data were collected from the C40 project [23] and the Millennium City Database [37]. There were no quantitative

quality evaluations, but only qualitative descriptions for the remainder of the data. Additional details regarding the data sources and data uncertainties are included in Appendixes A.1–A.3.

Table 3. Urban form and socio-economic variables used in this study and their references.

Regression Variables		Sources
Transportation $FFCO_2$ emissions		C40 project [23] and millennium city database [37]
Urban form factors	Population density (*PD*)	Spatial calculation based on LandScan [38]
	Road length per capita (*RLPC*)	Spatial calculation based on OpenStreetMap [39]
	Road density (*RD*)	Spatial calculation based on OpenStreetMap [39]
	Population centrality (*PC*)	Spatial calculation based on LandScan [38], OpenStreetMap [39], and global impervious surface area (ISA) [40]
Socio-economic factors	City GDP per capita (*GDPPC*)	Global Metro Monitor [41]
	Fuel economy (Miles per gallon, *MPG*)	Global Fuel Economy Initiative [42]
	Vehicles owned per capita (*VOPC*)	World Road Statistics [43]

Table 4. Statistics on the Urban Form/Socio-Economic Data for 58 Global Cities.

	Mean	Median	Standard Deviation	Min	Max	Coefficient of Variation	Coefficient of Skewness *
Vehicles owned per capita (number of vehicles)	0.53	0.57	0.42	0.02	1.53	0.79	−0.29
Road length per capita (km)	1.18	0.91	0.80	0.15	3.90	0.68	1.01
Population Centrality (no unit)	0.78	0.63	0.16	0.33	1.29	0.21	2.81
Road Density (km per km^2)	2295.72	2110.26	1142.60	383.41	8498.26	0.50	0.49
Population Density (per km^2)	3085.15	2356.91	3009.20	143.34	21,669.89	0.98	0.73
GDP per capita (dollars)	37,704.80	36,157.00	19,496.41	2500	102,941	0.52	0.24
Fuel economy (miles per gallon)	29.67	24.33	4.81	22.23	39.46	0.16	3.33
Area (km^2)	1313.86	859.67	1439.38	21.02	6755.83	1.10	0.95

* The coefficient of skewness is calculated following Pearson's second skewness coefficient [44].

Statistical summary results indicated that:

(1) These 58 cities cover a wide range of socio-economic/urban form characteristics.
(2) Except for the area, all other variables have a coefficient of variation less than 1, which suggests that, compared to a standard normal distribution, these variables are more concentrated toward their means.
(3) Except for vehicles owned per capita, all other variables have a positive coefficient of skewness (a right-skewed distribution). This indicates that there will be a high frequency of small values and low frequency of large values for these variables.
(4) Population centrality and fuel economy are the most and second most right-skewed variables. This indicates that extreme values or outliers are more likely on the right side of the distribution. Combining the information provided by the skewness and the variation indicates that caution should be exercised when interpreting regression results for cities that are less centralized and more fuel-efficient.

We first performed a series of single linear regressions between each of the variables in Table 3 to briefly examine their correlations as well as exclude explanatory variables with potential multicollinearities. These single linear regressions were log-transformed, so that the slope coefficients could be interpreted as an elasticity, which describes the percentage change of one variable in response to the percentage change of another variable.

A series of regressions were then conducted by running an ordinary least square (OLS) regression on every possible combination of the independent variables. For each combination, the following items were calculated:

- An adjusted R^2 of the regression equation;
- A *p*-value for each independent variable;
- A variance inflation factor (VIF) for the coefficient of each independent variable. VIF is calculated as $1/(1 - R_i^2)$, where R_i^2 is the coefficient of determination of a constructed regression equation with the corresponding independent variable *i* as a function of all the other independent variables. It is a factor that measures how much the variance of an estimated regression coefficient is increased because of collinearity with other independent variables;
- The Semi-partial R-squared. This is calculated as $R^2 - R_i^2$, where R^2 is the coefficient of determination of the regression equation with the combination of all independent variables, and R_i^2 is the coefficient of determination of the regression equation with the combination of all independent variables except variable i. This semi-partial R-squared value describes the amount of on-road $FFCO_2$ emission variation explained by the variation in the independent variable *i* while excluding the impacts from other independent variables (eliminating collinearity).

The final OLS regression model was chosen based on the following criteria:

- All independent variables in the OLS regression model are statistically significant with a *p*-value less than 0.05;
- All VIF values are less than 10 (according to [45], a VIF larger than 10 could be considered as having a high collinearity);
- While satisfying the previous two criteria, the regression model with the highest adjusted R^2 was chosen.

This study conducted a 10-fold cross validation to assess the model's sensitivity. The 58 C40/MCD cities were split into 10 subsets. Cities were randomly chosen for each subset. For one round of cross-validation, one subset was held out, and the model was trained based on the rest of the subsets. The trained model was then used to predict values for the subset that was removed. The predicted values were then compared with the actual values for the removed subset to calculate errors. This procedure was repeated until all 10 subsets had been used as the validation set, yielding 10 sets of coefficients and errors. The whole cross-validation process was, again, repeated 10 times to obtain a distribution of 100 coefficients (10 re-sampling rounds repeated 10 times).

Besides data-resampling, another attempt to assess the model sensitivity is by examining model specifications. Neumayer and Plümper [46] believed that uncertainty in social science data analysis is caused more by improper model specifications than random sampling errors. They argued that: " ... inferences become more valid if estimated results are sufficiently independent from the model specification, that is, if all plausible alternative specifications give similar results".

This study included seven socio-economic/urban form factors as independent variables in a multivariate regression. All possible combinations of these seven factors were regressed against per capita on-road $FFCO_2$ emissions. Thus, the number of different model specifications that contained a specific independent variable was $2^7 - 1 = 127$ (the NULL model without any explanatory variables was not considered).

A further sensitivity assessment was conducted by adding the errors and observing the variation of the coefficients and the significance for each independent variable. An artificially introduced error was generated as a random sample from a normal distribution (assuming that, in real world, datasets have uncertainties that follow normal distribution) with mean = 1 and standard deviation = σ. This error was then applied to per capita on-road $FFCO_2$ emissions and the socio-economic/urban form factors to simulate their uncertainties. The new dataset with the artificial error was then used to train the regression model. The whole process was repeated 100 times and yielded 100 sets of coefficients for all independent variables.

3. Results

3.1. Single Regression

Table 5 summarizes the single linear regression results among all variables in log-log form. The relationship between population density, *PD*, and on-road $FFCO_2$ emissions, *E*, exhibits a negative elasticity (slope = −0.26; $p < 0.01$; $R^2 = 0.13$). This is consistent with the previous literature, where increases in population density are correlated with declines in per capita on-road $FFCO_2$ emissions [9,47]. Previous research also found that *PD* had a significant relationship to vehicle ownership and vehicle use, while almost no relationship to fuel economy [25,26]. The single regression results here are consistent with those findings with an insignificant relationship ($p < 0.21$) between *PD* and fuel economy (miles per gallon - *MPG*), and a significant relationship ($p < 0.01$) between *PD* and vehicles owned per capita, *VOPC*. *PD* is positively correlated to road density, *RD* (slope = 0.42; $p < 0.001$; $R^2 = 0.44$). However, the elasticity is smaller than 1, meaning that for every 1% increase in *PD* there is a less than 1% increase in *RD*. There is a weak but significant relationship between *PD* and population centrality, *PC* (slope = 0.08; $p < 0.01$; $R^2 = 0.11$). This suggests that for larger values of *PD*, the population distribution is more centralized. However, the strength of the elasticity should not be over-interpreted because the correlation is weak. *PD* has a negative relationship to *VOPC* (slope = −0.36, $p < 0.01$, $R^2 = 0.12$). This supports the hypothesis of Karathodorou et al. [25] that urban density discourages people from purchasing vehicles.

Table 5. Results from single-variable linear regression models.

Independent \ Dependent	$FFCO_3$ Emissions per Capita	Population Density	Road Density	Miles per Gallon	Population Centrality	GDP per Capita	Vehicles Owned per Capita	Road Length per Capita
$FFCO_2$ emissions per capita	1 (1)	−0.26 (0.13) [0.01]	0.09 (0.01) [0.58]	−3.35 (0.4) [0]	0.68 (0.05) [0.09]	0.79 (0.45) [0]	0.46 (0.42) [0]	0.54 (0.29) [0]
Population density	−0.47 (0.13) [0.01]	1 (1)	1.04 (0.44) [0]	1.18 (0.03) [0.21]	1.34 (0.11) [0.01]	−0.37 (0.06) [0.07]	−0.33 (0.12) [0.01]	−1.03 (0.6) [0]
Road density	0.06 (0.01) [0.58]	0.42 (0.44) [0]	1 (1)	0.37 (0.01) [0.54]	0.73 (0.08) [0.03]	0.2 (0.04) [0.13]	0.05 (0.01) [0.55]	−0.03 (0) [0.78]
Miles per Gallon	−0.12 (0.4) [0]	0.02 (0.03) [0.21]	0.02 (0.01) [0.54]	1 (1)	−0.2 (0.12) [0.01]	−0.08 (0.12) [0.01]	−0.04 (0.07) [0.04]	−0.03 (0.02) [0.25]
Population Centrality	0.08 (0.05) [0.09]	0.08 (0.11) [0.01]	0.11 (0.08) [0.03]	−0.6 (0.12) [0.01]	1 (1)	0.11 (0.07) [0.04]	0.08 (0.11) [0.01]	−0.07 (0.04) [0.13]
GDP per Capita	0.57 (0.45) [0]	−0.15 (0.06) [0.07]	0.2 (0.04) [0.13]	−1.57 (0.12) [0.01]	0.7 (0.07) [0.04]	1 (1)	0.5 (0.68) [0]	0.41 (0.24) [0]
Vehicles Owned per Capita	0.91 (0.42) [0]	−0.36 (0.12) [0.01]	0.13 (0.01) [0.55]	−2.02 (0.07) [0.04]	1.38 (0.11) [0.01]	0.79 (0.68) [0]	1 (1)	0.75 (0.28) [0]
Road Length per Capita	0.54 (0.29) [0]	−0.58 (0.6) [0]	−0.04 (0) [0.78]	−0.81 (0.02) [0.25]	−0.61 (0.04) [0.13]	0.57 (0.24) [0]	0.38 (0.28) [0]	1 (1)

For each pairwise regression, the slope coefficient, the R^2 (in parentheses) and the *p*-values (in brackets) are provided. Statistically significant values ($p \leq 0.05$) are in red; insignificant values are in blue.

