Transportation Research Part D 140 (2025) 104593

Contents lists available at ScienceDirect

Transportation Research Part D

journal homepage: www.elsevier.com/locate/trd

Built environment and travel: Tackling non-linear residential self-selection with double machine learning

Florian Nachtigall [a,b,*], Felix Wagner [a,b], Peter Berrill [d], Felix Creutzig [a,b,c]

[a] *Technical University of Berlin, Berlin, Germany*
[b] *Potsdam Institute for Climate Impact Research (PIK), Potsdam, Germany*
[c] *University of Sussex, Falmer, UK*
[d] *Leiden University, Leiden, Netherlands*

ARTICLE INFO

Keywords:
Urban form
Compact development
Travel behavior
Travel-related CO_2 emissions
Moderating effects
Causal machine learning

ABSTRACT

Understanding how the built environment influences travel is key to low-carbon urban planning. However, previous cross-sectional studies lack a realistic operationalization of residential self-selection that accounts for its non-linear nature, limiting its applicability to urban planning. We propose a double machine learning (DML) approach that accounts for nonlinearities in residential self-selection and captures non-linear moderating effects. Using travel diaries of 32,201 Berlin residents, we estimate the built environment's impact on per capita travel-related CO_2 emissions. Our results indicate that neglecting nonlinearities overestimates this impact by 13%–18%, inflating the built environment proportion by 13%pt. Age, income, and car ownership also nonlinearly moderate the built environment's effect, with the effect being largest for middle-aged, high-income, car-owning households, a novel finding. Applying the method to urban planning reveals a 43%pt emissions reduction potential for 64,000 planned Berlin housing units, highlighting the need for evidence-based urban planning to effectively mitigate CO_2 emissions in cities.

1. Introduction

The built environment plays a critical role in facilitating or hindering the decarbonization of urban transport (Jaramillo et al., 2022). The location and compactness of new residential development shape and constrain residents' future travel behavior. As such, they limit the mitigation potential of mobility-related lifestyle changes and potentially lock in transport-related CO_2 emissions for decades (Creutzig et al., 2016). Urban sprawl and suburban development, in particular, have been widely criticized for increasing car dependence and travel demand, hindering the decarbonization of urban transport (Dieleman and Wegener, 2004). Many studies have observed that residents of urban (central, higher-density, mixed-use) neighborhoods tend to drive less and instead walk (van Holle et al., 2012; McCormack and Shiell, 2011; Barnett et al., 2017; Cerin et al., 2017), bike (van Holle et al., 2012), and use public transit more (Aston et al., 2021; Ibraeva et al., 2020) than residents of suburban (non-central, lower-density, single-use residential) neighborhoods. Yet, it is ambiguous to what extent these differences can be attributed to the built environment itself, as opposed to pre-existing differences in travel preferences between residents of urban and suburban neighborhoods. These differences are a result of a process known as residential self-selection (Cao et al., 2009), in which people choose their place of residence based on, among other things, locally available transport options matching their pre-existing travel preferences. Failing to account for

* Correspondence to: Potsdam Institute for Climate Impact Research (PIK), Member of the Leibniz Association, P.O. Box 60 12 03, D-14412, Potsdam, Germany.

E-mail address: nachtigall@tu-berlin.de (F. Nachtigall).

https://doi.org/10.1016/j.trd.2025.104593
Received 31 July 2024; Received in revised form 19 November 2024; Accepted 3 January 2025

Available online 31 January 2025

residential self-selection and the resulting differences in travel preferences between residents of different neighborhoods can lead to falsely attributing observed differences in travel behavior to the built environment alone, overestimating its impact and drawing biased conclusions (Guan et al., 2020).

Previous studies employed causal inference methods – most frequently statistical control, propensity score matching, sample selection, and structural equation modeling – to isolate the built environment's effect from the confounding influence of residential self-selection by controlling for attitudes and socio-demographics (Mokhtarian and Cao, 2008; Mokhtarian and van Herick, 2016; van Herick and Mokhtarian, 2020). These studies assume a pre-defined, mostly linear influence of attitudes and socio-demographics on residential choice and travel behavior. However, recent reviews argue that this assumption may be overly simplistic and potentially misleading (van Wee and Cao, 2022; Guan et al., 2020). Nonlinear dynamics – such as threshold effects, where only strong preferences significantly impact residential choice – appear to better capture the complexity of this relationship (De Vos and Witlox, 2016; Wolday et al., 2019; Ding et al., 2024). Similar threshold effects have been observed for residential choice based on neighborhood socio-demographics (Galster, 2018) and the influence of attitudes and socio-demographics on travel behavior (Berrill et al., 2024; Ding et al., 2024; Shao et al., 2023; Liu et al., 2021). Despite these insights, the topic received little attention in land use-travel research (Guan et al., 2020). With the exception of Ding et al. (2024), most studies continue to rely on linear models, potentially introducing bias and obscuring the true effect of the built environment.

In this study, we aim to quantify the extent of bias introduced by a linear operationalization of residential self-selection. We explore the potential of the novel causal inference method *double machine learning* (DML) (Chernozhukov et al., 2018) to address this research gap and account for the non-linear confounding influence of travel preferences. DML uses machine learning estimators to model these confounding influences, capturing threshold effects and other nonlinearities that previous studies missed, resulting in a more realistic operationalization of residential self-selection. We compare the estimated built environment effect to that of a linear model specification to quantify the extent of bias. We investigate this with a case study in Berlin, where we analyze cross-sectional travel survey data from 32,201 residents. We evaluate the built environment's impact across multiple, continuous dimensions, specifically the "5Ds" (Ewing and Cervero, 2010) – density, diversity, design, destination accessibility, and distance from transit – and assess the moderating, non-linear influence of socio-demographics and travel-related attitudes. We examine the impact on travel behavior in terms of average per capita travel-related CO_2 emissions to account for changes in both travel distance and mode choice, allowing for a comprehensive assessment of the climate-related impacts of the built environment. To illustrate the applicability of our approach to low-carbon urban planning, we evaluate different urban planning scenarios and currently planned housing projects in terms of their associated transport CO_2 emissions, and identify alternative locations for future housing development that would minimize travel and associated emissions.

Our main contribution is threefold: Using DML allows us to (1) achieve a more realistic operationalization of residential self-selection by accounting for its nonlinearities, (2) gain a more nuanced understanding of effect heterogeneity by accounting for the non-linear moderating effects of socio-demographics and travel-related attitudes, and (3) prospectively estimate the built environment's impact for new housing development and infrastructure changes to inform urban planning. Our research questions are the following:

1. How is the built environment effect biased if nonlinearities in residential self-selection are not considered?
2. How does accounting for nonlinearities in the moderating influence of socio-demographics and travel-related attitudes improve the understanding of who is most affected by the built environment?
3. What are the induced transport CO_2 emissions of currently planned housing projects in Berlin, and where should new housing be located to minimize travel?

