npj | sustainable mobility and transport

Article

https://doi.org/10.1038/s44333-025-00057-1

Understanding the scaling of transport energy use with operational density

Check for updates

Anupriya[1] ✉, Emma McCoy[2] & Daniel J. Graham[1]

Given the substantial contribution of transport operations to global energy demand, enhancing their energy efficiency is crucial for sustainable urban mobility. This study investigates whether intensifying the use of fixed transport networks, termed operational densification, reduces energy consumption. Grounded in economic theory, we develop a novel causal model to estimate the energy impacts of densification across two major commuting modes: urban rail transit (metro) and private car travel. Using a unique panel dataset of 27 metro operations worldwide, we find that a 10% increase in passenger-kilometres travelled on a fixed network reduces energy use per passenger-kilometre by 3.45%. These gains surpass what kinetic energy principles alone predict, as fixed energy inputs such as infrastructure and maintenance are distributed across increased usage. In contrast, analysis of the Millennium Cities Database reveals no significant energy savings from densification in private car travel, likely due to limited shared infrastructure or operational scale economies.

The COVID-19 pandemic initiated a massive, unplanned shift to remote work in many advanced economies. This shift has led to widespread discussions about the future of work, with many employees appreciating the benefits of working from home. The debate extends to the potential decline of traditional office spaces, which raises questions about the future viability of densely populated urban centres, or cities [see, for instance, refs. 1–3]. Such conjectures have created the need for a deeper understanding of how densification benefits society and the economy at large.

A substantial body of research, primarily from urban economists and geographers, has focused on the effects of densification on productivity and innovation. The prevailing evidence demonstrates that cities and industrial clusters have strong positive statistical associations with the economic performance of firms and workers [see ref. 4]. Recently, as data on cities has become more comprehensive, the literature has broadened to explore how various urban variables beyond socio-economic activity, including land use, urban infrastructure (for instance, road network), and even fundamental individual needs (for instance, household electricity and water consumption) scale with city size, typically measured by population[5–8]. Central to this analysis is an urban scaling law of the form, $Y = Y_0N^{\beta}$, where Y denotes the response of interest, N represents city population, Y_0 is a constant and β signifies the scaling exponent. Existing evidence suggests super-linear ($\beta > 1$), sub-linear ($\beta < 1$) and linear ($\beta = 1$) for socio-economic responses, infrastructure provision, and variables describing basic individual needs, respectively. Theoretical frameworks suggest that the density of social network interactions in urban environments influences the scaling of various parameters, with those benefiting from *economies of scale* exhibiting super-linear growth.

Pertinent to this theme, an emerging strand in the literature evaluates how city size and density influence climate change outcomes. For instance, using data from the UK[9], showed that larger cities are generally more efficient than smaller towns in terms of energy consumption and waste generation, offering strong evidence in support of green and *sustainable urbanisation*. Similarly, analysing data from 274 cities worldwide[10], found that compact urban forms, when combined with transport planning that promotes higher population densities, lead to substantial reductions in both residential and transport-related energy use. At the global scale[11], further confirmed that increased urban density is associated with lower overall energy consumption, reinforcing the sustainability case for densification. Extending this line of evidence,[12], based on data from 113 countries, found that per capita energy use declines with both population density and the area of built-up land per capita, with the latter emerging as an even stronger predictor. These studies collectively highlight the role of urban form and density in shaping energy demand and mitigating climate impacts.

The energy demanded by the transport sector remains a key theme in studies of urban sustainability, as it accounts for ~30 percent of global energy consumption and 20 percent of greenhouse gas emissions, making it a significant contributor to climate change[13]. A number of studies have found that transport energy consumption tends to scale negatively with population density[14–19], suggesting that denser cities may

[1]Transport Strategy Centre, Imperial College London, London, UK. [2]Department of Statistics, London School of Economics, London, UK.
✉e-mail: anupriya15@imperial.ac.uk

Table 1 | Summary of key literature on the impact of urban density and environmental sustainability of transport [adapted from ref. 20]

Study	Outcome of interest	Density	Country	Methodology	Estimated scaling
10	Transport energy use	GDP per capita	Multiple	OLS	-0.900–0.900
17	GHG from transport	Population density	Multiple	OLS	-0.769
47	CO_2 from commutes	Population density	Italy	OLS with IV	-0.2346
48	CO_2 from private driving	Population density	US	Correlation	-0.0821
	CO_2 from public transport	Population density	US	Correlation	0.3685
49	CO_2 from transport	Population density	US	OLS with IV	-0.3100

be inherently more energy-efficient for transport. Synthesising findings from twenty-one such studies,[20] report an average elasticity of -0.07 (standard deviation = 0.10) for combined domestic and private vehicle energy use with respect to density, implying that higher density is generally associated with lower per capita energy use. Interestingly, the elasticity estimate for public transit energy use is reported as 0.37, albeit based on a single study, indicating a positive association with density. This could reflect increased public transport usage in denser areas, though the authors caution that most of these estimates are only associational and potentially spurious, and establishing causal relationships remains a methodological challenge. Relevant estimates from previous studies on this theme are summarised in Table 1.

While the existing literature offers important insights primarily into the scaling of overall transport energy requirements with density, a critical gap remains in understanding how these effects differ across specific modes of travel. This paper addresses that gap by evaluating the causal implications of operational densification, that is, increasing the intensity of energy use of a fixed network for two dominant urban transport modes: (i) urban rail transit (metro) and (ii) road-based private vehicular travel.

The study has important policy implications. Firstly, as energy prices around the world continue to soar to new record highs, our results could enable governments to implement more targeted measures to lower the energy consumed in daily travel activities. Secondly, our results could enable more effective implementation of transport solutions and policies to help cities achieve their net-zero targets. Finally, yet importantly, the estimates delivered by this study could add a novel dimension to the appraisal of the environmental impacts of transport investments.

We first study the causal impact of operational density on the energy usage of metro systems, commonly recognised as a key pillar of sustainable transport in cities. There is now a great deal of evidence suggesting that metro operations with a high density of usage are highly productive and cost-efficient, thus highlighting key aspects of their economic sustainability[21–23]. However, evidence of their environmental sustainability, particularly in terms of their energy use efficiency remains limited. To address this gap in the literature, we develop a novel model, guided by the economic theory of production, to study the energy usage of metro operations. The objective is to assess how operational density, that is, the total passenger kilometres travelled on a given network, affects the relative demand for energy in metro operations. The model is developed using a distinctive and rigorously curated panel dataset encompassing 27 metro systems across various regions worldwide. This dataset, maintained by the Transport Strategy Centre (TSC) at Imperial College London, has been systematically gathered since 1994, ensuring consistency and reliability in metro system performance benchmarking. The parameters of the model are identified via the application of random effects (RE) estimation. It is worth emphasising that one may think of the relationship between operational density and energy consumption in terms of the kinetic energy equation, which suggests constant proportional change in energy for change in mass (density). However, the equation represents just one aspect of the total energy requirement of a metro operation. Economies of scale may significantly reduce per-passenger kilometre energy consumption as usage increases, by spreading the fixed component of energy use (such as infrastructure and maintenance) over more passenger kilometres and optimising operational efficiencies. Thus, higher density may lead to more efficient energy use overall, beyond the simple linear relationship suggested by kinetic energy alone.

Next, we assess the causal impact of operational density on the energy usage of private vehicular operations on urban road networks. Contrary to metros, urban road networks with a high density of usage have, on average, been found to be less productive and cost-efficient due to technical inefficiencies resulting from increased congestion[24–26]. It is thus interesting to assess whether such effects transpire in their energy-use efficiency too. We investigate how operational density affects the relative demand for energy in road-based private car travel. The model of interest is developed using the Millennium Cities database for sustainable transport compiled by the International Association of Public Transport (UITP) and the application of ordinary least squares (OLS) estimation.