It is expected that *MPG* will be negatively related to *E*, and the single regression results indicate that it has the highest elasticity (−3.35) among all the variables. As such, *MPG* alone can explain 40% of the variation in per capita on-road $FFCO_2$ emissions. The relationship between *MPG* and *PD* is weak ($R^2 = 0.03$). Karathodorou et al. [25] arrived at the same conclusion. Their explanation was that population density has both positive and negative impacts on fuel economy. For example, high urban density limits parking space and encourages people to purchase smaller, potentially more fuel-efficient cars. However, high density can also lead to congestion, inducing less efficient "stop and go" driving.

PC has a positive relationship to *E*, which suggests that decentralized cities tend to have greater per capita emissions. This would support the theory that a sprawling urban form increases on-road per capita $FFCO_2$ emissions [14,48]. However, the correlation is not very significant (*p*-value = 0.09).

Both per capita GDP and VOPC have positive, large and significant relationships to *E*. However, per capita GDP and VOPC are highly correlated (R^2 = 0.68). This indicates that it is not proper to include both as explanatory variables in the multiple regression model due to multi-collinearity.

3.2. Multivariate Regression

Table 6 summarizes the statistically significant variables that were included in the multivariate regression model construction (adjusted R^2 = 0.68). A combination of fuel economy (represented by miles per gallon, *MPG*), vehicles owned per capita, road density and population density remain the best predictors of total per capita on-road $FFCO_2$ emissions in the sample of cities examined. Figure 1 illustrates the corresponding correlation result from the single regression and multiple regression for these predictors. The circles in Figure 1 represent the variations of the dependent variable (DV) and the independent variables (IV). The top row of Figure 1 illustrates the relationships between the DV and IV in a single regression model (while ignoring the impacts of other IVs); the bottom row illustrates the relationships between the same combination of DV and IV in the multiple regression model (while controlling the rest of the IVs).

Table 6. Results of the Multiple Linear Regression Model. Adjusted R^2 = 0.68.

Variable	Coefficient	*P*-value	Semi-partial R^2
Fuel economy (MPG)	−2.67	3.10×10^{-8}	0.23
Vehicles Owned per Capita (VOPC)	0.25	3.05×10^{-4}	0.08
Road Density (RD)	0.39	4.41×10^{-3}	0.04
Population Density (PD)	−0.27	3.05×10^{-3}	0.05

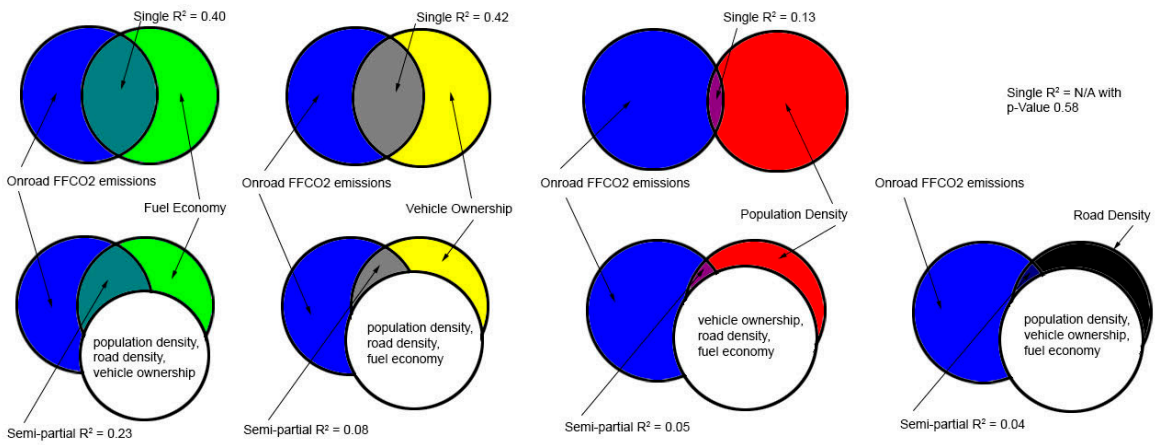

Figure 1. Correlations between on-road $FFCO_2$ emissions and fuel economy, vehicle ownership, population density and road density in single and multiple regression models.

The multivariate regression results indicate that the elasticity (−2.67) of *MPG* to per capita on-road $FFCO_2$ emissions (*E*) was the largest among all explanatory variables. Moreover, *MPG* is least influenced when comparing its relationship within the single regression (R^2 = 0.4 when ignoring other factors) versus the multiple regression (semi-partial R^2 = 0.23 when controlling other factors) models. This suggests that *MPG* is relatively independent of other explanatory variables. Sensitivity analysis results also indicated that *MPG* has the most robust relationship to on-road $FFCO_2$ emissions.

Both the elasticity and explanatory power of *VOPC* decreased in the multiple regression model (elasticity = 0.25, semi-partial R^2 = 0.08) when compared to the single regression model (elasticity = 0.91, R^2 = 0.42). Although the GDP per capita (*GDPPC*) was assumed to represent the affluence factor in the

IPAT model, it was excluded in the multivariate regression model in favor of *VOPC*. This was due to the fact that there is a high correlation between *GDPPC* and *VOPC* ($R^2 = 0.68$; $p < 0.001$; see Section 3.1). *VOPC* was selected not only because it yielded a mathematically better regression model with smaller residuals, but also because it is based on the physical relationship between VOPC and on-road $FFCO_2$ emissions. *GDPPC* does not directly and immediately impact on-road transportation $FFCO_2$ emissions but has a gradual rather than immediate influence on the vehicle market.

The elasticity of *RD* to on-road $FFCO_2$ emissions in the multiple regression model was estimated as 0.39, which was within the range of previously reported elasticities [20,22,47,49]. However, *RD* had an insignificant relationship to emissions in the single regression model (elasticity = 0.09, $R^2 = 0.01$; $p < 0.58$). Further examination of the correlation matrix revealed that *RD* was independent of all other socio-economic/urban form factors, except for *PD* and *PC*. This suggests that the proportion of variation of *E* caused by *RD* was masked by the correlation between *E* and *PD*. This has indicated that *RD* alone does not have a significant relationship to *E*, but when holding *PD* constant, *RD* became significant with a larger elasticity (elasticity = 0.39, $R^2 = 0.04$; $p < 0.001$).

The elasticity of *PD* to on-road $FFCO_2$ emissions (-0.27) in the multivariate results was within the reported range from previous studies [9,23–26]. These previous studies, mostly individual city case studies, reported a negative relationship between *PD* and on-road $FFCO_2$ emissions with weak to insignificant explanatory power. By contrast, we report here a significant relationship (p-value < 0.003). However, the relatively low R^2 (0.13 in single regression and 0.05 in multiple regression) suggests that *PD* would be a poor predictor of future emissions.

Figure 2 summarizes the regression residuals. These residuals were examined for patterns that would violate OLS assumptions. The quantile-quantile (QQ) plot between the residual and standard distribution suggests that residuals are normally distributed. The residual histogram has −0.14 skewness and −0.42 kurtosis, which suggests that the distribution of residuals is slightly left-skewed and slightly lighter-tailed than a normal distribution. The skewness and kurtosis of residuals yields a p-value of 0.99 for the Jarque–Bera normality test [50], which indicates that there is not a statistically significant difference between the residual distribution and normal distribution. All of these measures indicate that the log-transformed multivariate regression model is appropriate for this study.

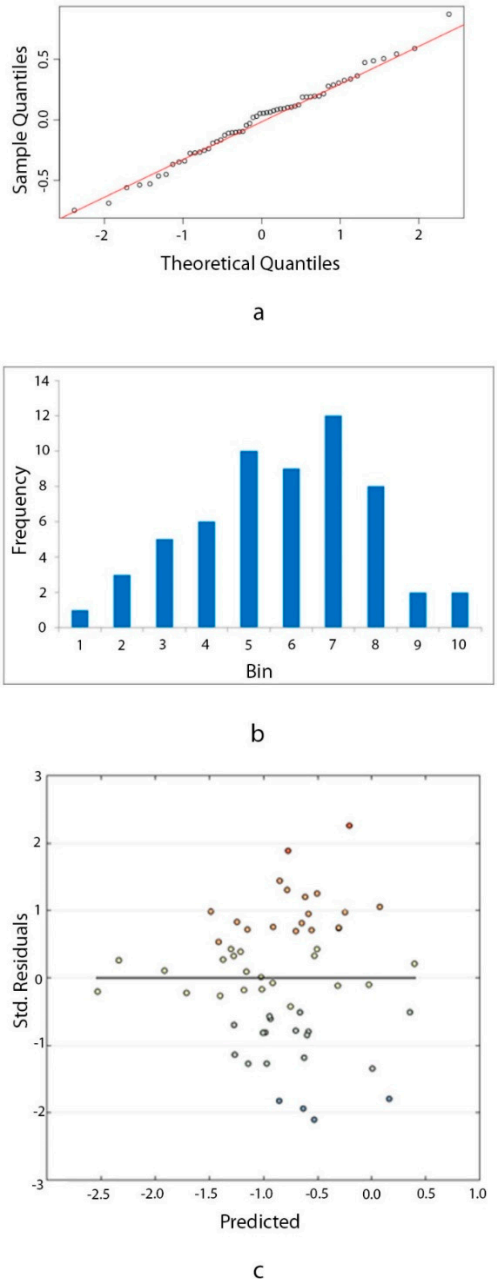

Figure 2. Summary of the regression residuals from the multivariate Stochastic Impacts by Regression on Population, Affluence and Technology (STIRPAT) regression. (**a**) QQ plot for regression residuals; (**b**) Histogram for regression residuals; (**c**) Scatterplot between standardized residuals and predicted values.