2. Literature review

2.1. Built environment

Characteristics of the built environment that influence travel behavior have been conceptualized by the "3Ds": density, diversity, and design (Cervero and Kockelman, 1997), later complemented by destination accessibility and distance to transit ("5Ds") (Ewing et al., 2009; Ewing and Cervero, 2010). Systematic reviews and meta-analyses consistently (Ewing and Cervero, 2010; Saelens and Handy, 2008; Stevens, 2017; Javaid et al., 2020; Aston et al., 2021; McCormack and Shiell, 2011) found that high destination accessibility and residential density, diverse land uses, well connected street networks, and accessible public transport are positively associated with sustainable transportation. A mix of different land uses shortens travel distances and reduces the need for motorized travel by bringing jobs and other destinations closer to residential locations (Duncan et al., 2010). High residential densities are needed to support transit services and the presence of retail and thus increase walkability of an area (Transportation Research Board and Institute of Medicine, 2005). A high street connectivity improves the directness of travel and proximity of destinations (Oakes et al., 2007), particularly benefiting active travel (Saelens et al., 2003), as it is very sensitive to travel distance (Liu et al., 2021). In addition to enabling sustainable transport, compact development is associated with further co-benefits such as reduced residential energy consumption, increased physical well-being, improved road safety, and increased social interaction, mobility and capital (Ewing and Cervero, 2017).

The literature differs on which dimension of the built environment is most influential on travel behavior. Systematic reviews and meta-analyses suggest that the influence is heterogeneous for different modes and travel purposes, with destination accessibility being most decisive for driving distance, followed by density and design (Ewing and Cervero, 2010; Stevens, 2017) ; design, diversity, and distance to transit are more relevant for public transit use (Javaid et al., 2020; Aston et al., 2021); while design, diversity, and density of the local neighborhood are most important for active travel (McCormack and Shiell, 2011). Altogether, for travel-related CO_2 emissions, recent case studies find destination accessibility (Liu et al., 2022; Shao et al., 2023; Ao et al., 2019; Zhang et al., 2022) or diversity (Wu et al., 2019; Cui et al., 2023) to be most important. In the absence of an up-to-date meta-analysis, it is difficult to assess whether the different results of case studies are due to geographic and cultural differences or to differences in methods and model specifications.

Recent studies have also shown that the influence of the built environment on travel behavior is non-linear (Cao and Tao, 2023; Aghaabbasi and Chalermpong, 2023), limiting previous findings to some extent. For example, density has threshold effects in that sufficient density is required to make public transit viable in a neighborhood (e.g. Guerra and Cervero, 2011). The effect of increasing density becomes weaker and possibly insignificant after a certain level of population density is reached (Hong, 2017; Ding et al., 2018). Similar non-linear effects have been found for other dimensions of the built environment. From a climate change perspective, one of the key findings was that new development should be concentrated within a 12 km radius of the city center to avoid a disproportionate increase in driving distances (Ding et al., 2018). The nonlinearity was also studied for other dimensions of travel behavior, e.g. active travel (Liu et al., 2021), elderly's walking (Cheng et al., 2020), walking to transit (Tao et al., 2020), metro ridership (Ding et al., 2019), ridesplitting (Tu et al., 2021), bike-sharing (Hu et al., 2024), car ownership (Berrill et al., 2024), commuting distance (Wagner et al., 2022), and travel-related CO_2 emissions (Wu et al., 2019).

Conceptualizing the built environment using the D variables is not uncriticized (e.g. Handy, 2018), and recent studies have partially moved toward accessibility-based measures (Handy, 2020) as the underlying mechanism of how the Ds affect travel behavior. For example, density enables local amenities and transit services, and its impact on travel is manifested through the improved accessibility to local destinations and transit. However, from an urban planning perspective, the original 3Ds are still interesting because they can be directly influenced by policies, for example through zoning regulations, building height restrictions, and compact development.

2.2. Residential self-selection bias

While many studies have explored the relationship between the built environment and travel behavior, the strength of these relationships is still contested, particularly the extent to which it is confounded by residential self-selection (Guan et al., 2020). Residential self-selection (RSS) describes a selection bias resulting from households choosing their residential location based on pre-existing preferences, attitudes, and needs that also influence their travel behavior (Cao et al., 2009). It recognizes that people who choose to live in certain areas may have inherently different travel preferences. Previous work identified socio-demographic traits and travel-related attitudes as the main drivers of residential self-selection (Mokhtarian and Cao, 2008). The mechanism that induces the residential self-selection bias is depicted in Fig. 1. Failing to account for residential self-selection can lead to falsely attributing observed differences in travel behavior solely to the built environment, overestimating its impact and drawing biased conclusions (Guan et al., 2020). For example, high cycling rates in bike-friendly neighborhoods may not be entirely due to existing residents taking up cycling, but to cycling enthusiasts moving into the area. Controlling for residential self-selection allows to isolate the true effects of the built environment on travel behavior. For the given example, the effect of bike-friendly infrastructure can be isolated by looking at the average increase in cycling among residents who moved there for reasons other than good cycling infrastructure.

Previous studies typically operationalize the confounding influence of residential self-selection using linear models. However, the influence of travel-related attitudes and socio-demographics on residential choice and travel behavior (see B & C in Fig. 1) is potentially non-linear and characterized by threshold effects (van Wee and Cao, 2022; Guan et al., 2020). Weak travel preferences may have no effect on residential choice, while strong preferences may have a dominant influence (De Vos and Witlox, 2016; Wolday et al., 2019). For example, De Vos and Witlox (2016) found that travel haters have strong preferences and self-select into urban neighborhoods, while travel lovers have no residential preferences. Similarly, Wolday et al. (2019) found that transit lovers have a strong preference of urban neighborhoods, whereas car lovers have no residential preferences. Galster (2018) observed similar threshold effects for residential choice based on neighborhood socio-demographics. Non-linear patterns are also evident in the relationship with travel behavior. Many machine learning approaches that examine the non-linear influence of the built environment on travel behavior also examine the non-linear influence of household characteristics on travel behavior (e.g. Shao et al., 2023; Berrill et al., 2024; Wagner et al., 2022; Xu et al., 2021; Liu et al., 2021). For example, Berrill et al. (2024) showed the prevalence of nonlinearities in the influence of age and income on travel across many European cities. Despite these broad findings from qualitative (e.g. Wolday et al., 2019) and quantitative studies (e.g. De Vos and Witlox, 2016) – spanning the traditional planning (e.g. Galster, 2018) and travel behavior literature (e.g. Lucas et al., 2016), as well as more recent influences from machine learning-based modeling (e.g. Berrill et al., 2024) – almost all previous studies assumed pre-defined, mostly linear relationships, resulting in an oversimplified operationalization of residential self-selection and potentially biased effect estimates. To our knowledge, only Ding et al. (2024) have accounted for the nonlinearity in a case study in Jinan, China, finding threshold effects for environmental attitudes, car ownership, and income, among others. To address this gap, van Wee and Cao (2022) recently proposed using gradient boosting decision trees to model these nonlinearities in travel preferences.