Methods

This section has four subsections. The first subsection describes the theoretical framework, followed by details of the empirical model and its estimation in the next subsection. The penultimate section discusses the causal identification strategy using the potential outcomes framework. The final subsection describes the data used in the empirical analysis.

Theoretical framework

Our approach to modelling energy use in transport operations is grounded in the economic theory of production, which explains the relationship between various inputs and outputs in a production process. Specifically, our model examines how the costs of running transport operations depend on input prices (such as energy and labour) and the level of output (such as the amount of goods transported). By using this model, we can analyse and understand the energy required for each unit of output, providing valuable insights into the efficiency and sustainability of transport operations.

We have the short-run variable cost function CV^s_{it} for a transport operation i at time t,

$$CV^s_{it} = f(y_{it}, N_{it}, w_{it}), \tag{1}$$

where y and N are measures of output (passenger-kilometres) and network size (operated route length), respectively, and w is a vector of prices for variable inputs, labour and energy. According to economic theory, the conditional factor share equations for input j, x_j can be derived from the short-run costs CV^s_{it} using Shepherd's lemma (the firm-year subscripts have been dropped for notational simplicity) as follows[27]:

$$\frac{\partial CV^s(y, w, N)}{\partial w_j} = x_j \tag{2}$$

Thus, requirement of input j by operation i at time t, $x_{j,it}$, can be represented as:

$$x_{j,it} = g(y_{it}, N_{it}, w_{it}) \tag{3}$$

Model specification and empirical estimation

Following from Eq. (3), the transport energy consumption equation can be represented using a flexible functional specification with second-order terms as follows:

$$\log E_{it} = \alpha_0 + \alpha_y \log y_{it} + \alpha_N \log N_{it} + \sum_{j=1}^{2} \alpha_j \log w_{j,it} + \beta_{yy} \log y_{it}^2 + \beta_{yN} \log y_{it} \log N_{it}$$
$$+ \beta_{NN} \log N_{it}^2 + \sum_{j=1}^{2} \beta_{jy} \log y_{it} \log w_{j,it} + \sum_{j=1}^{2} \beta_{jN} \log w_{j,it} \log N_{it}$$
$$+ \sum_{j=1}^{2}\sum_{k=1}^{2} \beta_{jk} \log w_{j,it} \log w_{k,it} + u_i + \delta_t + \epsilon_{it} \tag{4}$$

where δ_t are year dummies that capture the year-specific effects, u_i is a random effect associated with i-th operation (for instance, regional, topological, or managerial factors) and is assumed to be $u_i \sim \mathcal{N}(0, \sigma_u^2)$, and ϵ_{it} is the residual error term, assumed to be $\epsilon_{it} \sim \mathcal{N}(0, \sigma_\epsilon^2)$. Both u_i and ϵ_{it} are assumed to be uncorrelated with the covariates of the model. The adopted second-order polynomial specification often used to approximate an unknown functional form, is justified by Taylor's theorem. The specification can flexibly capture non-linearities and interaction effects between the covariates of the model. We estimate the parameters of this model using the RE estimator developed by ref. 28, also referred to as the Swami-Arora method for unbalanced panels. The analysis is conducted in `Stata` using the `xtreg` package.

From Eq. (4), the key estimand of interest that captures the impact of densification, or equivalently, the intensity of use of the transport network on the operational energy requirement can be determined as follows:

$$e_{it}^{en} = \frac{\partial \log E_{it}}{\partial \log y_{it}} = \alpha_y + 2\beta_{yy} \log y_{it} + \sum_{j=1}^{2} \beta_{jy} \log w_{j,it} + \beta_{yN} \log N_{it}. \tag{5}$$

The measure described above quantifies the rate at which energy consumption changes in response to variations in operational density, reflecting changes in output over a fixed network, ceteris paribus (conditional on other factors). In essence, it assesses the sensitivity of energy requirements to fluctuations in the density of operations. The measure can be referred to as the *elasticity* of energy consumption with respect to output. Importantly, the inclusion of higher-order terms in the energy model (Eq. (4)) enables this elasticity to vary systematically with output, network size, and input prices. This flexibility is critical for capturing non-linear effects and understanding how energy efficiency responds to differing levels of operational density and factor price environments, rather than imposing a constant elasticity across all observations.

Note that we are interested in understanding the causal impact of densification on the energy usage of metro and private vehicular operations. Unobserved factors may exist that simultaneously influence both energy usage and density of operations, thereby obscuring their actual causal relationship in the observed data. Such spurious influences are commonly referred to as sources of endogeneity or confounding. For instance, we recognise unobserved (to the analyst) managerial efficiency or productivity in the case of a metro operation, or unobserved technical efficiency related to driver's behaviour or vehicular characteristics in the case of road-based private car travel to be a likely source of endogeneity. We argue that unobserved productivity/ technical efficiency does not only influence the outputs and costs of the production process[21,29–34], but these effects also occur in factor demand models as they are transmitted to factor share equations (see, for instance, Eq. (2)) via optimising behaviour. As such, the factor plays an important role in determining the quantity of energy consumed in the production of a given level of output. Higher efficiency is likely to result in lower energy consumption per unit output. It is, therefore, important to adjust for potential biases from endogeneity in the estimation of the energy usage model.

Accordingly, as a robustness check, we also estimate Eq. (4) using instrumental variables (IV) estimation. Specifically, we employ a set of time-varying IVs that are strongly correlated with the model's endogenous covariates but do not directly affect the dependent variable. Given the absence of suitable external instruments, we leverage the panel structure of the dataset to construct relevant IVs. In particular, for the differenced equations, we use lagged levels of endogenous covariates as instruments and for the levels equation, we use lagged first differences of covariates[35,36]. The model parameters are then estimated using the dynamic panel generalised methods of moments (DPGMM) approach.

We emphasise at this point that the Millennium Cities database for sustainable transport used for estimation of the energy model for road-based private-vehicular travel only comprises the cross-sectional dimension, thereby hindering the ability to derive appropriate IVs from the dataset itself. We, therefore, estimate this model only via OLS estimation. However, to assess the robustness of our results in light of the endogeneity concerns discussed above, we use a panel dataset of road-based private vehicular travel across US states and compare the OLS and IV estimates.

Causal interpretation within the potential outcomes framework

While our empirical specification is grounded in the economic theory of production, our interest in identifying the causal impact of densification (measured as output per unit of network) on energy consumption motivates an explicit framing within the potential outcomes framework for causal inference introduced by [37] [refer to for a detailed review][38,39]. This framework allows us to clarify the conditions under which the parameter estimates from our outcome regression model (Eq. (4)) can be interpreted as causal.

We have E_{it} representing the energy consumption of transport operation i at time t. We define the potential outcome $E_{it}(y/N)$ as the energy that would be consumed if the operation were exposed to specific values of operational density y/N. Our key estimand of interest is the average potential outcome, $\mathbb{E}\left[E_{it}(y/N)\right]$ which captures the expected energy use under hypothetical interventions on operational density.

To estimate this quantity, we rely on a causal structural model that is consistent with the energy demand function in Eq. (3). We interpret Eq. (4) as a flexible outcome regression model that estimates the conditional expectation of energy usage given observed covariates. Under the assumption of unconfoundedness, that is, conditional on input prices (w_{it}), the potential outcome $E_{it}(y/N)$ is independent of the treatment variable y_{it}/N_{it}, formally expressed as $E_{it}(y/N) \perp\!\!\!\perp y_{it}/N_{it} | w_{it}$—the coefficients on y_{it}/N_{it} in Eq. (4) obtained via OLS estimation can be interpreted as causal effects.

We encode this identification strategy in a theory-guided causal diagram, that is, a directed acyclic graph (DAG), in which output, network size, and input prices influence energy consumption, possibly confounded by unobserved factors such as managerial efficiency or operational quality as shown in Fig. 1.