3.3. Sensitivity Analysis Result

This study used the coefficient of variance (CV), which is defined as the ratio of standard deviation over the mean, to measure distribution dispersion. There is no objective standard for assessing the CV value. However, CV can be used to compare the relative variability between variables with different units (given that all variables are measured on a ratio scale).

Statistics of coefficients based on data-resampling (Table 7) showed that all coefficients retained the same sign after 100 rounds of cross-validation. The coefficients from the original model were also within the 95% confidence interval for the mean of these 100 coefficient sets. This suggests that all coefficients have a robust relationship to per capita on-road $FFCO_2$ emissions [51]. Among

the regression coefficients, *MPG* had the smallest CV value. This suggested that compared to other independent variables, *MPG* is the least sensitive to data re-sampling. Results also showed that *MPG* and *VOPC* remained 100% significant in all rounds of data resampling compared to 70% for *PD* and *RD*. This indicated that *MPG* and *VOPC* are relatively more robust to data resampling compared to *RD* and *PD*.

Table 7. Statistical summary of regression coefficients after cross-validation.

Statistical Summary of Regression Coefficients	Miles per Gallon	Vehicles Owned per Capita	Population Density	Road Density
Minimum	−3.16	0.20	−0.32	0.31
Maximum	−2.29	0.31	−0.21	0.48
Median	−2.67	0.25	−0.27	0.39
Mean	−2.69	0.25	−0.27	0.39
Standard deviation	0.15	0.02	0.02	0.03
Coefficients from original model	−2.67	0.25	−0.27	0.39
95% confidence interval	[−2.72,−2.67]	[0.24,0.25]	[−0.27,−0.26]	[0.38,0.40]
Coefficient of variance	−0.05	0.07	−0.08	0.07
% of significant *p*-values in all rounds	100%	100%	66%	63%

Statistics of coefficients based on different model specifications (Table 8) showed that the final coefficient values for all four variables in the original model were outside of the 95% confidence interval for the mean of 127 model specifications. Only the *MPG* coefficient retained the same sign in all 127 models. This suggested that *MPG* had a more robust relationship with per capita on-road $FFCO_2$ emissions compared to vehicles owned per capita, population density and road density. The relatively smaller CV value also indicated that the *MPG* coefficient varied less about its mean than the other variables in the model. All of the *MPG* coefficients remained statistically significant in all models, while *VOPC*, *PD* and *RD* had 50%, 45% and 25% significant coefficients, respectively. This again indicated that *MPG* was relatively more robust by having relatively more significant relationships in different model specifications.

Table 8. Statistical summary of regression coefficients for all possible model specifications.

Statistical Summary of Regression Coefficients	Miles per Gallon	Vehicles Owned per Capita	Population Density	Road Density
Minimum	−3.49	−0.13	−0.64	−0.09
Maximum	−1.89	0.46	0.21	0.76
Median	−2.47	0.19	−0.13	0.12
Mean	−2.49	0.19	−0.13	0.17
Standard deviation	0.293	0.13	0.195	0.17
Coefficients from original model	−2.67	0.25	−0.27	0.39
95% confidence interval	[−2.53,−2.45]	[0.17,0.21]	[−0.15,−0.11]	[0.15,0.19]
Coefficient of Variance	−0.11	0.71	−1.46	1.02
% of significant *p*-values	100.00%	50.78%	45.31%	25.00%

The statistics of the regression coefficients subjected to artificially introduced errors (Table 9) showed that all variables retained the same sign in all repetitions at all error levels. When the standard deviation (σ) of the artificial error was at 0.05, all four independent variables remained 100% significant. With σ increasing, the coefficient of variation range increased for all independent variables. Regression coefficients for the vehicles owned per capita varied the least by having the smallest CV value at all error levels. In addition, VOPC was also the only variable that remained 100% significant in all repetitions when the standard deviation of artificial errors had increased to $\sigma = 0.3$. *MPG* and population density remained 100% significant until σ was raised above 0.1. Road density remained 100% significant only when σ was at or below 0.05.

Table 9. Statistical summary of regression coefficients with artificially introduced errors.

Statistical Summary of Regression Coefficients	Miles per Gallon	Vehicles Owned per Capita	Population Density	Road Density
Median	−2.63 (σ = 0.05) −2.56 (σ = 0.1) −1.82 (σ = 0.3) −1.03 (σ = 0.5)	0.25 (σ = 0.05) 0.25 (σ = 0.1) 0.25 (σ = 0.3) 0.24 (σ = 0.5)	−0.27 (σ = 0.05) −0.27 (σ = 0.1) −0.23 (σ = 0.3) −0.20 (σ = 0.5)	0.39 (σ = 0.05) 0.38 (σ = 0.1) 0.32 (σ = 0.3) 0.25 (σ = 0.5)
Mean	−2.59 (σ = 0.05) −2.46 (σ = 0.1) −1.83 (σ = 0.3) −1.16 (σ = 0.5)	0.25 (σ = 0.05) 0.25 (σ = 0.1) 0.25 (σ = 0.3) 0.25 (σ = 0.5)	−0.27 (σ = 0.05) −0.27 (σ = 0.1) −0.21 (σ = 0.3) −0.17 (σ = 0.5)	0.39 (σ = 0.05) 0.37 (σ = 0.1) 0.32 (σ = 0.3) 0.23 (σ = 0.5)
Standard deviation	0.10 (σ = 0.05) 0.29 (σ = 0.1) 0.71 (σ = 0.3) 0.95 (σ = 0.5)	0.00 (σ = 0.05) 0.01 (σ = 0.1) 0.03 (σ = 0.3) 0.06 (σ = 0.5)	0.01 (σ = 0.05) 0.02 (σ = 0.1) 0.07 (σ = 0.3) 0.14 (σ = 0.5)	0.01 (σ = 0.05) 0.03 (σ = 0.1) 0.12 (σ = 0.3) 0.21 (σ = 0.5)
Coefficients from original model	−2.67	0.25	−0.27	0.39
Coefficient of Variance	−0.04 (σ = 0.05) −0.12 (σ = 0.1) −0.39 (σ = 0.3) −0.82 (σ = 0.5)	0.02 (σ = 0.05) 0.04 (σ = 0.1) 0.14 (σ = 0.3) 0.24 (σ = 0.5)	−0.02 (σ = 0.05) −0.06 (σ = 0.1) −0.35 (σ = 0.3) −0.81 (σ = 0.5)	0.03 (σ = 0.05) 0.09 (σ = 0.1) 0.38 (σ = 0.3) 0.92 (σ = 0.5)
% of significant *p*-values	100% (σ = 0.05) 100% (σ = 0.1) 91% (σ = 0.3) 68% (σ = 0.5)	100% (σ = 0.05) 100% (σ = 0.1) 100% (σ = 0.3) 95% (σ = 0.5)	100% (σ = 0.05) 100% (σ = 0.1) 80% (σ = 0.3) 66% (σ = 0.5)	100% (σ = 0.05) 98% (σ = 0.1) 77% (σ = 0.3) 62% (σ = 0.5)

4. Discussion

Given the explanatory power, independence and robustness of the fuel economy within the regression results presented here, improving the fuel economy offers an effective and reliable tool for policymakers to mitigate $FFCO_2$ emissions. However, the fuel economy has not shown a significant improvement across the world. In Australia, the Survey of Motor Vehicle Usage (SMVU) conducted by the Australian Bureau of Statistics indicated that the fuel economy (12 L per 100 km) has not changed much between 1976 and 2018 [52]. In Britain, fuel economy improvement has significantly slowed since the mid-1980s [53]. Fuel economy improvements were potentially offset by the increased vehicle weight [53] or the increased vehicle use [54]. In the United States, the fuel economy has been stagnant for the past 30 years [7]. Vehicle technology developments were mostly focused on improving vehicle performance, and especially the engine size and maximum speed, rather than fuel efficiency. Furthermore, vehicles in the US have become 20% heavier and 25% faster, while the fuel economy has remained constant since 1985 [7]. The Global Fuel Economy Initiative (GFEI) report for the International Energy Agency [42] found that the global fuel economy increased 1.7% annually between 2005 and 2011. This indicated that the global fuel economy improvement would very likely lag behind the GFEI target of halving new vehicle fuel economy by 2030. On the other hand, the global production and purchase of hybrid/electric vehicles are still limited due to a lack of charging infrastructure [55] and insufficient mileage [56].