Quantitative estimates of the proportion of observed differences in travel behavior across neighborhoods that can be attributed to residential self-selection very strongly across case studies (van Herick and Mokhtarian, 2020; Mokhtarian and van Herick, 2016). Empirical comparisons are complicated by the fact that studies differ in the way they calculate the proportion (van Herick and Mokhtarian, 2020). A review of empirical findings by Mokhtarian and van Herick (2016) with a focus on propensity score and

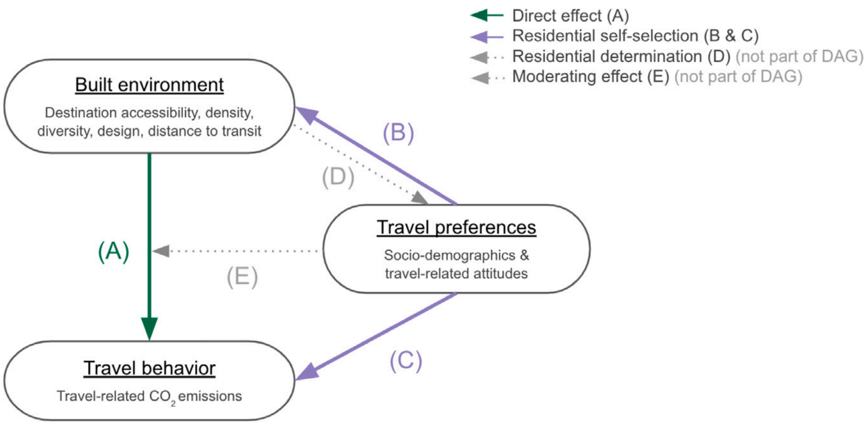

Fig. 1. Travel preferences confound the effect of the built environment on travel behavior. Graphical representation of the causal relationship between the built environment, travel preferences, and travel behavior. The gray text indicates the attributes used to measure the constructs in this study. (A) The neighborhood built environment influences residents' travel behavior by modulating mode choice and travel distances. (B) Travel preferences influence the choice of residential location and thus the surrounding built environment. (C) Travel preferences influence travel behavior and (E) moderate the influence of the built environment on travel behavior. (D) Travel preferences can also be influenced by the built environment itself, e.g. through transport cultures. (B) and (C) together induce the residential self-selection bias. (A), (B), and (C) form the directed acyclic graph (DAG) as used in this study. Specifically, we estimate the effect of (A) while accounting for the non-linear confounding influence of (B) and (C) and the moderating influence of (E).

sample selection approaches finds estimates ranging from 2% (Cao et al., 2010) to 66% (Cao and Fan, 2012). Among models with a high goodness-of-fit, the average proportion is 38% (van Herick and Mokhtarian, 2020). Yet, the proportion also varies between different dimensions of travel behavior, e.g. Cao and Fan (2012) estimate the proportion to be 66% for driving time and 28% for distance traveled. While machine learning approaches cannot quantify the proportion of the apparent built environment effect attributable to residential self-selection, they can quantify the combined impact of socio-demographics and travel-related attitudes on travel behavior. Studies have found this impact to be between 31% (Tao et al., 2020) and 69% (Liu et al., 2021) for active travel, 50% for mode choice (Cheng et al., 2019), and 35% for car commuting (Ding et al., 2018) in their respective local settings. The extent to which residential self-selection is driven by attitudes in addition to socio-demographics remains disputed. While many studies attempt to show its importance quantitatively (Mokhtarian and van Herick, 2016), others argue conceptually that its relevance is overstated (e.g. Næss, 2014).

2.3. *Operationalizing travel preferences: confounder, moderator, or mediator?*

More recent studies (van de Coevering et al., 2016; Scheiner, 2018; Kroesen, 2019; van de Coevering et al., 2018; Lin et al., 2017; Guan et al., 2020) have challenged the notion of travel preferences being exogenous and instead suggested that they may have a mediating and moderating effect on the influence of the built environment on travel behavior.

Kroesen (2019) and van de Coevering et al. (2018) provide evidence that travel attitudes itself are affected by the built environment and past travel behavior, a process sometimes referred to as reverse causation (van de Coevering et al., 2018; Kroesen, 2019) or residential determination (Lin et al., 2017) (see D in Fig. 1). This mediating effect can be explained by transport cultures (Klinger and Lanzendorf, 2016), availability of travel options (van de Coevering et al., 2016; Bamberg, 2006), and post-hoc rationalization to reduce cognitive dissonance (Tertoolen et al., 1998; Chatman, 2009), among others. For example, Tertoolen et al. (1998) find that to reduce cognitive dissonance households tend to reduce their pro-environmental attitudes rather than adjust their driving behavior when confronted with the environmental impacts of their suburban-style driving. Similarly, Bamberg (2006) and van de Coevering et al. (2016) find that people tend to adjust their attitudes in favor of public transit after moving to an area with better transit accessibility. In addition to attitudes, car ownership strongly mediates the influence of the built environment on travel behavior (Ding et al., 2017; van Acker and Witlox, 2010). Failing to account for mediating effects may lead to an overestimation of the influence of residential self-selection and an underestimation of the influence of the built environment on travel behavior (Lin et al., 2017; Næss, 2009).

Travel attitudes furthermore moderate the influence of the built environment on travel behavior (e.g. good transit accessibility may have a lower impact on car-lovers Guan et al., 2020) (see E in Fig. 1). Socio-demographics also have a moderating influence, in several respects: They can strengthen or weaken the effect of the built environment on travel behavior (e.g. young adults may prioritize active travel almost independently of neighborhood walkability Christiansen et al., 2014), the effect of travel attitudes on travel behavior (e.g. low-income households may be not able to realize their preference for driving Næss, 2014), and on attitude-based residential self-selection (e.g. low-income households may not be able to afford housing in a transit-oriented neighborhood Wolday et al., 2018). Failing to consider these interaction effects may mask important equity issues and drivers of heterogeneity in the built environment effect. While they have been well studied for active travel from a health perspective, suggesting that the relationship is stronger for non-senior (Richardson et al., 2017), high-income (Manaugh and El-Geneidy, 2011), highly-educated (Owen et al., 2007) households, they have received less attention in low-carbon urban planning. As a result, there is limited evidence on moderation effects for travel-related CO_2 emissions.

2.4. Methods and study designs

Methodologically, two strands have emerged in the literature to quantify the built environment's influence on travel behavior. First, a more recent development using machine learning methods to better capture the non-linearity of the effect (Cao and Tao, 2023; Aghaabbasi and Chalermpong, 2023). Mainly (extreme) gradient boosting regression trees (e.g. Ding et al., 2018) and random forests are used (e.g. Cheng et al., 2020). These machine learning approaches often examine the effect using SHAP values (Lundberg and Lee, 2017) or partial dependence plots, which allow for spatially explicit estimates of each built environment's dimension on travel behavior. However, they can only detect associations, not causal relationships, because they neglect the influence of travel preferences on residential choice (see B in Fig. 1) and the interdependencies between built environment characteristics. For urban planning decisions, however, such causal insights are indispensable.