This justifies the inclusion of panel data methods such as RE estimation, which models U_{it} as a time-invariant random effect $U_i \sim \mathcal{N}(0, \sigma_U^2)$, and DPGMM estimation, which employs instrumental variables Z_{it} to adjust for the potential endogeneity of the regressors.

Taken together, our approach combines structural economic modelling, causal DAG reasoning, and potential outcomes-based identification, providing a rigorous foundation for interpreting the elasticity estimates in Eq. (5) as capturing the causal effect of densification on energy consumption.

Data

To analyse metro energy consumption, we utilise data compiled by a consortium of metro operators known as the Community of Metros (https://communityofmetros.org/), which is overseen by the TSC. This consortium specialises in performance benchmarking, drawing on a comprehensive dataset that includes key indicators from 45 metro networks across 42 global cities marked in Fig. 2. However, the dataset contains gaps due to variations in reporting by individual operators each year. As a result, our study relies on

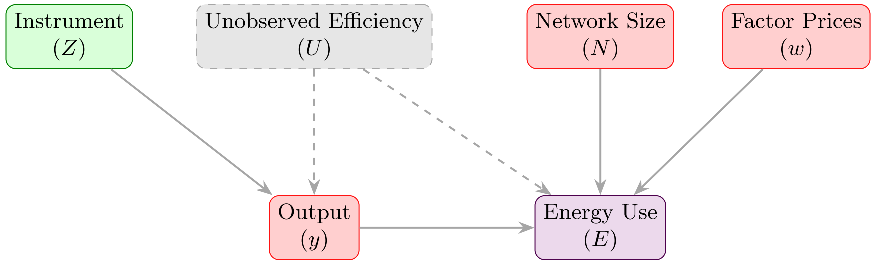

Fig. 1 | Theory-guided causal DAG for energy use in transport operations. This figure presents a causal diagram used to guide the identification strategy for estimating the determinants of energy consumption in transport operations. The diagram encodes a structural hypothesis about the relationship between key variables, informed by economic theory. The outcome variable of interest, energy use (purple rounded rectangle on the right), is influenced by multiple observed and unobserved factors. Observed covariates include output, network size, and factor prices (pink rectangles). Output refers to the scale of transport service provided, represented by passenger-kilometers. Network size represents infrastructure or service coverage via track length. Factor prices capture the cost of inputs including labour or energy. An instrumental variable, labelled as instrument (green rounded rectangle on the left), is used to account for endogeneity in output. This variable, denoted as Z, affects energy use only through its influence on output, satisfying exclusion restrictions under standard instrumental variable assumptions. Unobserved confounding is represented by a latent variable labelled unobserved efficiency (grey dashed rectangle at the top), which includes unmeasured managerial quality, or regional factors that may simultaneously influence both output and energy use. This confounding is depicted by grey dashed arrows from unobserved efficiency to both output and energy use. Causal pathways are shown using solid grey arrows. The solid grey arrow from output to energy use represents the main causal relationship of interest: the effect of producing more output on energy consumption, *ceteris paribus*. Additional solid grey arrows capture direct influences of network size and factor prices on energy use, representing other known energy use drivers in transport operations. The dashed grey arrows from the unobserved efficiency node to output and energy use depict the presence of omitted variable bias, which motivates the use of methods such as RE estimation, treating unobserved efficiency as a time-invariant, normally distributed factor, and DPGMM estimation, which allows for flexible modelling of latent heterogeneity and supports the use of instrumental variables to isolate exogenous variation in output. Together, this DAG clarifies the assumptions and identification strategy underpinning the empirical analysis and justifies the econometric approaches employed in the study.

Fig. 2 | Metro operations reported in the community of metros data. This figure presents the locations of metro systems included in the Community of Metros (CoMET), a global benchmarking consortium of large urban rail operators coordinated by the TSC at Imperial College London. Each blue markers denotes metro systems that contribute operational and performance data to the CoMET dataset. The map shows broad geographic representation, with systems located across all inhabited continents. In particular, metro networks are well represented in Europe and East Asia, alongside key systems in North America, Latin America, Oceania, South Asia, and select cities in the Middle East. The distribution captures systems in cities such as Tokyo, Seoul, New York, London, São Paulo, and Delhi, reflecting the diversity of CoMET's member base. The data used in this study were compiled from this consortium and span from 2006 to 2019, covering 27 systems with valid annual data on energy consumption and output. These data were subjected to rigorous validation procedures, including direct correspondence with member operators and systematic outlier detection, ensuring the reliability of the dataset. However, due to differences in reporting practices and participation across years, the resulting panel is unbalanced. This anonymised global map supports the empirical analysis of metro energy efficiency presented in the study and illustrates the range of operating environments, from dense, mature networks in high-income cities to rapidly expanding systems in emerging urban areas. It contextualises the generalisability of findings derived from this internationally diverse, yet operationally comparable, set of metro systems.

an unbalanced panel dataset with 174 observations, covering 27 metro systems over a 15 year period from 2006 to 2019. Notably, this dataset has undergone rigorous cleaning, including direct validation with operators and systematic verification procedures, ensuring high data quality. Given the commercially sensitive nature of the information, we present our findings in an anonymised format.

Our analysis uses total energy consumption, measured in megawatt hours, as the dependent variable. Passenger-kilometres serve as the measure of output, representing the total distance travelled by all passengers, including fare evaders. Network size is defined as the length of track actively used by in-service trains, excluding depot, yard, and siding tracks, as well as those designated for turning movements. We calculate factor prices for labour and energy by dividing total labour costs by total labour hours and total energy costs by total energy consumption, respectively. These prices are then adjusted to their 2016 international dollar equivalents. Consistent with the discussion presented in the previous subsection, we treat both output and input prices as endogenous variables. Table 2 provides descriptive statistics for all variables used in the analysis.

The energy consumption model for private vehicle use on urban road networks is estimated using cross-sectional data from the Millennium Cities Database for Sustainable Transport (https://www.uitp.org/publications/mobility-in-cities-database/), compiled by UITP. This dataset includes information on 100 cities worldwide marked in Fig. 3 for the year 1995, though only 84 cities from 42 countries were used in the estimation due to incomplete data for some locations. The cross-sectional nature of the dataset raises concerns about omitted variable bias in regression analysis. However, the database was designed with a strong emphasis on data consistency. Additionally, addressing the research question in this study requires substantial variation in urban density, which is difficult to capture in time-series data due to its persistence over time. To our knowledge, no existing panel dataset provides a sufficiently broad cross-section at the city level to meet these requirements.

Table 2 | Summary statistics for variables used in the metro analysis

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Total energy consumption (MWh)	232	559.59	486.06	50.33	2517.55
Variable energy consumption (MWh)	232	342.20	291.37	37.94	1687.61
Passenger kilometres (m)	232	7440.03	8205.71	289.08	37839.90
Network length (km)	232	152.21	134.70	31.62	675.80
Labour price (PPP US$/h)	232	32.45	15.88	6.15	73.79
Energy price (PPP US$/MWh)	232	0.18	0.07	0.05	0.48

Obs. Observations, *Std. Dev.* Standard Deviation, *m* millions, *MWh* Megawatt hours.

The dependent variable, total energy consumption (measured in megawatt hours), is calculated by multiplying reported energy use per passenger kilometre by the total number of passenger kilometres. Output is represented by passenger kilometres, which is determined by multiplying the reported private passenger kilometres per capita by the city's population. The unit price of labour is approximated using metropolitan gross domestic product per capita as recorded in the dataset. The unit price of energy is derived by dividing the cost of fuel per kilometre by the energy consumption per private passenger vehicle kilometre. Network size is estimated by multiplying the reported road length per 1000 residents by the city's total population. Table 3 provides descriptive statistics for all variables used in the analysis.