The decline of the elasticity and explanatory power of *VOPC* in the multivariate results versus the single regression model indicates that the impact of *VOPC* on $FFCO_2$ emissions is subject to the influence of other socio-economic/urban form factors. Similar conclusions were drawn by Karathodorou et al. [25], who regressed the vehicle ownership against a series of factors and found significant relationships with population density, GDP, road length per capita, public transit per capita and external costs (such as maintenance, taxes, parking, tolls and depreciation). The *VOPC* elasticity (0.25) derived from the multiple regression model in this study is close to the elasticity (0.338) calculated by Jones and

Kamman [57], who regressed the household carbon footprint against vehicle ownership controlling for income, carbon intensity of electricity, home size, persons per household and population density.

The four variables identified in this analysis as significant and robust have potential policy interventions, assuming an interest in mitigating GHG emissions. However, the availability of policy options is likely a reflection of the interactions between local and provincial or national governments, not to mention the general nature of government purview and structure. For example, the fuel economy has typically been established at the national/federal level, given the inefficiency and challenges associated with attempting to regulate the fuel economy in local markets. The control over the number of vehicles owned per capita or vehicle ownership in general is rare and likely difficult to enact. Road density and population density, however, are attributes often considered part of the urban planning effort, though not necessarily in relation to $FFCO_2$ emissions. Given the fact that these two urban form attributes are significant variables in explaining on-road $FFCO_2$ emissions, they may take on added meaning when considering urban form in GHG mitigation planning. As with much consideration of GHG mitigation at the urban scale, these on-road $FFCO_2$ emission drivers are one factor in a number of competing interests and needs of the planning processing cities.

The quantification of on-road $FFCO_2$ emissions at a high spatiotemporal resolution adds useful information content to urban GHG inventories. Traditional "bottom-up" approaches require extensive local traffic activity data, such as traffic counts, VMT, fleet composition and fuel types. However, collecting, assimilating and analyzing these data are usually labor-intensive and costly. Although more cities have begun to participate in measuring and reporting their local GHG inventories [58], detailed local transportation data are still not universally available. This has limited most bottom-up on-road $FFCO_2$ inventories to local efforts and case studies within the developed cities.

This study explored the relationships between on-road $FFCO_2$ emissions and socio-economic/urban form factors from a global perspective. All input data used in this study are globally available from either existing public databases or remote sensing data products. The findings of this study might assist in generating on-road $FFCO_2$ emission estimates in developing cities with insufficient local transportation data. Given the importance of on-road $FFCO_2$ emissions in cities and the potential for local government influence over aspects of on-road travel, this study demonstrates a more amenable approach to establishing baseline on-road emissions and setting mitigation targets for areas where the traditional bottom-up approach is unsuitable.

5. Conclusions and Future Work

This study aims to enhance the understanding of driving factors for on-road $FFCO_2$ emissions for global cities. Beyond individual single-city case studies, the most common form of analysis to date, it offers general outcomes across a diverse sample of cities. Using a STIRPAT regression model approach, the results identified the fuel economy, vehicle ownership, population density and road density as having significant relationships with on-road $FFCO_2$ emissions. Among these factors, the fuel economy has the strongest, most independent and robust relationship with on-road $FFCO_2$ emissions. The relationship between vehicle ownership and on-road $FFCO_2$ emissions appears strong in a single regression model, but when controlling for other factors, declines in marginal explanatory power (semi-partial $R^2 = 0.08$). Population density and road density have significant but relatively weak relationships to on-road $FFCO_2$ emissions.

The outcome of the study may provide a potential method for scientists in the field of carbon studies to quantify carbon footprints in urban areas where alternative methods prove too costly or not possible due to local data constraints.

The characteristics of the relationships identified in this study can also provide valuable information support for policy making, especially for global city mayors. In order to achieve the goal of sustainability and mitigate on-road $FFCO_2$ emissions, an emphasis on improving the fuel economy and public transportation (to reduce vehicle ownership per capita) should still be the focus in the near future. On the other hand, although increasing population density and reducing road density still have a

positive impact on reducing on-road $FFCO_2$ emissions, priority should be given to the improvement of fuel economy and public transportation, since their relationships are relatively more significant and robust.

A number of caveats to the work presented here are worth mentioning. First, it takes time to observe the significant changes caused by socio-economic/urban form factors. For example, the impacts of GDP or income on fuel consumption are mainly reflected through the variation of vehicle ownership and distance traveled [12,13]. These effects could take years to become significant. Due to this lagged effect, long-term elasticity is usually larger than short-term elasticity. This study was based on cross-sectional data. Therefore, not all the estimated elasticities reflect long-term effects. This could lead to an underestimation of elasticities when compared with other studies that used panel data. In future studies, both the short-term and long-term elasticities should be considered and reported based on both spatial and temporal data.

Data quality is another major caveat for this study. Although a careful examination and prudent approximation were implemented during the data assimilation process, the wide range of data types, formats, vintages and spatial/temporal resolutions can still result in considerable output uncertainty. This study took the approach of addressing the uncertainty issue by conducting sensitivity analyses based on the artificially introduced errors. Given enough information regarding the data quality assessment, a more detailed uncertainty analysis could be conducted to improve the model precision.

Another limitation of this study is that it primarily focused on the explanatory variables that were included in the final multiple regression model. This was purely based on the mathematical reason that these explanatory variables yielded the least residuals and most significances without multi-collinearity. This does not mean that other socio-economic factors were not important. Future work should focus more on other factors that had strong and significant relationships with per capita on-road $FFCO_2$ emissions but were excluded in this study due to multi-collinearity, such as GDP per capita and road length per capita.

Author Contributions: Conceptualization, Y.S. and K.R.G.; methodology, Y.S. and K.R.G.; software, Y.S.; validation, Y.S.; formal analysis, Y.S.; investigation, Y.S.; resources, Y.S.; data curation, Y.S.; writing—original draft preparation, Y.S.; writing—review and editing, Y.S. and K.R.G.; visualization, Y.S.; supervision, K.R.G.; project administration, K.R.G.; funding acquisition, K.R.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Aeronautics and Space Administration (NASA), CMS grant NNX12AP52G and National Science Foundation (NSF) CAREER award 0846358." and "The APC was funded by Kevin Gurney.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

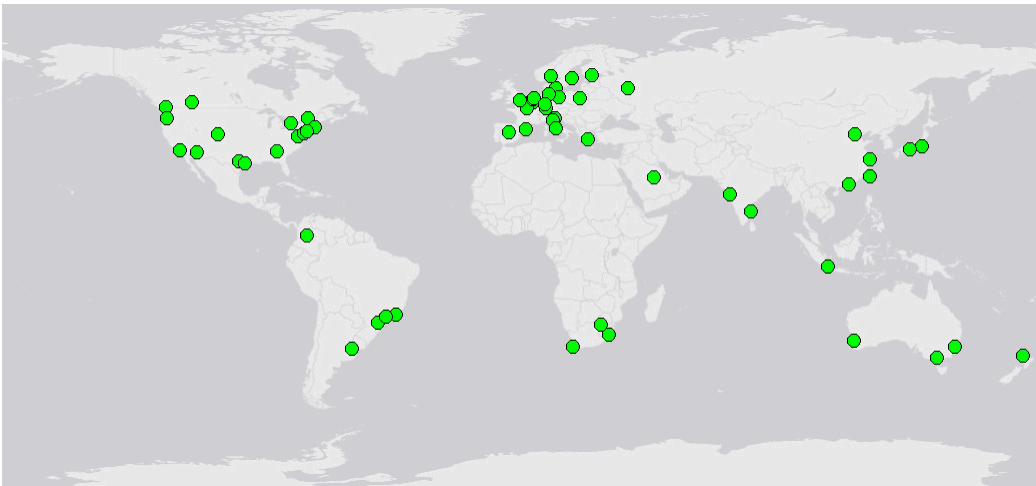

Figure A1. The geographical locations of 58 global cities used in this study.

Table A1. The 58 global cities used in this study.

City	Country	City	Country	City	Country
Buenos Aries	Argentina	Athens	Greece	Durban	South Africa
Melbourne	Australia	Chennai	India	Johannesburg	South Africa
Sydney	Australia	Mumbai	India	Madrid	Spain
Perth	Australia	Jakarta	Indonesia	Barcelona	Spain
Brussels	Belgium	Venice	Italy	Stockholm	Sweden
Curitiba	Brazil	Bologna	Italy	Taipei	China Taiwan
Rio de Janeiro	Brazil	Rome	Italy	London	United Kingdom
Sao Paulo	Brazil	Tokyo	Japan	Atlanta	United States
Vancouver	Canada	Osaka	Japan	Portland	United States
Calgary	Canada	Rotterdam	Netherlands	Denver	United States
Toronto	Canada	Amsterdam	Netherlands	Boston	United States
Montreal	Canada	Auckland	New Zealand	Washington DC	United States
Bogota	Colombia	Oslo	Norway	Austin	United States
Copenhagen	Denmark	Warsaw	Poland	Phoenix	United States
Helsinki	Finland	Hong Kong	China	Philadelphia	United States
Paris	France	Beijing	China	Houston	United States
Stuttgart	Germany	Shanghai	China	Los Angeles	United States
Hamburg	Germany	Moscow	Russia	New York	United States
Frankfurt	Germany	Riyadh	Saudi Arabia		
Berlin	Germany	Cape Town	South Africa		

Appendix A.1. Data Sources for Urban Form Factors

Appendix A.1.1. Population Density

Population density is defined in this study as population divided by city area in square kilometers. Population density is calculated based on LandScan [38] which is a global population dataset developed by Oak Ridge National Laboratory (ORNL) using spatial data and imagery analysis. It disaggregates sub-national census counts to a 1 km^2 grid based on multi-variable spatial modeling. Many datasets are included in the modeling process, such as administrative boundaries, land cover, elevation, slope, coastlines and high-resolution imagery. Each grid cell is assigned a likelihood factor based on these data to predict the possible occurrence of population. The census result is then further disaggregated based on each cell's likelihood factor. Transportation on-road emissions are more likely associated with people's total movements during a day rather than where people sleep. Therefore, compared to traditional household survey statistics, LandScan is better suited for this proposed study due to the fact that it represents the ambient population (average population within 24 h).