In contrast, the second strand aims at uncovering causal relationships by using statistical methods to disentangle the direct influence from the confounding effect of residential self-selection. As panel data are scarce, most studies use cross-sectional study designs (Cao et al., 2007a), weakening the causal explanatory power (van de Coevering et al., 2015). Mainly statistical control (e.g. Kitamura et al., 1997), propensity score matching (e.g. Cao et al., 2010), sample selection (e.g. Zhou and Kockelman, 2008), structural equation (e.g. Bagley and Mokhtarian, 2002), and instrumental variable models (e.g. Heres-Del-Valle and Niemeier, 2011) have been used. Yet, constrained by linearity assumptions and often ignoring moderating factors (Guan et al., 2020), these approaches cannot fully capture the true impact either. To our knowledge, only Ding et al. (2024) have made initial attempts to account for the nonlinearity of confounding factors when estimating the impact of bus rapid transit areas using DML.

To strengthen the causal explanatory power, some studies use repeated cross-sectional data (e.g. Dai et al., 2020) or construct pseudo-panels (e.g. Cao et al., 2007b). To account for non-linear relationships, a notable recent study by Zhou et al. (2024) also employed a DML approach, specifically an extension of the DML framework to difference-in-differences modeling, as developed by Chang (2020). Despite such methodological advances, multiple reviews advocate for true longitudinal study designs as the most effective means of addressing residential self-selection bias (e.g. van de Coevering et al., 2015; Cao et al., 2009; Guan et al., 2020). Existing longitudinal studies mostly employ structural equation modeling (e.g. Wang and Lin, 2019), linear or binomial regressions (e.g. Kamruzzaman et al., 2016), difference-in-differences techniques (e.g. Deng and Zhao, 2022), or mixed-methods approaches (e.g. Scheiner et al., 2024). However, these studies remain rare and are typically very localized with small sample sizes (van de Coevering et al., 2015), limiting their potential for the large-scale assessment of low-carbon urban planning.

In summary, using the causal graph in Fig. 1 as reference, statistical methods typically estimate the linear effect of (A) while accounting for (B) and (C). DML similarly estimates the effect of (A), but accounts for the nonlinearity of (B) and (C). ML estimates the non-linear effect of (A) and/or (C), but completely ignores (B). Recent studies focusing on travel-related CO_2 emissions predominantly use structural equation modeling (Zhang et al., 2022; Ao et al., 2019) and gradient boosting regression trees (Shao et al., 2023; Liu et al., 2022; Wu et al., 2019). DML has not been applied here.

2.5. Unit of analysis

Previous studies have primarily chosen two units of analysis to examine the impact of the built environment on travel behavior: individuals and neighborhoods. Studies that control for residential self-selection mostly look at individuals or households to directly incorporate their socio-demographic traits and travel-related attitudes as confounding factors (Mokhtarian and van Herick, 2016). Recent studies, however, have recommended conducting the analysis of residential self-selection at the household level to address the issue of intra-household interactions (Guan and Wang, 2019; Yang et al., 2019). For both individual- and household-level studies, the built environment is then characterized based on the neighborhood (e.g. zip code or tract) or using a buffer around the residential or travel location. Other studies directly aggregate travel behavior within a spatial unit and look at the built environment of distinct, non-overlapping neighborhoods (e.g. Wagner et al., 2022; Hankey and Marshall, 2010) or cities (e.g. Ewing et al., 2018). In all cases, the spatial definition of a neighborhood or buffer size affects the outcome of the spatial analysis (Zhang and Kukadia, 2005), also known as the Modifiable Areal Unit Problem (MAUP) (Fotheringham and Wong, 1991). For example, aggregating data may hide, smooth, or distort spatial heterogeneities and nuances, leading to biased estimates of neighborhood effects. Therefore, all insights are bound to the definition of spatial units. Another limitation of existing studies, especially disaggregated ones, is that spatial spillover effects between neighborhoods are often not comprehensively accounted for Kolak and Anselin (2020) and Credit and Lehnert (2023). For example, good cycling infrastructure in a nearby neighborhood can attract cyclists and increase the overall cycling mode share in surrounding areas. As Ewing et al. (2018) point out, the choice of unit of analysis is also a matter of research focus, with disaggregated analyses better suited to evaluating local and short-term policies, and aggregated city-level analyses better suited to long-term policy aspirations.

2.6. Goodness-of-fit

When performing analysis on an individual or household level the goodness-of-fit of the respective models tends to be low (van Herick and Mokhtarian, 2020), with a coefficient of determination R^2 of 0.06 – 0.19 using classical statistical approaches (van Herick and Mokhtarian, 2020) and 0.22 (Liu et al., 2022) – 0.30 (Wu et al., 2019) using machine learning.[1] While it is inherently difficult to model travel behavior of a brief snapshot – given the many unobserved, highly-personal, and volatile factors that influence travel decisions during the observed period – the limited overall explanatory power translates into a high degree of uncertainty in estimating the impact of the built environment (van Herick and Mokhtarian, 2020). In contrast, studies that examine aggregate travel behavior in neighborhoods or tracts tend to achieve higher R^2, e.g. 0.60 for car commuting distance (Wagner et al., 2022).

3. Methods

Our goal is to obtain an unbiased estimate of the travel-related neighborhood specific CO_2 emissions that can solely be attributed to the built environment. Our starting point is observed differences in travel-related CO_2 emissions according to surveyed trip data. However, these differences are biased estimates of the true impact of the built environment because they do not account for individuals' pre-existing travel preferences that influenced their residential choices (residential self-selection). Here, we use DML to control for these confounding effects and estimate the direct effect of the built environment.

3.1. Data & preprocessing

Study area. Berlin, the capital of Germany, has about 3.8 million inhabitants on 890 km^2, a population density of on average 4300 people per km^2 and up to 25,000 people per km^2 in urban neighborhoods (see Fig. 2). It is growing, with nearly 200,000 housing units announced to be built by 2030 (SenStadt, 2023). It is a relatively monocentric city, with an urban centrality index of 0.38. For context, values closer to 1 indicate higher monocentricity, while values closer to 0 indicate greater polycentricity. By comparison, Paris has an index of 0.51, São Paulo 0.20, and Los Angeles 0.05 (Pereira et al., 2013). Berlin has a good public transport network and a built environment that facilitates active travel. About 78% of residents live within 500 m of a public transport stop with service at least every 10 min, ranking 14th out of 1095 cities globally (ITDP, 2024). High mixed-use development provides very good local accessibility, with 86% of residents living within 1 km of both healthcare and education services, ranking 14th globally. In terms of protected bike lane coverage, Berlin ranks 22nd. Car traffic is comparatively slow in the Berlin metro area, ranking 47th out of 387 cities in the TomTom Traffic Index (TomTom, 2024). While the city center has a low-emission zone, driving is generally permitted throughout Berlin with no city tolls or congestion pricing. Parking is either free with a residential parking permit or inexpensive compared to other major capitals. The implementation of car-free streets and low-traffic neighborhoods is currently a contentious issue in Berlin.