Additionally, for the purpose of a robustness check, we assemble a panel dataset covering road-based private vehicular travel across fifty US states and the District of Columbia from 1994 to 2019. We compile state-level data on annual vehicle miles travelled (as a measure of output), lane miles of rural and urban roads (representing network length), and total

Fig. 3 | Private vehicular operations reported in the Millennium Cities Database for Sustainable Transport. Global distribution of cities included in the Millennium Cities Database for Sustainable Transport. This figure shows the geographic coverage of the Millennium Cities Database for Sustainable Transport (MCD), compiled by the International Association of Public Transport (UITP). Each red location pin on the map represents a city included in the MCD dataset, which comprises cross-sectional data on urban transport and land use indicators from the year 1995. A total of 100 cities across all inhabited continents are represented in the dataset. The map highlights the broad global scope of the MCD, with a strong concentration of cities in Europe, North America, East Asia, and Oceania. Additional coverage spans major metropolitan areas in Latin America, Africa, South Asia, and the Middle East, ensuring a wide diversity in terms of development levels, geography, and urban form. The red pins mark the cities for which the MCD provides information on energy consumption, infrastructure characteristics, and other transport energy use-related metrics. However, due to missing or incomplete data for some variables in a subset of cities, only 84 cities from 42 countries were retained for the analysis of private vehicle energy consumption. The selection process ensured the analytical sample remained representative of different urban typologies while maintaining data quality. This figure is used to contextualise the spatial diversity of the cities included in the estimation of road transport energy use. The global spread supports comparative analysis and cross-regional insights, although the single-year nature of the dataset limits temporal inference. Despite its age, the MCD remains one of the most comprehensive and internationally harmonised sources of urban transport data, offering a valuable empirical foundation for understanding the structural drivers of energy demand in private vehicle use.

Table 3 | Summary statistics for variables used in the private vehicular travel analysis

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Total energy consumption (MWh)	86	71503.51	120988.80	472.83	767134.40
Passenger kilometres (m)	86	25439.08	38863.15	487.45	226962.80
Network length (km)	86	11527.69	19529.95	815.52	134815.50
GDP per capita	86	21789.17	14804.92	395.64	54692.08
Energy price	86	24.65	41.09	1.73	263.38

Obs. Observations, *Std. Dev*. Standard Deviation, *m* millions, *MWh* Megawatt hours.

Table 4 | The estimated total energy consumption model for metro operations

Covariate	Coefficient	Std. Error	*p*-value
Output	0.754	0.253	0.003
Network Length	1.555	0.282	0.000
Labour Price	0.085	0.373	0.820
Energy Price	-1.172	0.494	0.019
Output^2	0.015	0.014	0.288
Output x Network Length	-0.173	0.046	0.000
Output x Labour Price	0.030	0.048	0.536
Output x Energy Price	-0.201	0.051	0.000
Network Length^2	0.083	0.045	0.067
Network Length x Labour Price	0.107	0.065	0.098
Network Length x Energy Price	0.526	0.082	0.000
Labour Price^2	-0.099	0.047	0.038
Energy Price^2	0.034	0.080	0.673
Labour Price x Energy Price	0.052	0.103	0.611
Constant	-5.478	1.518	0.000
Year Effects		Yes	
Elasticity w.r.t. output	0.655	0.022	0.000
Elasticity w.r.t. network size	0.304	0.032	0.000
Elasticity w.r.t. energy price	-0.298	0.035	0.000
No. of Observations			232
R-squared			0.978
Within-group variance, σ_u			0.000
Between-group variance, σ_ϵ			0.066

annual fuel consumption from the Highway Statistics series published by the Federal Highway Administration (Available at https://tinyurl.com/yfezptbp). State-level annual real GDP data are obtained from the US Bureau of Economic Analysis (Available at https://tinyurl.com/y6s4usev), and a proxy for labour price is constructed by dividing GDP by employment figures from the US Bureau of Labor Statistics (Available at https://tinyurl.com/2p9h7ce9). Annualised fuel prices are sourced from open datasets published by the US Energy Information Administration (Available at https://www.eia.gov/opendata/).

Results

This section has four subsections. The first and second subsections discuss the estimates of the energy consumption model (Eq. (4)) for urban rail transit (metro) and private vehicular operations on urban road networks, respectively, which is followed by a detailed discussion of the energy-use efficiency estimates in the penultimate subsection. The final subsection converts the energy-use efficiency estimates into associated climate change impacts.

Energy consumption model for metro operations

Our analysis reveals that high-density metro systems are significantly more energy-efficient, as evidenced by the elasticity estimates in Table 4, which presents the parameter estimates of the energy consumption model for metro operations (Eq. (4)) obtained via RE estimation. The reported elasticities of energy consumption with respect to key covariates are evaluated at the mean values of the data. The elasticity estimate for output is 0.655 (standard error = 0.022), indicating that a 10% increase in density-related factors leads to only a 6.55% rise in energy consumption, highlighting the substantial energy efficiency gained from increasing operated passenger kilometres on a fixed network. Meanwhile, the elasticity of energy consumption with respect to network length is estimated at 0.304 (standard error = 0.032). Taken together, these results suggest that if both output and network size are increased equi-proportionately by 10%, overall energy use increases by just 9.59% (95% confidence interval = [9.25%, 9.94%]), pointing to modest energy savings even when metro systems are expanded. To assess robustness, we also estimated the model using dynamic panel generalised methods of moments (DPGMM) to account for potential endogeneity. The similarity of results across both methods suggests that our findings are robust to endogeneity concerns. For brevity, the DPGMM estimates are not reported.

The elasticity of energy consumption with respect to energy price provides further insight into operator behaviour. As shown in Table 4, the elasticity estimate is -0.298 (standard error = 0.035), indicating that a 10% increase in energy prices, holding other factors constant, leads to a 2.98% reduction in energy consumed by metro operations. This suggests that rising energy costs prompt operators to adopt energy-saving measures. These may include technological upgrades (such as modern, energy-efficient trains), operational optimisations (like reducing unnecessary acceleration and deceleration, maintaining consistent speeds, minimising idle times, or implementing automated train control systems), and strategic planning initiatives (such as optimising track layouts and using smoother tracks).

It is often argued that higher-density metro systems, that is, those with more passenger kilometres per route kilometre, are more energy efficient because they deliver more output relative to the fixed component of energy costs (for instance, from stations, ventilation systems), thereby spreading energy use over a larger number of passenger kilometres. However, the estimates in Table 4 do not indicate whether such efficiencies extend to the variable, traction-related component of energy use. In practice, high-density metro systems are more likely to incorporate advanced control technologies, such as communications-based train control, which reduce energy use by minimising acceleration and deceleration and improving overall system efficiency. On the other hand, larger systems may also operate heavier trains, which can be less energy-efficient.

To examine this further, we analyse the relationship between operational density and variable energy use. The key parameter of interest, the elasticity of variable energy use with respect to output is estimated at 0.678 (standard error = 0.030), as summarised in Table 5. The estimate implies that a 10% increase in passenger kilometres on a fixed network results in only a 6.78% increase in traction energy consumption. These findings confirm that denser metro systems are highly energy efficient, with efficiencies evident in both fixed and variable components of energy use. This provides strong support for the continued provision and expansion of metro services, particularly in city centres where high-frequency operations are critical to meeting travel demand.