For verification and validation purposes, Dobson et al. [34] utilized the LandScan data based on state population with a 30 s by 30 s grid for the southwestern United States. The gridded results were aggregated to counties and compared to the county census data (which was a local census in comparison with the sub-national census for LandScan). Results showed that 87.8% of simulated population corresponded to the census data (meaning 12.2% of the population was placed in an incorrect county). Dobson et al. [34] concluded that LandScan was the most suitable global database to estimate population compared to previous population databases.

Appendix A.1.2. Road Network Characteristics

Road Density was estimated using city area and a road network basemap, OpenStreetMap (OSM). Per capita road length was estimated using the OSM road network and urban population derived from LandScan. OSM is a crowd-source project under the Open Database License where individuals collaborate and contribute road-related geospatial data worldwide [39]. OSM data sources include aerial imagery, GPS, and field maps based on local knowledge. These data sources are integrated into geospatial databases and updated regularly by the public OSM community (consists of worldwide GIS professionals, cartographers, and engineers).

Various quality assessments of OSM have been conducted worldwide [35,36,59]. Haklay [35] compared OSM in England versus products from the Ordinance Survey (British government mapping agency). Results showed that OSM was, on average, within 6 m of the ground truth and covered over 80% of motorways. Girres and Touya [36] extended Haklay's [35] work to France by comparing OSM to an institutionalized reference map from the French National Mapping Agency. Girres and Touya [36] expressed conclusions similar to those of Haklay [35], specifically that OSM has fairly good locational accuracy (0–6 m away from ground truth on average) but the attribute accuracy was rather poor. For example, only 43% of roads were named, and 49% of residential and local roads were misclassified. Highways and primary roads on the other hand, achieved almost 100% correct classification. Forghani and Delavar [59] compared OSM to the official reference map published by the Municipality of Tehran separated in grids. They concluded that about 80% of Tehran was covered with fairly good quality OSM data. However, the spatial distribution of uncertainty varies by region. This may be, again, due to the heterogeneous coverage and lack of standards in the submitted geospatial data format.

Many of these assessments have concluded that OSM is a good representation of roads, yet all share concern about the data quality in specific regions. Due to the fact that OSM is "volunteered geographical information (VGI)" [60], its data quality varies by location and topic. Developed, populated, and urbanized places receive more data with good quality, while rural areas in general are less represented with relatively poorer quality [35].

This study focuses on major global cities and primary roads, which are represented with high accuracy in OSM (0–6 m away from ground truth on average) according to Girres and Touya [36]. Plus, the spatial calculation approach used in this study requires spatial accuracy rather than attribute accuracy. Therefore, OSM is considered well-suited for this research.

Appendix A.1.3. Population Centrality

There were many definitions of population centrality. The Wheaton index [61] measures how fast populations accumulate along the route between local activity centers and the city edge. The Wheaton index can be described by the formula:

$$Wheaton\ Index = \int_0^b \frac{f(t)}{b} dt \tag{A1}$$

where t is the distance to the local CBD, is the cumulative proportion of population at distance t; b is the distance from the CBD to city edge.

The Wheaton index can be interpreted as the area of cumulative distribution of population from 0 to t. When 100% of inhabitants are concentrated at the CBD, the Wheaton index equals 1. When the inhabitants are uniformly distributed, the Wheaton index equals 0.

Massey and Denton [62] proposed an area weighted index to describe population centrality. It measures the cumulative proportion of population compared to the cumulative proportion of land area around the CBD. This index can be described by the following formula in discrete form, where land areas in different zones are ordered from close to faraway based on their distance to CBD:

$$area\ weighted\ centrality\ index = \sum_{i=1}^{n} Pop_{i-1}Area_i - \sum_{i=1}^{n} Pop_i Area_{i-1} \quad (A2)$$

where is the cumulative proportion of population at zone i; is the cumulative proportion of land area through zone i.

The area weighted centrality index ranges from −1 to 1. It equals 0 when the population is uniformly distributed. A positive value indicates people are more concentrated around the CBD compared to a uniform distribution. Negative values indicate inhabitants are further away from CBD compared to a uniform distribution. Therefore, the area-weighted centrality index can be interpreted as the proportion of people needed to relocate in order to achieve a uniform distribution around CBD.

Bento et al. [47] used an index of population centrality which examines the percentage of population living within a percentage of distance from local activity centers to the edge of the city. For example, if city A has 80% of people living within the 20% of distance from local activity centers to the edge of the city, and city B has 20% of people living within 80% of that distance, city A will be considered more centralized than city B. This spatial index is independent of city and population size.

Estimation of population centrality requires information about the location and size of local activity centers. Various studies have been conducted to delineate activity centers according to their thermal, spectral, spatial and physical characteristics. Examples include identifying the ground surface temperature difference between an activity center and non-activity center areas [63]; detecting the shadows caused by high buildings [64]; comparing the building density with the aid of remote sensing images [65]; and mapping the socio-economic activities by the constraint of road network [66]. However, these are all case studies that use either high spatial/spectral resolution remote sensing images or local activity data, which are either expensive or challenging to adopt for global scale research. For the purpose of this study, there are no known comprehensive datasets that record the geographical locations of activity centers at the global scale.

This study proposes a fast and economic way to delineate activity centers based on impervious surface area (ISA) and road network data. A previous study by Wu and Murray [67] confirmed that, compared to other land use types, activity centers have the largest share of impervious surface and least share of vegetation coverage. Moreover, since activity centers are the center of socio-economic activities in an urban area, major roads converge to provide accessibility for suburban areas [68]. Combining the information of ISA and road density, a GIS hot-spot analysis was performed to identify the locations and sizes of clustered pixels with large ISA values and high road density (Figure A2).

As Taubenböck et al. [69] pointed out, the producer's accuracy for CBD classification (which measures the probability that ground features are accurately classified as such) cannot be calculated due to the lack of a comprehensive reference dataset or ground-truth. This study followed the approach of Taubenböck et al. [69] to calculate the user's accuracy (which measures the probability that map classifications are actually present on the ground) through visual inspection with map [70]. Results showed that out of 145 global cities, 106 had their identified activity centers within 3 km of local ground truths, yielding a user's accuracy of 73%. Figure A2 showed the locations of the largest activity center that was identified by the hot-spot spatial analysis for several global cities.

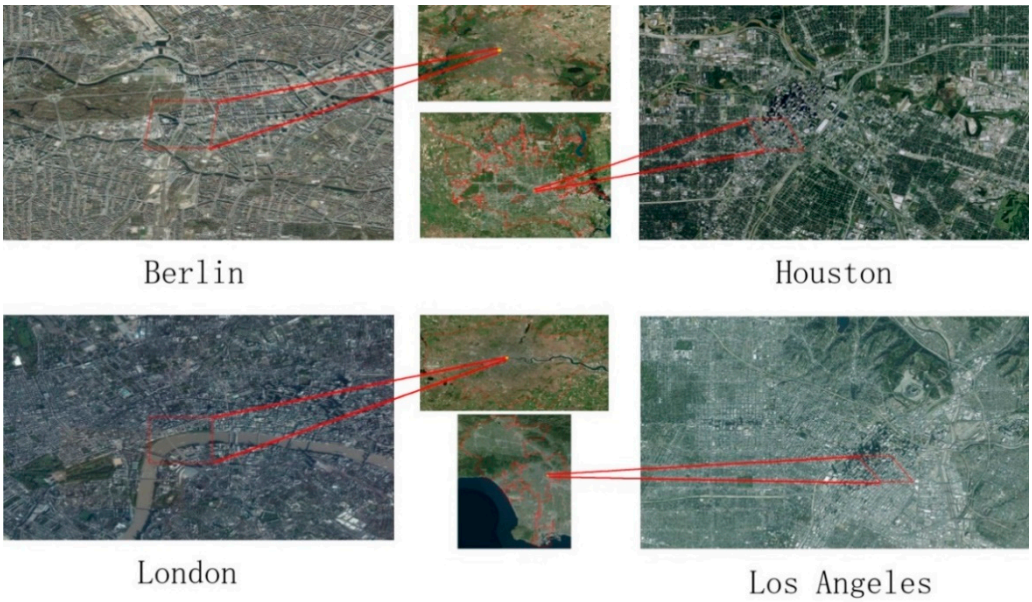

Figure A2. Four Global Activity centers identified by hot-spot analysis.

This study defines population centrality following the Bento et al.'s [47] definition, which is the ratio of the cumulate population at distances from activity centers to the average population at distances from activity centers. Suppose a city has m activity centers and is divided into n grid cells of equal area:

$$Population\ Centrality = \frac{\sum_{x=1}^{n}\sum_{y=1}^{m}\left(Pop_{xy} * Distance_{xy}\right)}{\sum_{x=1}^{n}\sum_{y=1}^{m}\overline{Pop} * Distance_{xy}} \tag{A3}$$

where Pop_{xy} is the proportion of population that was attracted by the yth activity center for the xth grid cell, $Distance_{xy}$ is the distance from the xth grid cell to the yth activity center, $\overline{Pop}$ is the average population for a single cell.