Overall, despite the limited policies to systematically reduce car usage, Berlin's built environment supports sustainable transport modes relatively well. This results in the lowest household car ownership rate in Germany (56.7%) and high shares of active mobility (47.2% of trips) and transit use (42.5% of mileage) (Gerike et al., 2020).

Data sources. Representative workday travel behavior for all 190 zip code areas in Berlin is obtained from travel diaries of the German national mobility survey "Mobilität in Städten — SrV 2018" (Hubrich et al., 2019), which includes about 118,000 trips by 32,201 residents from 15,484 Berlin households on midweek days. To describe the built environment, we use only publicly available data from OpenStreetMap (OpenStreetMap contributors, 2017) on the street network and points of interest, from the public transit operator "Verkehrsverbund Berlin-Brandenburg" on transit accessibility (GTFS data) (VBB, 2023) and from Berlin's open data portal (Open Data Berlin, 2023) on population density, land-use, and election results. Lastly, to convert mode-specific travel kilometers to CO_2 emissions, we use emissions factors from the International Transport Forum (Cazzola and Crist, 2020).

Target variable. To comprehensively examine the climate-relevant impact of the built environment on travel, taking into account both a reduction in travel distances and a shift to sustainable modes, we evaluate travel behavior in terms of its induced CO_2 emissions. This allows to address the interdependency between travels by different modes. Similar to previous studies (Shao et al., 2023; Liu et al., 2022; Wu et al., 2019; Zhang et al., 2022; Ao et al., 2019), we convert mode-specific travel distances to CO_2 emissions using emission factors (see Appendix Table 7) and aggregate the trips of respondents across all modes to determine their total daily travel-related CO_2 emissions on a working day. We conduct the analysis at the household level to account for intra-household interactions, dependencies, and decision-making power structures (Yang et al., 2019; Guan and Wang, 2019). Therefore, we calculate the average per capita emissions for each household by aggregating the travel-related CO_2 emissions of all household members and dividing by the number of household members to facilitate comparability across households.

Neighborhood types. To examine how travel behavior and the influence of the built environment differ between neighborhood types, such as urban and suburban, we need to classify neighborhoods accordingly. We use k-means clustering to group neighborhoods with similar built environment characteristics (see Table 2), i.e. destination accessibility, density, diversity, design, accessibility to transit (see Appendix A.1).

[1] Only considering studies on overall travel-related CO_2 emissions, as the goodness-of-fit of machine learning approaches differs substantially across dimensions of travel behavior (Aghaabbasi and Chalermpong, 2023).

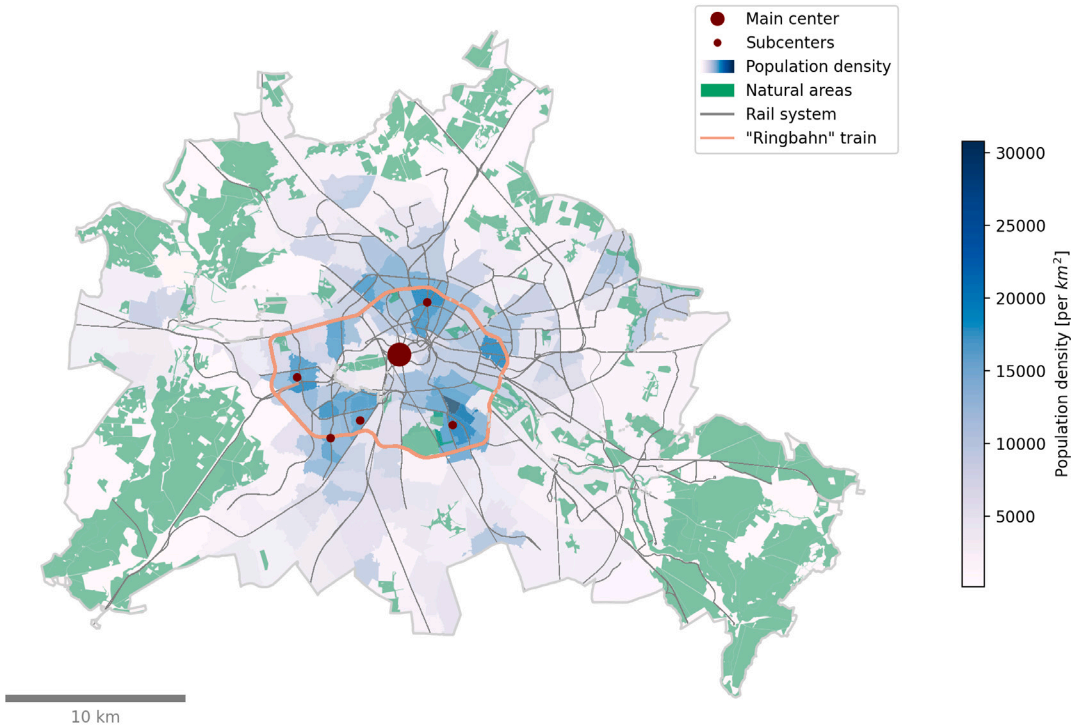

Fig. 2. Built environment of Berlin, Germany. The center and subcenters, based on points of interest density, are indicated as dark red circles. Population density of 190 neighborhoods is color coded in blue. Natural areas according to Berlin land use data (Open Data Berlin, 2023) are marked in green. The public transportation rail network is drawn in gray, with the exception of the so-called "Ringbahn", an urban rail line that circles central Berlin, which is highlighted in orange. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.2. Conceptualization

Confounding effects. Travel preferences confound the influence of the built environment on travel behavior because of residential self-selection. The relationship is visualized as a directed acyclic graph (DAG) in Fig. 1. The confounding effect can be accounted for by controlling for socio-demographic traits and travel-related attitudes (Mokhtarian and Cao, 2008). In terms of socio-demographics, we use information on age, income, education, and household size from the mobility survey. Because travel-related attitudes are not directly captured in the survey, we use information on ownership of transport means as proxies, specifically car, bike, driving license, and transit subscription ownership. While this is not a comprehensive characterization of travel-related attitudes, it has the advantage of capturing the temporal lag of travel-related attitudes through vehicle ownership, particularly car ownership (van Acker et al., 2014), that not only locks in travel behavior but also influences residential choice (Lin et al., 2017; Scheiner, 2014). Complementary to transport ownership, we capture general attitudes towards sustainable mobility and other environmental issues according to recent election results where the provision of sustainable transport infrastructure was in the public focus. Refer to Table 1 for a brief description of all attributes used to control for residential self-selection.