Energy consumption model for private vehicular operations on urban road networks

Our analysis finds no evidence of energy savings from higher density in private vehicular operations. As shown in Table 6, the estimated elasticity of energy consumption with respect to output is 0.926 (standard error = 0.053), which is not significantly different from 1 at the 95% confidence level. This

Table 5 | The estimated variable energy consumption model for metro operations

Covariate	Coefficient	Std. Error	*p*-value
Output	1.026	0.343	0.003
Network Length	1.370	0.382	0.000
Labour Price	-0.789	0.506	0.120
Energy Price	-2.640	0.671	0.000
Output^2	0.133	0.020	0.000
Output x Network Length	-0.576	0.063	0.000
Output x Labour Price	-0.019	0.065	0.765
Output x Energy Price	-0.125	0.069	0.073
Network Length^2	0.432	0.061	0.000
Network Length x Labour Price	0.232	0.088	0.009
Network Length x Energy Price	0.675	0.111	0.000
Labour Price^2	0.007	0.064	0.908
Energy Price^2	-0.058	0.109	0.594
Labour Price x Energy Price	0.011	0.140	0.938
Constant	-6.369	2.058	0.002
Year Effects		Yes	
Elasticity w.r.t. output	0.678	0.030	0.000
Elasticity w.r.t. network size	0.217	0.043	0.000
Elasticity w.r.t. energy price	-0.239	0.048	0.000
No. of Observations			232
R-squared			0.955
Within-group variance, σ_u			0.001
Between-group variance, σ_ϵ			0.070

Table 6 | The estimated energy consumption model for road-based private vehicular operations

Covariate	Coefficient	Std. Error	*p*-value
Output	-1.879	1.480	0.208
Network Length	2.693	1.402	0.059
Labour Price	-0.462	1.610	0.775
Energy Price	-2.821	1.780	0.117
Output^2	-0.080	0.094	0.398
Output x Network Length	0.154	0.181	0.396
Output x Labour Price	0.262	0.105	0.016
Output x Energy Price	0.194	0.129	0.138
Network Length^2	-0.071	0.087	0.423
Network Length x Labour Price	-0.266	0.098	0.008
Network Length x Energy Price	-0.101	0.118	0.395
Labour Price^2	-0.006	0.067	0.923
Energy Price^2	0.118	0.073	0.110
Labour Price x Energy Price	0.103	0.123	0.404
Constant	9.507	11.534	0.413
Elasticity w.r.t. output	0.926	0.053	0.000
Elasticity w.r.t. network size	0.135	0.060	0.024
Elasticity w.r.t. energy price	-0.258	0.050	0.000
No. of Observations			86
R-squared7D2			0.980

Table 7 | Comparison of energy consumption elasticities obtained from different estimators for road-based private vehicular operations in the US states

	OLS estimation		DPGMM estimation	
Covariate	Coef.	Std. Err.	Coef.	Std. Err.
Elasticity w.r.t. output	0.957	0.014	1.019	0.039
Elasticity w.r.t. network size	0.044	0.014	-0.006	0.089
Elasticity w.r.t. energy price	0.090	0.048	0.134	0.121
No. of Observations	1275		1173	

suggests that increases in the density of private car travel on urban road networks do not yield statistically significant improvements in energy efficiency. When considered alongside the elasticity with respect to network length, the results further imply that if both the output and network size are increased equi-proportionately by 10%, energy consumption rises by 10.60% (95% confidence interval = [10.01%, 11.20%]). Thus, even the expansion of road networks for private vehicle use does not appear to offer any meaningful energy savings. Based on the sensitivity analysis provided in the next subsection, we assert that our estimates are robust to endogeneity concerns discussed in the Methods section, further reinforcing the reliability of the findings presented in Table 6.

We acknowledge that these findings are derived from cross-sectional UITP compiled in the mid-1990s, which may not fully reflect recent technological shifts such as the growing penetration of hybrid and electric vehicles. However, our supplementary analysis attached in the next subsection, based on a panel dataset of private vehicular operations across US states from 1994 to 2019, which encompasses later periods of vehicle electrification, produces elasticity estimates that are consistent in both magnitude and statistical significance with those obtained from the UITP data. This provides reassurance that our conclusions are not artefacts of the older dataset. Nonetheless, it is important to note that the US dataset reflects state-level travel patterns and energy use, and is not restricted to urban areas specifically.

As in the metro case, the elasticity with respect to energy price provides further insight into behavioural responses. The estimated price elasticity of energy consumption is -0.258 (standard error = 0.050), indicating that a 10% increase in energy prices, holding other factors constant, leads to a 2.58% reduction in energy use. This finding suggests that drivers respond to higher energy costs by adopting energy-saving practices, such as more efficient route planning, reducing idling, maintaining engines better, or switching to more fuel-efficient vehicles. Notably, our estimate aligns with a substantial body of existing literature, which places the price elasticity of energy consumption for private vehicular travel around -0.30 [see, ref. 40]. This consistency supports the well-established view that energy demand for private car use is price inelastic, that is, relatively unresponsive to changes in energy prices.

Robustness test

To assess the robustness of our energy model estimates for road-based private vehicular travel, we address the endogeneity concerns discussed in the Methods section by comparing results from OLS and IV-based estimation methods applied to a panel data relating to private car operations in US states between 1994−2019.

We begin by estimating the parameters of the energy consumption model for private vehicular travel using OLS. To address potential endogeneity bias, we also apply DPGMM estimation. The results from both approaches are closely aligned, indicating that our findings are robust to concerns about endogeneity. For brevity, Table 7 reports only the elasticities, which are central to our analysis. Notably, a two-sample *t*-test confirms that the elasticity estimates with respect to output are not statistically different between the two estimation methods at the 95% confidence level. This statistical similarity reinforces the robustness of our conclusions regarding the relationship between output and energy consumption in private vehicular operations.

Energy-efficiency

To further assess the energy efficiency of the two urban travel modes, we examine the elasticity of unit energy consumption, that is, energy consumed per passenger-kilometre, with respect to output (measured in passenger kilometres). The estimates, presented in Table 8, highlight a notable contrast between the two modes. For metro operations, the elasticity is statistically significant at the 95% confidence level and indicates that a 10% increase in output leads to a 3.45% reduction in energy consumption per unit of output. This underscores the energy savings achieved through more intensive use of a fixed metro network, where increased passenger kilometres lead to greater operational efficiency. In contrast, the elasticity estimate for private vehicular travel is statistically insignificant, indicating that increased intensity of road use does not yield comparable energy savings.

These differences are further illustrated in Fig. 4a, b, which plot estimated elasticities of unit energy consumption with respect to output across the full range of operational densities for metro and private vehicular operations, respectively. Both figures are based on model-derived elasticity estimates, incorporating parameter uncertainty through simulation. Shaded 95% confidence bands are shown to reflect the precision of these estimates.

Table 8 | The estimated unit energy consumption elasticities at mean levels of the data

Mode	Coefficient	Std. Error	95% Conf. Int.	
Urban rail transit	-0.345	0.022	-0.388	-0.301
Private vehicular travel	-0.074	0.053	-0.179	0.030

In the case of metro operations (Fig. 4a), the elasticity curve remains statistically significantly negative throughout the entire density range. However, the curve rises steadily with increasing density, indicating that while unit energy consumption continues to decline with higher output, the magnitude of efficiency gains diminishes. That is, the system exhibits economies of density, but with declining marginal returns as usage intensifies. At higher density levels, the curve flattens, suggesting a tendency toward constant energy efficiency intensity. The figure also disaggregates the elasticity curves by network size: small, medium, and large, as defined by the lower quartile, median, and upper quartile of the network length distribution. Notably, the elasticity curve for large networks lies above those for smaller systems, particularly at lower densities, as evidenced by the absence of overlapping confidence intervals. This pattern implies that larger metro systems may experience weaker energy efficiency gains from densification, potentially due to managerial inefficiencies, operational complexity, or reduced coordination effectiveness in more expansive networks.