The calculation of Pop_{xy} uses the Landscan population data at 1 km^2 spatial resolution. Considering that cities may be polycentric with multiple activity centers, this study further divided Pop_{xy} into portions by the attractive power of each nearby activity center. The definition of attractive power follows the Newton's law of universal gravitation. For any two objects with mass m_1 and m_2, the attracting force between them is:

$$F = G\frac{m_1 m_2}{r^2} \tag{A4}$$

In this case, m_2 is represented by the intensity value for each activity center (size*ISA*road density). Assuming a city has n activity centers, each activity center will have its own attracting power F_y (k = 1,2, . . . n) to the local population. Thus, for the xth 1 km^2 grid cell with population P,

$$Pop_{xy} = p * \frac{F_y}{\sum_{y=1}^{n} F_y} \tag{A5}$$

After the Pop_{xy} and $Distance_{xy}$ have been quantified, the overall population centrality can be calculated. When the value of population centrality is small it indicates people are living in a clustered or compact pattern near the activity center. When this value is large it indicates people live far away from the activity center, or in a sprawl pattern.

Calgary, Canada and Naples, Italy are two examples that illustrate the difference of population centrality (Figure A3). Calgary is a relatively more centralized city with population centrality score as 0.403. Naples as a relatively decentralized city has population centrality score as 1.002. These two cities represent the upper and lower boundaries of population centrality for 145 global cities.

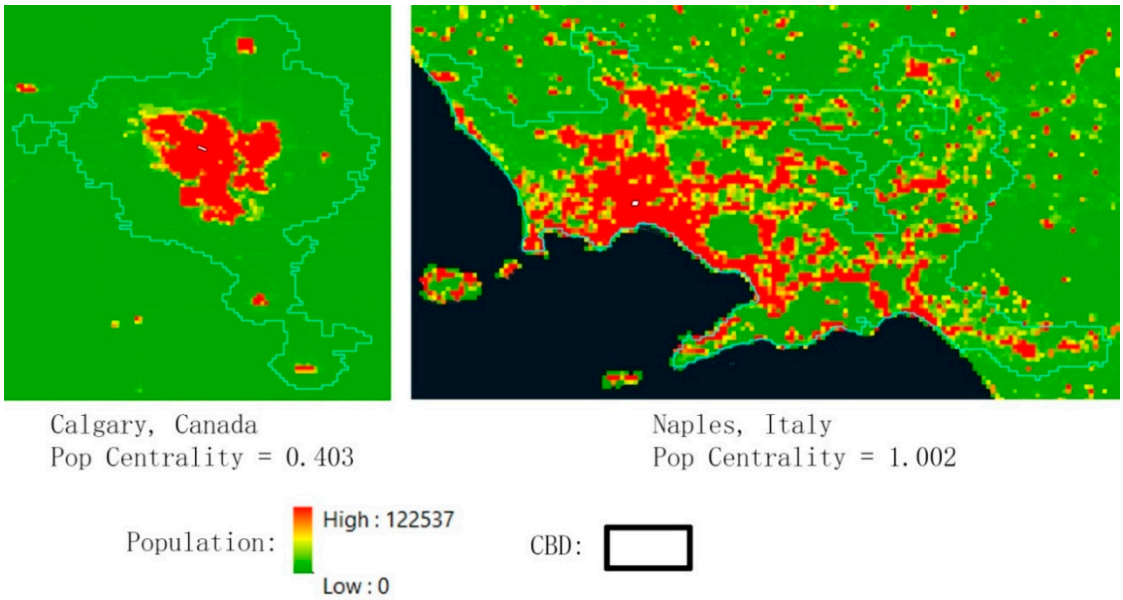

Figure A3. Population distribution in Calgary, Canada and Naples, Italy.

Appendix A.2. Data Sources for Socio-Economic Factors

Appendix A.2.1. Gross Domestic Product (GDP)

The urban GDP data were obtained from the Global Metro Monitor [41]. This publication is a product of the Brookings institution's metropolitan policy program (www.brookings.edu).

Global Metro Monitor has various data sources across the world. For U.S. metropolitan areas, the employment data come from the Local Area Unemployment Statistics (LAUS) program supervised by the U.S. Bureau of Labor Statistics [41]. The LAUS program provides monthly and annual-average employment estimates at the county and city scale based on the Current Population Survey, which is a monthly sample survey of approximately 60,000 households conducted by the US Census Bureau [71]. The population data is measured by U.S. Census Bureau's Population Estimates Program, which relies on decennial census data to provide population estimates [71].

For Europe, the economic, population and labor force data are provided by Cambridge Econometrics, which is a consulting agency in the United Kingdom. Their data are primarily based on the Eurostat [41].

Eurostat data are collected from national statistics following the Eurostat quality assessments [41]. This process is enforced by the European Statistical Law to guarantee the validity and consistency of the national reports.

For cities that are outside of United States and Europe, economic data are obtained from Oxford Economics, which is a consulting agency who provides data and analytical results for governments and Non-Government Organizations (NGO) at the global scale. According to Berube et al. [41], data sources for Oxford Economics come from various national statistical agencies or local data providers.

Per capita GDP for global cities were calculated based on these data by the Global Metro Monitor program. Complicated data sources make it very difficult to quantify the overall uncertainty. Although different agencies should use the best available data, uncertainties or errors will inevitably propagate and aggregate. Caution should be exercised when using and interpreting the GDP data for this study. This study used the per capita GDP data for the year 2011, which is the same year as vehicle ownership and fuel economy data (described next).

Appendix A.2.2. Vehicle Ownership

The Vehicles owned per capita (*VOPC*) data were obtained from World Road Statistics (https://worldroadstatistics.org/contact.html) [43]. This is a document published by the International Road Federation (IRF), a non-profit global organization with headquarters in Washington DC. Through their network, IRF sends out questionnaires to over 200 countries to collect statistical data related to road networks, traffic, and vehicles. The survey is conducted annually at the national scale as their primary data source. The secondary supplemental data source is the national year book from countries. Data are categorized by vehicle types and time periods, including passenger cars, lorries and vans, buses and motorcycles from 2007 to 2012. This study uses the light duty passenger vehicle counts for the year 2011, which is coincident with the GDP and fuel economy data (next section). According to IRF's documentation, the vehicle ownership data are validated by comparing them to other data sources (which vary by country and are not listed), comparing them with historical data, and checking/confirming outliers with contacts in their network. There hasn't been a quantitative quality assessment on this national vehicle ownership data.

Appendix A.2.3. Fuel Economy (MPG)

Cuenot et al. [42] published an updated global fuel economy database for the International Energy Agency (IEA). This follows the approach of the worldwide average fuel economy trends reported by the Global Fuel Economy Initiative (GFEI) started in 2005. The fuel economy data is derived from the average fleet composition and officially tested fuel economy for vehicle models sold from 2005 to 2011 in different countries.

The average fleet composition data for each country are based on the vehicle registration records provided by government statistical departments and worldwide third-party agencies (a detailed list of sources for each country can be found in the table Annex I quoted from [42]). The registration data were collected in CY 2005 and 2008 and were comprised of almost 50 million records in each year covering 94% and 88% of vehicles produced, respectively. Market sales data have been used to stratify these data. The market sales data are from 21 countries (for a detailed country list see Table 1 in [42]) covering 88% of the world total vehicle sales in 2008. The sales-weighted vehicle registration data are then summarized into a national fleet composition.

For vehicle models with various age, engine type and fuel type), the fuel economy is tested in a laboratory based on the result of fuel consumed per kilometer in different testing routines. The tested results and national fleet composition data are then compiled into fuel economy by vehicle model and country. This study uses the national average fuel economy across all models vehicle for each country at year 2011.

Cuenot and Fulton [42] acknowledged the caveats of these data such as the difference between the ideal fuel economy tested in lab and actual fuel economy under real-world conditions, not all data are from official government reports, and the testing methods may vary by country. No quantitative uncertainty analysis was done in their reports.

Appendix A.3. Data Sources for on-Road $FFCO_2$ Emission

This study collected data from 58 global cities with on-road $FFCO_2$ emission data available for year 2011. Half of the data are derived from $FFCO_2$ submissions to the C40 project [27], another half were included in the Millennium City Database [37].

Appendix A.3.1. C40 Cities

C40 is a network of the world's megacities committed to addressing climate change. The C40 cities are cities for which mayors have committed to reduce greenhouse gas emissions [27]. Annual city-wide greenhouse gas emissions are reported by C40 cities based on guidelines provided by the

Intergovernmental Panel on Climate Change (IPCC), Global Protocol for Community (GPC), and other local protocols for emission inventories.

The IPCC guidelines provide methodologies for estimating inventories of greenhouse gas emissions resulting from human activities [72]. For the transportation sector, two methods have been recommended by IPCC. "Top down" refers to an approach that uses fuel sales and an assumption regarding the carbon content of the fuel. "Bottom up" refers to an approach that uses VKT, an estimate of vehicle fuel economy, and the carbon content of fuel. Both methods have been used by Kennedy et al. [73]. They concluded that these methods resulted in similar outcomes. For example, Kennedy reported that for Bangkok's gasoline use in 2006, "the estimate of 2662 million liters based on VKT is 2.9% lower than the estimate of 2741 million liters from fuel sales" [73].

The GPC is a project developed by Local Governments for Sustainability (previously known as International Council for Local Environmental Initiatives, ICLEI), the World Resources Institute (WRI), and C40 cities. GPC advocates four data collection methods: (1) the fuel sales approach; (2) city-induced activity trips that either have their origin or destination in the city (passing through trips are excluded); (3) geographic boundary, includes all traffic occurring within city boundaries, regardless of origin or destination; and (4) resident activity, based on household survey of transportation activity. City induced activity is the methodology recommended by GPC.