Treatment encoding. To examine how the exposure to a particular built environment influences travel behavior, we first determine how the built environment differs from the rest of the city. We characterize the difference in terms of the 5Ds, density, diversity, design, destination accessibility, and distance from transit (Ewing and Cervero, 2010). For each D, we develop at least one feature, guided by previous studies (see Table 2). For destination accessibility, we develop several characteristics to distinguish between local and regional accessibility and account for the mode-specific destination accessibility that the D variables have been criticized for neglecting (Handy, 2018). The differences in the 5Ds at the residential location from the city average represent the degree of treatment. In other words, we estimate a counterfactual world in which the built environment is homogeneous across the city and assess how this would change travel behavior. Since the built environment of a given neighborhood can influence travel behavior in neighboring neighborhoods, we include spatial lag variables to capture these spillover effects, similar to a spatially-lagged X model (SLX) (Halleck Vega and Elhorst, 2015). We transform the population, points of interest, and intersection density features using a natural log function ($ln(1 + x)$) to capture sublinear effect of increasing density despite using a partial linear treatment model. See Appendix A.1 for more details on how we calculate the individual built environment characteristics and Appendix A.3 how we mitigate spatial spillover effects.

Table 1
Travel preferences. Overview of all socio-demographic traits and proxies for travel-related attitudes considered in the study including a brief description. Children are excluded from the calculation of average age, education, employment status, and driver's license ownership to reduce the correlation between the variables and ensure that each variable describes a separate aspect.

Category	Variable name	Description
Socio-demographics	Income	Household income
	Age	Average age of adult (>18 years) household members
	Higher education	Share of household members (>25 years) with university degree
	Employment status	Share household members (>18 years) working part- or full-time
	Mobility constraints	Share household members with mobility impairments
	Household size	Number of persons living in a household
	Household type	Distinction between single households over and under 65 and family households
Proxies for travel-related attitudes	Car ownership	Average number of private & company cars owned per person
	Bike ownership	Average number of bicycles owned per person
	Driving license	Share household members (>18 years) with driving license
	Transit subscription	Share of household members with monthly transit subscription (incl. children and people with disabilities with free ride tickets)
	Environmental policy preferences	Electoral share of the Green Party in constituencies intersecting the neighborhood in the last regional elections

Table 2
Built environment characteristics. Overview of all built environment characteristics considered in the study including a brief description and mapping to the corresponding D of compact development. The street network and points of interest (POIs) are extracted from OpenStreetMap. Shortest and fastest routes are determined using OSMnx (Boeing, 2017). Accessibility to transit is based on the survey or calculated using GTFS feeds from the local transit operator VBB, accessibility to destinations by transit using GoogleMaps. Density measures are calculated using the built-up area only. See Appendix A.1 for more details on feature engineering of built environment characteristics.

5D's of compact development	Feature name	Description
Destination accessibility	Distance to center	Distance of the fastest car route to the central business district
	Distance to subcenter	Shortest distance to one of the 5 areas with the highest POI density via the fastest car route
	Transit travel time	Average travel time to POIs by public transit
	POI density	Local POI density of retail, healthcare, second and third places
Density	Population density	Population density in the built-up area
Diversity	Mixed-use share	Share of mixed-use areas
Design	Street connectivity	Intersection density in the built-up area (proxy for walkability)
Distance to transit	Walking time to rail	Walking time to the nearest rail-based public transport

3.3. Model specification

Given potential non-parametric confounding effects, non-parametric moderating effects, and a continuous treatment along multiple treatment dimensions, we choose a Causal Forest-based DML approach to estimate the effect of the built environment on travel behavior.

Fundamentals. The DML framework is based on the Frisch–Waugh–Lovell (FWL) theorem, which shows that a regression coefficient of interest can also be obtained sequentially by separating the regression into two stages: First, the independent variable of interest and the dependent variable are regressed on the other control variables; second, the residuals of these regressions are regressed on each other, yielding the same partial regression coefficient as in the multiple regression approach. More formally, suppose we have a multiple regression model with q independent variables:

$$\hat{Y} = \hat{\beta}_1 X_1 + \cdots + \hat{\beta}_q X_q + \epsilon \tag{1}$$

The FWL theorem states that any regression coefficient $\hat{\beta}_j$ for $j \in \{1, 2, \ldots, q\}$ is equal to the corresponding regression coefficient $\hat{\beta}_j^*$ in

$$\hat{\tilde{Y}} = \hat{\beta}_j^* \tilde{X}_j + \epsilon^*, \tag{2}$$

where $\hat{\tilde{Y}}$ is the residualized $\hat{Y}$ and $\tilde{X}_j$ is the residualized X_j and $\epsilon = \epsilon^*$, meaning the same unobservable stochastic disturbances.

This sequentialization has the advantage that any method can be used for the first stage of residualization, including machine learning methods, which have the advantage of not having to make any parametric assumptions about the functional form. DML leverages this for causal effect estimation: The outcome and treatment are being predicted from the controls using any appropriate machine learning model, called a nuisance model. The treatment residuals are then used to fit the outcome residuals, yielding a debiased treatment effect estimate. The controls are also included to account for any moderating effects in explaining the heterogeneity of the treatment effect.

Formal model specification. For a partially linear model, the following structural equation assumptions can be made on the data generating process:

$$Y = \theta_0 D_0 + g_0(X) + \epsilon \quad \mathbb{E}[\epsilon | X, D] = 0 \tag{3}$$

$$D = m_0(X) + \zeta \quad \mathbb{E}[\zeta|X] = 0 \tag{4}$$

where Y is the outcome (e.g. per capita travel-related CO_2 emissions), D is the treatment (e.g. residential built environment density), X are the covariates (e.g. travel preferences), θ_0 is the causal parameter of interest, $g_0(X)$ and $m_0(X)$ are the nuisance functions describing the confounding effect of X on the outcome and treatment, respectively, $\mathbb{E}[\cdot]$ is an empirical expectation function, and ζ and ϵ are stochastic disturbances. No further structural assumptions have to be made about $g_0(X)$ and $m_0(X)$.

A naive approach to estimating the causal parameter θ_0 by combining machine learning estimation of g_0 with least squares regression of Eq. (3) would lead to two problems. First, regularization in machine learning introduces a bias in $\hat{g}_0$, causing the estimator to converge slower than $\frac{1}{\sqrt{n}}$ for a sample size n, with $\mathbb{E}[\hat{g}_0] \neq g_0$, and a diverging bias with increasing number of covariates p, namely $|\sqrt{n}(\hat{\theta}_0 - \theta_0)| \xrightarrow{p} \infty$. Second, because $\hat{g}_0$ was fit to the data $W = (Y, D, X)$, it is not statistically independent of W, leading to an overfitting bias when using W to estimate θ_0, specifically $\mathbb{E}[\partial_g \psi(W; \theta_0, g_0) \cdot (\hat{g}_0 - \mathbb{E}[\hat{g}_0])] \neq 0$. See Chernozhukov et al. (2018) for details.