For private vehicular operations (Fig. 4b), the energy efficiency curve exhibits a monotonically decreasing trend across the entire density range. However, the elasticity estimates remain statistically significantly different from zero at the 95% confidence level, reinforcing the conclusion that increased usage of road networks does not yield improvements in energy performance. In other words, despite higher utilisation, energy consumption per unit of output does not become more efficient. Furthermore, as

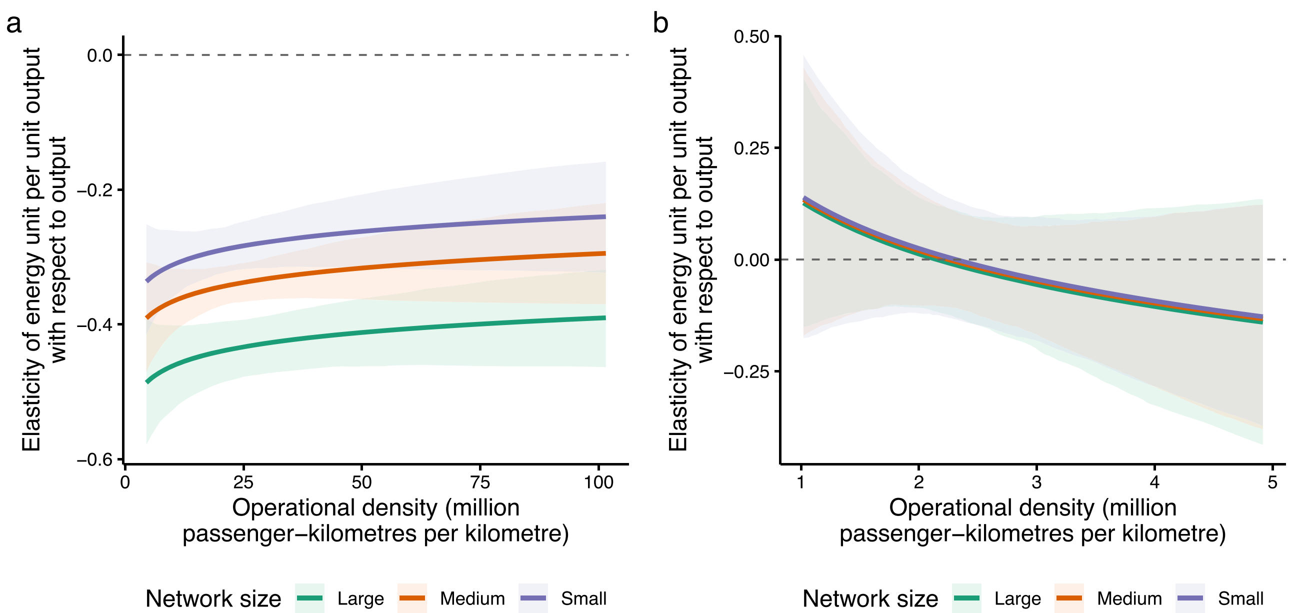

Fig. 4 | Variation of energy efficiency estimates of transport operations over density. **a** Metro operations. **b** Private vehicular travel. This figure consists of two panels comparing the elasticity of unit energy consumption with respect to output for two modes of urban transport: metro operations and private vehicular travel. Elasticity is defined as the percentage change in energy use per passenger-kilometre resulting from a one percent change in total output, measured in million passenger-kilometres. Negative values indicate energy efficiency improvements as output increases. **a** (distinct coloured lines with shaded regions) presents elasticity estimates for metro operations, disaggregated by network size. Three separate curves are shown, corresponding to small, medium, and large networks, defined using the lower quartile, median, and upper quartile of the network length distribution, respectively. The curves are generated from model-based simulations using a flexible translog specification, with 95% confidence bands displayed as shaded ribbons. All three curves remain negative across the full range of operational density, indicating that increases in output lead to reductions in unit energy consumption. However, the curves rise with density, suggesting that the magnitude of efficiency gains declines as usage increases. The flattening of the curves at high densities points to a tendency toward constant energy intensity in heavily used systems. Notably, larger networks exhibit less negative elasticity values, that is, smaller energy efficiency gains, possibly reflecting operational complexity or scale-related inefficiencies. **b** (overlapping coloured lines with shaded region) presents the corresponding analysis for private vehicular travel. Unlike the metro case, only a single curve appears to emerge, as disaggregation by network size yielded no meaningful variation. In contrast to metro systems, the elasticity curve remains statistically not different from zero across the density range. This suggests that road-based private vehicle operations do not exhibit gains from densification, that is, higher utilisation is associated with declining energy performance, and network size has no discernible effect on these estimates. Overall, these visualisations highlight a key contrast between transport modes. While metro systems benefit from densification, particularly smaller networks, private vehicular travel displays diseconomies of density, with increased usage associated with similar levels of unit energy consumption. The figure underscores the differing energy efficiency dynamics of infrastructure-fixed versus demand-responsive transport systems in urban contexts.

illustrated in the figure, variation in network size has no discernible effect on the elasticity estimates, suggesting that scale does not influence energy efficiency in the context of private vehicular travel on road networks.

Climate change impacts

To illustrate the energy efficiency benefits of operational densification, we compare two metro systems from our dataset that operate on networks of the same length (200 kilometres) but serve very different passenger volumes. The first system, which we refer to as 'Metro A', delivers 3000 million passenger-kilometres annually, while the second system, 'Metro B', delivers 18,000 million passenger-kilometres. Labour and energy prices for both operations are fixed at the mean values of the data. We use our estimated energy model (presented in Table 4) to predict the unit energy consumption for each of these operations.

This is not a simulation or counterfactual exercise, but rather an application of our model to two empirically observed operational profiles to demonstrate how energy consumption per unit of output varies with operational density. Based on model estimates, Metro A and Metro B consume 0.211 (standard error = 0.008) and 0.096 (standard error = 0.003) Watt-hours per passenger-kilometre, respectively. Thus, Metro B requires only 45.52% of the energy used by Metro A to produce one passenger-kilometre, reflecting the substantial energy savings associated with denser utilisation of the metro network.

To convey the scale of these energy savings in real-world terms, we project these per-unit energy figures onto a common benchmark of 11,600 million passenger-kilometres, which corresponds to the average annual metro travel across all systems in our sample in 2019. Based on this benchmark, Metro B would require 1330 MWh (standard error = 76 MWh) less energy than Metro A to produce the same level of output. According to the United States Environmental Protection Agency (US-EPA)'s Greenhouse Gas Equivalencies Calculator[41], this energy saving translates into a reduction of ~555 metric tons (standard error = 29 metric tons) of CO_2 emissions.

Furthermore, according to our energy consumption model for private-vehicular travel on road networks (see Table 5), we estimate the energy required to produce an output equivalent to Metro B, that is, 18,000 million passenger-kilometres. For this comparison, we assume a road network length of 800 kilometres. This choice is supported by conservative capacity estimates: even under idealised assumptions (for instance, 4-lane roads, 1.5 passengers per vehicle, and 10% of total travel occurring during the peak hour), a road network would require at least 685 kilometres to deliver 18 billion passenger-kilometres annually while operating at 80% of the theoretically maximum capacity (that is, 1500 vehicles/lane/hour capacity). Under these assumptions, private vehicles consume 1.704 (standard error = 1.092) Watt-hours per unit of output, meaning that the energy required to produce one passenger-kilometre is 1676.81% of the energy used by Metro B. To produce 11,600 million passenger-kilometres, private vehicles would require 18,650 MWh (standard error = 11,950 MWh) more energy than Metro B. According to the US-EPA's Greenhouse Gas Equivalencies Calculator[41], this additional energy corresponds to 12,530 metric tons (standard error = 4611 metric tons) of CO_2 emissions.

Discussion

The Energy Information Administration (EIA)'s *International Energy Outlook* (https://www.eia.gov/outlooks/ieo/) forecasts a 34% increase in global energy consumption between 2022 and 2050, primarily driven by population growth, economic development, and rising living standards. This growth is expected to result in a 15% rise in global carbon dioxide emissions from energy use. Critically, the EIA projects that energy demand will continue to outpace improvements in energy efficiency, suggesting that fundamental changes in consumption patterns will be essential to achieving "Net Zero Emissions" by 2050. These global projections underscore the urgent need for effective policy interventions to curb urban energy consumption, with urban transport emerging as a central domain for action.