The categorical difference in these methodologies and the potential for diverse outcomes is acknowledged. Chavez et al. [74] found that, in the city of Delhi, only 3% of VKT occurred across city boundaries. The authors theorized that "in megacities, VKT values attributed to trans-boundary traffic may be negligible due to the large amounts of concurrent in-boundary traffic". This implies that, at least for the city of Delhi and possibly for other megacities, the result of using the city-induced method may not be significantly different from using the geographic boundary method.

Of the 29 C40 cities, seven cited the IPCC guidelines as the primary methodological source for estimating transportation emission data, with two additional cities stating that they used "other" methodologies but were influenced by the IPCC. Eight cities identified the GPC as the primary methodological source with four additional cities stating that they were influenced by the GPC. Thirteen of the C40 cities stated that sources other than IPCC or GPC guided their transportation CO_2 emissions estimation. For example, Los Angeles obtained transportation data from the Southern California Association of Governments, the regional planning authority. Toronto used the City of Toronto's Transportation Services Division as their data source. Some municipalities used a combination of sources. For example, Rio de Janeiro, while using IPCC as their primary source of methodology, also used input from Academia, the World Resources Institute (WRI), and the World Bank.

Approximately thirty percent of C40 cities reported the detailed methodology used for data collection and $FFCO_2$ emission estimation. Of those cities reporting specific methodologies, Stockholm, Berlin, Portland and Vancouver used top down method. Stockholm and Berlin disaggregated fuel consumption data based on fuel sales data from road traffic. Portland disaggregated fuel sales data, taken from EIA Energy Information Administration (EIA). Vancouver disaggregated fuel sales data, but did not identify the source. Toronto and London used bottom up method based on traffic activity data. Oslo used a combination of fuel sales and activity data.

Appendix A.3.2. Millennium City Database

The Millennium City Database (MCD) is compiled by the International Association of Public Transport (UITP) in collaboration with Murdoch University in Perth, Australia [37]. MCD contains over 100 worldwide cities with data on land use, transportation, economics and energy use for the year 1995/1996. MCD data were collected from published sources (such as national censuses and annual reports) and many emails exchanged with individuals working in various corresponding organizations in each city [37]. To synchronize with C40 data for the year 2011, MCD data were scaled using a growth factor of national-scale CO_2 emissions from the transportation sector in the International

Energy Agency greenhouse gas reports [2]. This is a compromise of data quality since the IEA national transportation GHG emission growth rate is used to scale city emission data.

References

1. Owen, W. Transportation and Economic Development. *Am. Econ. Rev.* **1959**, *49*, 179–187.
2. World Energy Outlook 2017 Home Page. Available online: https://doi.org/10.1787/weo-2017-en (accessed on 14 November 2019).
3. Gurney, K.R.; Mendoza, D.L.; Zhou, Y.; Fischer, M.L.; Miller, C.C.; Geethakumar, S.; de la Rue du Can, S. High Resolution Fossil Fuel Combustion CO_2 Emission Fluxes for the United States. *Environ. Sci. Tech.* **2009**, *43*, 5535–5541. [CrossRef] [PubMed]
4. Lakshmanan, T.R.; Han, X. Factors Underlying Transportation CO_2 Emissions in the USA: A Decomposition Analysis. *Transp. Res. Part D Transp. Environ.* **1997**, *2*, 1–15. [CrossRef]
5. Cervero, R.; Hansen, M. Induced Travel Demand and Induced Road Investment: A Simultaneous Equation Analysis. *J. Transp. Econ. Policy* **2007**, *36*, 469–490.
6. Lu, I.J.; Lin, S.J.; Lewis, C. Decomposition and Decoupling Effects of Carbon Dioxide Emission from Highway Transportation in Taiwan, Germany, Japan and South Korea. *Energy Policy* **2007**, *35*, 3226–3235. [CrossRef]
7. Liddle, B. Long-run Relationship among Transport Demand, Income, and Gasoline Price for the US. *Transp. Res. Part D Transp. Environ.* **2009**, *14*, 73–82. [CrossRef]
8. Timilsina, G.R.; Shrestha, A. Transport Sector CO_2 Emissions Growth in Asia: Underlying Factors and Policy Options. *Energy Policy* **2009**, *37*, 4523–4539. [CrossRef]
9. Barla, P.; Miranda-Moreno, L.F.; Lee-Gosselin, M. Urban Travel CO2 Emissions and Land Use: A Case Study for Quebec City. *Transp. Res. Part D Transp. Environ.* **2011**, *16*, 423–428. [CrossRef]
10. Wang, H.; Zhang, R.; Liu, M.; Bi, J. The Carbon Emissions of Chinese Cities. *Atmos. Chem. Phys.* **2012**, *12*, 6197–6206. [CrossRef]
11. Newman, P.; Kenworthy, J.R. Gasoline Consumption and Cities: A Comparison of US Cities with a Global Survey. *J. Am. Plan. Assoc.* **1989**, *55*, 24–37. [CrossRef]
12. Eltony, M.N.; Al-Mutairi, N.H. Demand for Gasoline in Kuwait: An Empirical Analysis Using Cointegration Techniques. *Energy Econ.* **1995**, *17*, 249–253. [CrossRef]
13. Espey, M. Gasoline Demand Revisited: An International Meta-Analysis of Elasticities. *Energy Econ.* **1998**, *20*, 273–295. [CrossRef]
14. Cervero, R.; Landis, J. Suburbanization of Jobs and the Journey to Work. *J. Adv. Transp. Banner* **1991**, *26*, 275–297. [CrossRef]
15. Wang, S.; Ma, H.; Zhao, Y. Exploring the Relationship between Urbanization and the Eco-environment—A Case Study of Beijing–Tianjin–Hebei Region. *Ecol. Indic.* **2014**, *45*, 171–183. [CrossRef]
16. Wheaton, W.C. Land Use and Density in Cities with Congestion. *J. Urban Econ.* **1998**, *43*, 258–272. [CrossRef]
17. Holden, E.; Norland, I. Three Challenges for the Compact City as a Sustainable Urban Form: Household Consumption of Energy and Transport in Eight Residential Areas in the Greater Oslo Region. *Urban Stud.* **2005**, *42*, 2145–2166. [CrossRef]
18. Bouwman, M.E.; Moll, H.C. Environmental Analyses of Land Transportation Systems in The Netherlands. *Transp. Res. Part D Transp. Environ.* **2002**, *7*, 331–345. [CrossRef]
19. Echenique, M.H.; Hargreaves, A.J.; Mitchell, G.; Namdeo, A. Growing Cities Sustainably: Does Urban Form Really Matter? *J. Am. Plan. Assoc.* **2012**, *78*, 121–137. [CrossRef]
20. Hansen, M.; Huang, Y. Road Supply and Traffic in California Urban Areas. *Transp. Res. Part A* **1997**, *31*, 205–218. [CrossRef]
21. Noland, R.B.; Lem, L.L. A review of the Evidence for Induced Travel and Changes in Transportation and Environmental Policy in the US and the UK. *Transp. Res. Part D Transp. Environ.* **2002**, *7*, 1–26. [CrossRef]
22. Fulton, L.; Meszler, D.; Noland, R.; Thomas, J. A Statistical Analysis of Induced Travel Effects in the US Mid-Atlantic Region. *J. Transp. Stat.* **2000**, *3*, 1–14.
23. Johansson, O.; Schipper, L. Measuring the Long-Run Fuel Demand of Cars: Separate Estimations of Vehicle Stock, Mean Fuel Intensity, and Mean Annual Driving Distance. *J. Transp. Econ. Policy* **1997**, *31*, 277–292.
24. Ewing, R.; Bartholomew, K.; Winkelman, S.; Walters, J.; Chen, D. *Growing Cooler: The Evidence on Urban Development and Climate Change*; Urban Land Institute: Washington, DC, USA, 2008.