Orthogonalization & cross-fitting. Chernozhukov et al. suggest using orthogonalization to mitigate the regularization bias, and cross-fitting to mitigate the overfitting bias. Orthogonalization ensures that the estimation of θ_0 remains robust to small errors in the nuisance function estimates $\eta_0 = (g_0, m_0)$ by constructing a score function $\psi(W; \theta_0, \eta_0)$ that satisfies the Neyman orthogonality condition $\partial_\eta \mathbb{E}[\psi(W; \theta_0, \eta_0)][\eta - \eta_0] = 0$. This robustness can be achieved by partialling out the effect of X from Y and D, effectively removing the confounding influence and isolating the exogenous variation in D:

$$\tilde{D} = D - \hat{m}_0(X), \quad \tilde{Y} = Y - \hat{g}_0(X) \tag{5}$$

To mitigate overfitting bias, cross-fitting divides the data into K folds, fitting nuisance functions to out-of-fold samples, yielding $\hat{\eta}_{0,-k}$, and computing residuals $\tilde{Y}_k$ and $\tilde{D}_k$ for the held-out fold $k \in K$. Using the generalized method of moments, $\check{\theta}_0$ can be consistently estimated by solving the moment condition of the empirical expectations of the different folds:

$$\frac{1}{K} \sum_{k \in K} \mathbb{E}_k[\psi(W_k; \check{\theta}_0, \hat{\eta}_{0,-k})] = 0 \tag{6}$$

See Appendix B for a more detailed definition of orthogonalization and cross-fitting, including an exemplary score function and solved estimator for θ_0.

Estimation approach by EconML (Microsoft Research, 2019). The structural equations in (3) can be re-written using the residualized treatment and outcome as $\tilde{Y} = \theta_0 \tilde{D} + \epsilon$. When assuming $\mathbb{E}[\zeta \cdot \epsilon | X] = 0$, meaning that the error term ϵ is uncorrelated with the residualized treatment $\tilde{D}$, conditional on X, we can estimate θ_0 independently of the error term ϵ by minimizing the squared difference between the residualized outcome $\tilde{Y}$ and the predicted outcome $\check{\theta}_0 \tilde{D}$:

$$\check{\theta}_0 = \arg\min_{\theta} \mathbb{E}\left[(\tilde{Y} - \check{\theta}_0 \tilde{D})^2\right] \tag{7}$$

Intuitively, we search for a function that best explains the variation in the outcome that other covariates could not explain using the exogenous variation in the treatment ζ, which we have approximated using orthogonalization, i.e. by partialling out the effect of other covariates.

Conditional treatment effects. We extend the previously introduced double machine learning framework of Chernozhukov et al. (2018) to estimate a constant treatment effect θ in a partially linear model by allowing the treatment effect θ to be conditional on X and employing a Causal Forest (Athey et al., 2018) to model these moderating effects nonparametrically. The estimation of the conditional average treatment effect (CATE) $\theta_0(x)$ for sample $x \in X$ is based on a set of residualized local moment conditions

$$\check{\theta}_0(x) = \arg\min_{\theta} \mathbb{E}\left[K_x(X) \cdot \left(\tilde{Y} - \check{\theta}_0(X) \cdot \tilde{D}\right)^2\right], \tag{8}$$

where the kernel $K_x(X_i)$ is a similarity metric that is calculated by constructing a variant of random forests incorporating "honesty" (Athey et al., 2018) and sub-sampling techniques. Using sample averages and the fitted nuisance functions $\hat{\eta}_0$, we can solve the moment equation and estimate $\check{\theta}_0(x)$:

$$\check{\theta}_0(x) = \arg\min_{\theta} \sum_{i=1}^{n} K_x(X_i) \cdot \left(Y_i - \hat{g}_0(X_i) - \check{\theta}_0(X_i) \cdot (D_i - \hat{m}_0(X_i))\right)^2 \tag{9}$$

Model implementation & hyperparameters. We use the open source implementation of the double machine learning framework provided by the EconML package (Microsoft Research, 2019). For estimating the nuisance functions η_0, we use gradient boosting decision trees, i.e. the XGBoost implementation (Chen and Guestrin, 2016). We perform a multi-output regression to predict the multiple treatment dimensions simultaneously, taking into account their interdependencies (Xu et al., 2019). We choose hyperparameters based on a random search with 5-fold cross validation, resulting in 600 tree estimators, a tree depth of 4, a learning rate of 0.01, a subsample of 0.8, and a fraction of features used per tree of 0.9 for the treatment nuisance model, and for the outcome nuisance model 400 tree estimators, a tree depth of 5, a learning rate of 0.01, a subsample of 0.6, and a fraction of features used per tree of 0.7. For the final estimation of θ_0, we use a Causal Forest (Athey et al., 2018) with 100 trees to estimate the non-parametric heterogeneous treatment effect by minimizing the mean squared error of a linear moment estimation tree. Hyperparameter optimization based on the out-of-sample R-score performance yields a maximum tree depth of 3, a minimum variance fraction per leaf of 0.01, and a minimum weighted fraction of the sum of weights per leaf node of 0.01. Compared to the default hyperparameters, this reduces the complexity of the model and makes it more conservative in modeling effect heterogeneity, preventing overfitting to the noise in the data. We perform 5-fold spatial cross-validation at the neighborhood level when estimating nuisance parameters and treatment effect to account for the inherent spatial structure of the data and thus avoid overestimating predictive performance (Roberts et al., 2017).

Model assumptions. To obtain a reliable treatment estimate from observational data using causal inference methods, several assumptions must hold, most importantly unconfoundedness, ignorability, positivity, and the Stable Unit Treatment Value Assumption (SUTVA). Unconfoundedness states that all factors that simultaneously influence treatment assignment and outcome are included in the covariates leaving no confounders unobserved and thus eliminating the possibility of confounding bias. Here, this means that we observe all socio-demographics and attitudes that influence both residential choice and travel decisions. Ignorability states that the treatment assignment is independent of potential outcomes given the observed covariates. As such, units could be swapped between the treatment and control group without affecting the experiment (exchangeability). Including bad controls, i.e. variables that are not exogenous but influenced by the built environment, such as car ownership, violates the assumption and biases the effect estimate. Positivity states that across the range of covariates there must be a positive probability that observations are present in the treatment group. In other words, there should be a reasonable overlap between the covariate distribution of the treatment and control group. For example, there must be car-lovers in both urban and suburban areas to be able to disentangle the effect from living in a suburban area from being a car-lover. Lastly, SUTVA states that the treatment assignment of one unit does not affect the potential outcomes of other units. It assumes no interference or spillover effects between units. It also states that there is only one consistent version of treatment. Here, this means that the built environment of one neighborhood has no influence on the travel behavior of the residents of another neighborhood. We introduce spatial lag variables to capture these spillover effects and slightly relax the SUTVA of no interactions between units (Kolak and Anselin, 2020) (see Appendix A.3 for more details). For the moment, we assume that these assumptions sufficiently hold and discuss potential violations and implications for the robustness of our results and applicability of the model in the discussion Section 5.4.