In this context, our study offers novel insights into how operational densification influences the energy efficiency of two dominant modes of urban transport: metro rail systems and road-based private vehicles. Using theory-guided causal models and high-quality multi-city data, we find that densification of operations on metro networks leads to substantial energy savings. Specifically, we estimate that a 10% increase in passenger kilometres on a fixed metro network is associated with a 3.45% reduction in energy use per passenger kilometre. In contrast, we find no statistically significant energy efficiency gains from increased operational density in private vehicular travel. These findings align with the broader literature on production efficiency in urban transport systems: metro operations have been widely shown to exhibit economies of scale, whereby higher usage spreads the fixed component of energy and infrastructure costs more efficiently[21–23], while private road transport often suffers from diseconomies due to congestion and diminishing speed with increased volume[24–26]. This asymmetry underscores the limitations of private motorisation as a sustainable transport strategy in denser urban contexts and reinforces the case for investment in high-capacity public transit.

These findings have important implications for urban transport policy and planning. First, they suggest that urban densification alone is not sufficient for reducing energy demand unless it is coupled with investments in efficient, high-density public transport infrastructure, particularly high-capacity rail transit systems. Governments at the municipal, regional, and national levels can leverage these findings to inform policies that promote compact urban growth alongside public transit expansion. Examples include targeted incentives for high-frequency metro operations, integration of land use and transport planning to support transit-oriented development, and restrictions on private vehicle use in core urban areas through congestion pricing, parking limits, or low-emission zones (as seen in Stockholm or London).

In addition to informing operational and policy design, the results also offer a pathway for integrating energy efficiency into the formal appraisal of transport interventions. A useful analogy can be drawn with how wider economic impacts (WEIs) such as productivity gains from agglomeration [for details, refer to [4]] are currently included in cost-benefit analyses. In such frameworks, changes in generalised travel costs caused by a transport intervention are first used to estimate changes in effective economic density. These changes are then combined with empirically estimated agglomeration elasticities to derive the corresponding productivity effects. A similar two-step logic can be applied to energy-related environmental impacts. Specifically, model-based predictions of usage changes (for instance, increases in passenger-kilometres or shifts in operational density) resulting from a proposed transport scheme can be combined with the estimated elasticities of unit energy consumption presented in this study to quantify the likely energy efficiency benefits of densification. This provides a tractable and empirically grounded pathway for integrating energy savings and environmental performance into the broader impact assessment of transport policies and infrastructure investments.

Second, while our analysis focuses on metro rail and private vehicle modes, the underlying mechanism, that is, economies of scale in energy use due to increased operational density, is likely to extend to other high-capacity transit systems such as bus rapid transit (BRT) and light rail Transit (LRT). Prior studies have documented scale economies in the production of services for BRT and LRT systems [refs. 42–45] which suggests that densification-induced energy efficiency is also plausible in these modes, though the magnitude of savings may differ. For instance, the BRT systems in Curitiba and Bogotá, known for their high throughput, may experience similar benefits under increased passenger loads. A comparative analysis of energy performance across transit modes under different density conditions would be a valuable direction for future research.

It is also worth emphasising that although our findings for private vehicular travel are consistent across different datasets and estimation strategies, we recognise that the primary data used in the main analysis, that is, the Millennium Cities Database from the mid-1990s, has limitations in capturing the energy performance of contemporary vehicle fleets. In particular, the dataset predates the large-scale diffusion of hybrid and electric vehicles, as well as more recent advancements in engine efficiency and vehicle automation. Although our sensitivity analysis using US state-level panel data offers additional reassurance, these data are not urban-specific. Accordingly, we emphasise that future research should seek to revisit this analysis using updated and disaggregated datasets that better reflect the evolving energy profile of private vehicular travel in urban settings. This would not only improve the empirical basis for comparing modes, but also inform more nuanced policy design in the context of rapidly changing vehicle technologies.

Third, while our unit of analysis is the transport operation rather than the city, our results are highly relevant to the broader literature on urban scaling and sustainable development. The previous studies and policy reports, for instance, Intergovernmental Panel on Climate Change's *Mitigation of Climate Change* report (https://www.ipcc.ch/report/ar6/wg3/) and the Lancet's series on energy and health[46], emphasise the role of compact urban form and public transit accessibility in reducing emissions and improving population health. Our results complement these findings by showing how densification improves the energy performance of public transit operations which is a key operational link in the broader sustainability chain. More broadly, this research supports the argument that compact cities, when supported by efficient transit systems, can deliver both environmental and operational co-benefits. From a global policy perspective, our findings are generalisable to diverse urban contexts, provided that transit systems operate under conditions where increasing scale can be leveraged to improve efficiency. Policymakers in rapidly urbanising regions, particularly in the Global South, can apply these insights when making infrastructure investment decisions that align with climate goals and urban development priorities.

Finally, our study highlights several promising avenues for future work. One important extension involves linking the *structural density* of cities (for instance, population or employment density) to the *operational density* of transit systems. A richer understanding of how spatial concentration of people translates into demand intensity, and how that, in turn, affects energy use across travel modes, would help identify city-wide thresholds for sustainable modal shifts. For example, it is plausible that part of the energy savings observed in denser cities arises from a behavioural shift from private vehicles to metro, walking, or cycling. While beyond the scope of this paper, this hypothesis warrants empirical investigation and could serve as the basis for integrated modelling of urban form, travel behaviour, and energy outcomes.

In sum, our study provides robust empirical evidence that the intensity of use in metro systems leads to significant energy efficiency gains, while road-based private travel does not benefit similarly from densification. These results reinforce the case for policy frameworks that prioritise public transport investment in dense urban environments as a strategy for achieving sustainable and low-carbon cities.

Data availability

The metro operations dataset used in this study is commercially sensitive and proprietary to the Community of Metros, managed by the Transport Strategy Centre at Imperial College London. Access is restricted due to confidentiality agreements with participating operators and cannot be made publicly available or shared under approval-based mechanisms. The Millennium Cities Database for Sustainable Transport, compiled by the International Association of Public Transport, is also proprietary and accessible only through institutional purchase or license. Publicly available data sources used in the study are cited within the manuscript, and download links have been provided where applicable.

Code availability

The code developed for the analyses in this study, along with synthetic datasets that replicate the structure of the original data, can be made available upon request to the corresponding author.

Received: 24 February 2025; Accepted: 1 August 2025;
Published online: 01 September 2025