25. Karathodorou, N.; Graham, D.J.; Noland, R.B. Estimating the Effect of Urban Density on Fuel Demand. *Energy Econ.* **2010**, *32*, 86–92. [CrossRef]
26. Lee, S.; Lee, B. The Influence of Urban Form on GHG Emissions in the US Household Sector. *Energy Policy* **2014**, *68*, 534–549. [CrossRef]
27. Greenhouse Gas Interactive Dashboard Data, C40 Cities Climate Leadership Group. 2018. Available online: http://www.c40.org/other/gpc-dashboard (accessed on 14 November 2019).
28. Kenworthy, J. Trends in Low Carbon Transport and Urban Development in 33 Cities, 1995/1996 to 2005/2006. In *Low Carbon Cities*; Steffen, L., Ed.; Routledge: London, UK, 2014; pp. 113–130.
29. World Bank. *World Development Indicators*; The World Bank: Washington, DC, USA, 2012; Available online: http://data.worldbank.org/data-catalog/world-development-indicators (accessed on 14 November 2019).
30. York, R.; Rosa, E.A.; Dietz, T. STIRPAT, IPAT and ImPACT: Analytic Tools for Unpacking the Driving Forces of Environmental Impacts. *Ecol. Econ.* **2003**, *46*, 351–365. [CrossRef]
31. Liddle, B. Urban Density and Climate Change: A STIRPAT Analysis Using City-level Data. *J. Transp. Geogr.* **2013**, *28*, 22–29. [CrossRef]
32. Dietz, T.; Rosa, E.A. Rethinking the Environmental Impacts of Population, Affluence and Technology. *Hum. Ecol. Rev.* **1994**, *1*, 277–300.
33. Dietz, T.; Rosa, E.A. Effects of Population and Affluence on CO2 Emissions. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 175–179. [CrossRef]
34. Dobson, J.E.; Bright, E.A.; Coleman, P.R.; Durfee, R.C.; Worley, B.A. A Global Population Database for Estimating Populations at Risk. *Photogramm. Eng. Remote Sens.* **2000**, *66*, 849–857.
35. Haklay, M. How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets. *Environ. Plan. B Plan. Des.* **2010**, *37*, 682–703. [CrossRef]
36. Girres, J.-F.; Touya, G. Quality Assessment of the French OpenStreetMap Dataset. *Trans. GIS* **2010**, *14*, 435–459. [CrossRef]
37. Kenworthy, J.; Laube, F. Millennium in Cities Database. 2001. Available online: https://www.uitp.org/MCD (accessed on 14 November 2019).
38. LandScan, (2011), This product was made utilizing the LandScan (2011)™ High Resolution global Population Data Set copyrighted by UT-Battelle, LLC, operator of Oak Ridge National Laboratory under Contract No. DE-AC05-00OR22725 with the United States Department of Energy. The United States Government has certain rights in this Data Set. Neither UT-BATTELLE, LLC NOR THE UNITED STATES DEPARTMENT OF ENERGY, NOR ANY OF THEIR EMPLOYEES, MAKES ANY WARRANTY, EXPRESS OR IMPLIED, OR ASSUMES ANY LEGAL LIABILITY OR RESPONSIBILITY FOR THE ACCURACY, COMPLETENESS, OR USEFULNESS OF THE DATA SET. Available online: https://landscan.ornl.gov/landscan-datasets (accessed on 14 November 2019).
39. OpenStreetMap Contributors, Planet Dump. 2011. Available online: https://planet.openstreetmap.org (accessed on 14 November 2019).
40. Elvidge, C.D.; Tuttle, B.T.; Sutton, P.C.; Baugh, K.E.; Howard, A.T.; Milesi, C.; Bhaduri, B.; Nemani, R. Global Distribution and Density of Constructed Impervious Surfaces. *Sensors* **2007**, *7*, 1962–1979. [CrossRef] [PubMed]
41. Berube, A.; Friefhoff, A.; Nadeau, C.A.; Rode, P.; Paccoud, A.; Kandt, J.; Just, T.; Schemm-Gregory, R. Global Metro Monitor. *Polit. Sci.* **2010**, 150857366.
42. Cuenot, F.; Fulton, L. *International Comparison of Light-duty Vehicle Fuel Economy and Related Characteristics*; Global Fuel Economy Initiative, International Energy Agency, Eds.; International Energy Agency: Paris, France, 2011.
43. World Road Statistics 2014. International Road Federation Home Page. Available online: https://irfnet.ch/ (accessed on 14 November 2019).
44. Doane, D.P.; Seward, L.E. Measuring Skewness: A Forgotten Statistic? *J. Stat. Educ.* **2011**, *19*. [CrossRef]
45. Kutner, M.H.; Nachtsheim, C.J.; Neter, J. *Applied Linear Regression Models*, 4th ed.; Higher Education Press: Beijing, China, 2004.
46. Neumayer, E.; Plümper, T. *Robustness Tests for Quantitative Research*; Cambridge University Press: Cambridge, UK, 2017.
47. Bento, A.; Cropper, M.; Mobarak, A.; Vinha, K. The Impact of Urban Spatial Structure on Travel Demand in the United States. *Worldbank Policy Res. Work. Paper Ser.* **2003**, *3007*. [CrossRef]

48. Levinson, D.; Kumar, A. The Rational Locator: Why Travel Times Have Remained Stable. *J. Am. Plan. Assoc.* **1994**, *60*, 319–332. [CrossRef]
49. Qing, S. The Effect of Population Density, Road Network Density, and Congestion on Household Gasoline Consumption in U.S. Urban Sreas. *Energy Econ.* **2011**, *33*, 445–452.
50. Jarque, C.M.; Bera, A.K. Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals. *Econ. Lett.* **1980**, *6*, 255–259. [CrossRef]
51. Lu, X.; White, H. Robustness Checks and Robustness Tests in Applied Economics. *J. Econom.* **2014**, *178*, 194–206. [CrossRef]
52. Australian Bureau of Statistics Home Page. Available online: https://www.abs.gov.au/AUSSTATS/abs@.nsf/DOSSbyTopic/1FF8657E5F8E2AC6CA256BD00027143D?OpenDocument (accessed on 14 November 2019).
53. Bonilla, D.; Foxon, T. Demand for New Car Fuel Economy in the UK, 1970−2005. *J. Transp. Econ. Policy* **2009**, *43*, 55–83.
54. Small, K.A.; Van Dender, K. Fuel Efficiency and Motor Vehicle Travel: The Declining Rebound Effect. *Energy J.* **2007**, *28*, 25–52. [CrossRef]
55. Funke, S.Á.; Sprei, F.; Gnann, T.; Plötz, P. How Much Charging Infrastructure Do Electric Vehicles Need? A Review of the Evidence and International Comparison. *Transp. Res. Part D Transp. Environ.* **2019**, *77*, 224–242. [CrossRef]
56. Shao, L.; Karci, A.E.H.; Tavernini, D.; Sorniotti, A.; Cheng, M. Design Approaches and Control Strategies for Energy-Efficient Electric Machines for Electric Vehicles—A Review. *IEEE Access* **2020**, *8*. [CrossRef]
57. Jones, C.; Kammen, D.M. Spatial Distribution of US Household Carbon Footprints Reveals Suburbanization Undermines Greenhouse Gas Benefits of Urban Population Density. *Environ. Sci. Technol.* **2014**, *48*, 895–902. [CrossRef] [PubMed]
58. Matisoff, D.C.; Noonan, D.S.; O'Brien, J.J. Convergence in Environmental Reporting: Assessing the Carbon Disclosure Project. *Bus. Strategy Environ.* **2013**, *22*, 285–305. [CrossRef]
59. Forghani, M.; Delavar, M.R. A Quality Study of the OpenStreetMap Dataset for Tehran. *Int. J. Geo-Inf.* **2014**, *3*, 750–763. [CrossRef]
60. Goodchild, M.F. Citizens as Sensors: The World of Volunteered Geography. *GeoJournal* **2007**, *69*, 211–221. [CrossRef]
61. Wheaton, W.C. Commuting, Congestion, and Employment Dispersal in Cities with Mixed Land Use. *J. Urban Econ.* **2004**, *55*, 417–438. [CrossRef]
62. Massey, D.S.; Denton, N. The Dimensions of Residential Segregation. *Soc. Forces* **1988**, *67*, 281–313. [CrossRef]
63. Lo, C.P. Testing Urban Theories Using Remote Sensing. *GIScience Remote Sens.* **2004**, *41*, 95–115. [CrossRef]
64. Dare, P.M. Shadow Analysis in High-resolution Satellite Imagery of Urban Areas. *Photogramm. Eng. Remote Sens.* **2005**, *71*, 169–177. [CrossRef]
65. Pan, X.; Zhao, Q.; Chen, J.; Liang, Y.; Sun, B. Analyzing the Variation of Building Density Using High Spatial Resolution Satellite Images: The Example of Shanghai City. *Sensors* **2008**, *8*, 2541–2550. [CrossRef] [PubMed]
66. Yu, W.; Ai, T.; Shao, S. The Analysis and delimitation of Central Business District Using Network Kernel Density Estimation. *J. Transp. Geogr.* **2015**, *45*, 32–47. [CrossRef]
67. Wu, C.; Murray, A.T. Estimating Impervious Surface Distribution by Spectral Mixture Analysis. *Remote Sens. Environ.* **2003**, *84*, 493–505. [CrossRef]
68. Li, Z.C.; Chen, Y.J.; Wang, Y.D.; Lam, W.H.; Wong, S.C. Optimal Density of Radial Major Roads in a Two-dimensional Monocentric City with Endogenous Residential Distribution and Housing Prices. *Reg. Sci. Urban Econ.* **2013**, *43*, 927–937. [CrossRef]
69. Taubenböck, H.; Klotz, M.; Wurm, M.; Schmieder, J.; Wagner, B.; Wooster, M.; Esh, T.; Dech, S. Delineation of Central Business Districts in Mega City Regions Using Remotely Sensed Data. *Remote Sens. Environ.* **2013**, *136*, 386–401. [CrossRef]
70. Google Earth Home Page. Available online: www.google.com/earth (accessed on 14 November 2019).
71. Population and Housing Unit Estimates Home Page. Available online: https://susy.mdpi.com/manuscript/summary/f3c707bab1f3854f0c2fc6fb4aefa25c (accessed on 14 November 2019).
72. The Intergovernmental Panel on Climate Change. *IPCC Guidelines for National Greenhouse Gas Inventories*; Eggleston, H.S., Buendia, L., Miwa, K., Ngara, T., Tanabe, K., Eds.; IGES: Kanagawa, Japan, 2006.

73. Kennedy, C.; Steinberger, J.; Gasson, B.; Hansen, Y.; Hillman, T.; Havránek, M.; Mendez, G.V. Methodology for Inventorying Greenhouse Gas Emissions from Global Cities. *Energy Policy* **2010**, *38*, 4828–4837. [CrossRef]
74. Chavez, A.; Ramaswami, A.; Nath, D.; Guru, R.; Kumar, E. Implementing Trans-boundary Infrastructure-Based Greenhouse Gas Accounting for Delhi, India: Data Availability and Methods. *J. Ind. Ecol.* **2012**, *16*, 814–828. [CrossRef]