References

1. Swinney, P. et al. *Office Politics: London And The Rise of Home Working*. https://www.centreforcities.org/reader/office-politics/introduction/ (2023).
2. Althoff, L., Eckert, F., Ganapati, S. & Walsh, C. The geography of remote work. *Reg. Sci. Urban Econ.* **93**, 103770 (2022).
3. Florida, R., Rodríguez-Pose, A. & Storper, M. Cities in a post-covid world. *Urban Stud.* **60**, 1509−1531 (2021).
4. Graham, D. J. & Gibbons, S. Quantifying wider economic impacts of agglomeration for transport appraisal: existing evidence and future directions. *Econ. Transp.* **19**, 100121 (2019).
5. Bettencourt, L. M., Lobo, J., Helbing, D., Kühnert, C. & West, G. B. Growth, innovation, scaling, and the pace of life in cities. *Proc. Natl. Acad. Sci. USA* **104**, 7301–7306 (2007).
6. Bettencourt, L. M. The origins of scaling in cities. *science* **340**, 1438–1441 (2013).
7. Bettencourt, L. M. & Lobo, J. Urban scaling in europe. *J. R. Soc. Interface* **13**, 20160005 (2016).
8. Bettencourt, L. M. et al. The interpretation of urban scaling analysis in time. *J. R. Soc. Interface* **17**, 20190846 (2020).
9. Eeckhout, J. & Hedtrich, C. Green urbanization. *PLoS One* **16**, e0260393 (2021).
10. Creutzig, F., Baiocchi, G., Bierkandt, R., Pichler, P.-P. & Seto, K. C. Global typology of urban energy use and potentials for an urbanization mitigation wedge. *Proc. Natl. Acad. Sci. USA* **112**, 6283–6288 (2015).
11. Güneralp, B. et al. Global scenarios of urban density and its impacts on building energy use through 2050. *Proc. Natl. Acad. Sci. USA* **114**, 8945–8950 (2017).
12. Haberl, H. et al. Built structures influence patterns of energy demand and CO_2 emissions across countries. *Nat. Commun.* **14**, 3898 (2023).
13. Seto, K. C. et al. From low-to net-zero carbon cities: The next global agenda. *Annu. Rev. Environ Resour.* **46**, 377–415 (2021).
14. Newman, P. W. & Kenworthy, J. R. Gasoline consumption and cities: a comparison of us cities with a global survey. *J. Am. Plan. Assoc.* **55**, 24–37 (1989).
15. Brown, M. A., Southworth, F. & Sarzynski, A. *Shrinking the Carbon Footprint of Metropolitan America* (Brookings Institution Washington, DC, 2008).
16. Marshall, J. D. Energy-efficient urban form. *Environ. Sci. Technol*. **9**, 3133–3137 (2008).
17. Kennedy, C. et al. Greenhouse gas emissions from global cities. *Environ. Sci. Technol*. **19**, 7297–7302 (2009).
18. Parshall, L. et al. Modeling energy consumption and co2 emissions at the urban scale: methodological challenges and insights from the united states. *Energy Policy* **38**, 4765–4782 (2010).
19. Weisz, H. & Steinberger, J. K. Reducing energy and material flows in cities. *Curr. Opin. Environ. Sustain.* **2**, 185–192 (2010).
20. Ahlfeldt, G. M. & Pietrostefani, E. The economic effects of density: a synthesis. *J. Urban Econ.* **111**, 93–107 (2019).
21. Anupriya, Graham, D. J., Carbo, J. M., Anderson, R. J. & Bansal, P. Understanding the costs of urban rail transport operations. *Transp. Res. B. Methodol.* **138**, 292–316 (2020).
22. Graham, D. J. Productivity and efficiency in urban railways: parametric and non-parametric estimates. *Transp. Res. E. Logistics Transp. Rev.* **44**, 84–99 (2008).
23. Savage, I. Scale economies in United States rail transit systems. *Transp. Res. A. Policy Practice* **31**, 459–473 (1997).
24. Anupriya, Bansal, P. & Graham, D. J. Congestion in cities: can road capacity expansions provide a solution? *Transp. Res. A. Policy Practice* **174**, 103726 (2023).
25. Loder, A., Ambühl, L., Menendez, M. & Axhausen, K. W. Understanding traffic capacity of urban networks. *Sci. Rep.* **9**, 1–10 (2019).
26. Couture, V., Duranton, G. & Turner, M. A. Speed. *Rev. Econ. Stat.* **100**, 725–739 (2018).

27. Shepherd, R. W. *Theory of Cost and Production Functions* (Princeton University Press, 2015).
28. Baltagi, B. H. & Chang, Y.-J. Incomplete panels: a comparative study of alternative estimators for the unbalanced one-way error component regression model. *J. Econometr.* **62**, 67–89 (1994).
29. Beveren, I. V. Total factor productivity estimation: a practical review. *J. Econ. Surveys* **26**, 98–128 (2012).
30. Katayama, H., Lu, S. & Tybout, J. R. Firm-level productivity studies: illusions and a solution. *Int. J. Industrial Organ.* **27**, 403–413 (2009).
31. Ackerberg, D., Benkard, C. L., Berry, S. & Pakes, A. Econometric tools for analyzing market outcomes. In *Handbook of Econometrics*, (eds. Heckman, J. J. & Leamer, E. E.) 4171–4276 (Amsterdam and Boston: Elsevier, North-Holland, 2007).
32. De-Loecker, J. Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity. *Econometrica* **79**, 1407–1451 (2011).
33. Levinsohn, J. & Petrin, A. Estimating production functions using inputs to control for unobservables. *Rev. Econ. Stud.* **70**, 317–341 (2003).
34. Olley, G. S. & Pakes, A. The dynamics of productivity in the telecommunications equipment industry. *Econometrica* **64**, 1263–1297 (1996).
35. Arellano, M. & Bond, S. Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *Rev. Econ. Stud.* **58**, 277–297 (1991).
36. Arellano, M. & Bover, O. Another look at the instrumental variable estimation of error-components models. *J. Econometr.* **68**, 29–51 (1995).
37. Rubin, D. B. Inference and missing data. *Biometrika* **63**, 581–592 (1976).
38. Imbens, G. W. & Wooldridge, J. M. Recent developments in the econometrics of program evaluation. *J. Econ. Lit.* **47**, 5–86 (2009).
39. Graham, D. J. Causal inference for transport research. *Transp. Res. A. Policy Practice* **192**, 104324 (2025).
40. Graham, D. J. & Glaister, S. The demand for automobile fuel: a survey of elasticities. *J. Transp. Econ. Policy* **36**, 1–25 (2002).
41. US-EPA. *Greenhouse Gas Equivalencies Calculator.* https://www.epa.gov/energy/greenhouse-gas-equivalencies-calculator (2024).
42. Berechman, J. & Giuliano, G. Economies of scale in bus transit: a review of concepts and evidence. *Transportation* **12**, 313–332 (1985).
43. Cowie, J. Acquisition, efficiency and scale economies: an analysis of the british bus industry. *Transp. Rev.* **22**, 147–157 (2002).
44. Graham, D. J., Couto, A., Adeney, W. E. & Glaister, S. Economies of scale and density in urban rail transport: effects on productivity. *Transp. Res. E. Logistics Transp. Rev.* **39**, 443–458 (2003).
45. Farsi, M., Fetz, A. & Filippini, M. Economies of scale and scope in local public transportation. *J. Transp. Econ. Policy* **41**, 345–361 (2007).
46. Woodcock, J., Banister, D., Edwards, P., Prentice, A. M. & Roberts, I. Energy and transport. *Lancet* **370**, 1078–1088 (2007).
47. Cirilli, A. & Veneri, P. Spatial structure and carbon dioxide (Co2) emissions due to commuting: an analysis of Italian urban areas. *Reg. Stud.* **48**, 1993–2005 (2014).
48. Glaeser, E. L. & Kahn, M. E. The greenness of cities: Carbon dioxide emissions and urban development. *J. Urban Econ.* **67**, 404–418 (2010).
49. Hong, J. & Shen, Q. Residential density and transportation emissions: examining the connection by addressing spatial autocorrelation and self-selection. *Transp. Res. Part D. Transp. Environ.* **22**, 75–79 (2013).

Acknowledgements

This study was conducted as part of the Alan Turing Institute's Data-centric Engineering Challenge 1.0, led by Professor Emma McCoy. The authors are grateful to Mr. Richard Anderson and Mr. Alexander Barron from the Transport Strategy Centre for their valuable insights, which enhanced the practical relevance of the study's findings. We also thank the four anonymous reviewers from NPJ Sustainable Mobility and Transport for their constructive feedback, which significantly improved the quality of this manuscript.

Author contributions

Conceptualisation: A., D.J.G. Methodology: A., D.J.G. Investigation: A. Writing—original draft: A. Writing—review and editing: E.M. and D.J.G. Supervision: D.J.G. Project administration: D.J.G. Funding acquisition: E.M. and D.J.G.

Competing interests

The authors declare no competing interests.

Additional information

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