3.4. Evaluation

Treatment effect metrics. To facilitate interpretation of the treatment effect, we report it as a percentage point change compared to the average per capita travel-related CO_2 emissions. We aggregate the individual treatment effects (ITEs) and calculate the group average treatment effects (GATEs) at the neighborhood and neighborhood type level to obtain a measure of the average impact of a particular built environment. We further report what share of the observed variability in travel-related CO_2 emissions between households and between neighborhoods can be explained by the built environment using the coefficient of determination (R^2). Finally, to compare different model specifications, we report the average of the absolute individual treatment effects (ATE) and the marginal average treatment effect (MATE) of one standard deviation in the direction of a more sustainable built environment (e.g. higher population density and lower distance to city center). Both can be understood as a measure of the overall impact of the built environment, with the ATE giving an indication of how much existing differences in the city's built environment affect travel, and the MATE giving an indication of how marginal changes in the city's overall built environment, such as an overall increase in population density or transit frequency, would affect travel.

To answer the question of how a relocation within the city affects travel-related CO_2 emissions, we calculate the average treatment effect for moving from one neighborhood type to another, for example from a mixed-use urban to a suburban neighborhood. For new housing developments, the travel-related CO_2 emissions of residents are calculated as follows: For all households in our survey, we first calculate the individual treatment effect of moving from their current location to the location of the housing development. To estimate their new emissions, we then offset their observed emissions with the treatment effect. Finally, we calculate the average of the estimated new emissions across all households to capture the full range of covariates.

Effect composition. We decompose the total effect of the built environment into the 5D dimensions and discuss the main drivers for low-carbon urban development. To make the 5D dimensions comparable in terms of their contribution to the overall marginal effect, we standardize all built environment characteristics by subtracting the mean, scaling to unit variance and ensuring consistent direction of feature values.

Confounding effects. To examine the importance of residential self-selection, we calculate the "built environment proportion" (BEP) as described by van Herick and Mokhtarian (2020). It describes the ratio of the true effect of the built environment to the apparent effect, which is biased by residential self-selection. Given the lack of a standardized calculation method, we propose two approaches for determining this proportion within the DML framework: First, based on the treatment effects, specifically as the ratio of the MATE when controlling for residential self-selection compared to when ignoring it. Second, based on the variance explained, specifically as the ratio of R^2 at the neighborhood level when controlling for residential self-selection compared to when ignoring it. We use dummy mean-predicting nuisance models to mimic the case of absent confounding in the residualization stage of the DML model. To provide further insight into the strength of confounding, we report the R^2 values of the nuisance regression models. To identify which socio-demographic characteristics and travel-related attitudes drive residential self-selection, we calculate the gain-based feature importance of the nuisance XGBoost models.

Nonlinearity of residential self-selection. To examine how our methodological contribution of accounting for potential nonlinearities in residential self-selection affects the estimated treatment effects, we compare our results to those obtained using a multiple linear regression model for confounding effects, similar to previous studies. This model specification will be referred to as "Household-level (linear)". We analyze how the magnitude and direction of the built environment effects differ between the two approaches and compare the R^2 values of the machine learning and linear nuisance models to assess which method more accurately captures the confounding effects.

Moderating effects. To analyze the heterogeneity of the built environment's effect on travel behavior, we calculate SHAP (SHapley Additive exPlanations) (Lundberg and Lee, 2017) values of the final stage Causal Forest model. SHAP values indicate the influence of each feature on the model's prediction by distributing the difference between the prediction and the mean value across the features based on cooperative game theory. This allows us to examine the moderating, non-linear influence of confounding socio-demographic traits and travel-related attitudes on the effect. We also analyze the heterogeneity associated with covariate quartiles, as done in previous studies. We calculate the conditional MATE difference between the 1st and 4th quartile of the covariates as well as the differences in covariates between the 1st and 4th quartile of the estimated marginal individual treatment effects.

Robustness checks. To examine the robustness of our results, we conduct four experiments: (1) We assess the existence of threshold effects and other nonlinearities of the built environment's influence in Berlin using a XGBoost model and discuss their implications for our DML approach. (2) We test whether a categorical treatment based on neighborhood types, which avoids linearity assumptions, yields different effect estimates compared to our baseline DML approach that uses multiple, continuous treatments along the 5Ds and suffers from multicollinearity. (3) To address the low explanatory power and the resulting high uncertainty of treatment effect estimates, we perform the analysis at the neighborhood level using average household travel behavior to mitigate the volatility of daily travel. (4) Finally, we assess how including neighborhood-level environmental policy preferences affects effect estimates, addressing the issue of unobserved attitudinal confounders. Please refer to Appendix C.1 for more details.

3.5. Urban planning case study

To highlight the practical importance for urban climate mitigation, we evaluate current plans for 64,000 new residential units (SenStadt, 2023) in Berlin in terms of total induced transport CO_2 emissions. For each planned housing project, we characterize the built environment for the specific site and in a 1.2 km buffer around it, equal to the average size of a zip code neighborhood (see Appendix A.2 for more details). To determine the expected travel-related CO_2 emissions of new residents, we calculate the average across all households whose actual emissions are offset by the estimated individual treatment effect of moving to the new location. We compare the induced CO_2 emissions of three alternative densification scenarios: (1) transit-oriented development focused on neighborhoods whose residents have an average walk time of less than 7 min to the nearest rail station, (2) compact development within the urban rail line that circles central Berlin, the so-called "Ringbahn", and (3) spatially differentiated compact development of low-emission neighborhoods, where the built environment is estimated to have the largest reducing impact on CO_2 emissions. We assume that the housing units will be evenly distributed among the targeted neighborhoods for all three scenarios. We estimate the CO_2 emissions of new residents based on the average treatment effect of moving to these neighborhoods. We characterize the potential CO_2 emissions savings from each strategy relative to the average per capita CO_2 emissions. We assume that the built environment remains unchanged except for the increase in population density, which means that we do not take into account potential changes, such as road infrastructure or access to public transportation, associated with the new housing development.

To more broadly demonstrate how the method can be used to prospectively evaluate changes in the built environment and inform low-carbon urban planning, we exemplary model two scenarios: (1) We estimate the effect of a improved accessibility to public transport, namely a maximum walking time to the nearest rail station of 5 min, and (2) a densification of the neighborhoods within or in close proximity to the "Ringbahn" so that the population density is at the level of the currently densest neighborhoods of Moabit, Neukölln, and Prenzlauerberg.

4. Results

4.1. Descriptive analysis

According to the our dataset, the average trip distance in Berlin is 5.4 km (see Appendix Table 8). 30% of trips are done on foot, 29% by car, 21% by transit, and 19% by bike. In terms of distance, 43% of kilometers are traveled by car, 38% by transit, 13% by bike and 5% by foot. In the center the shares of active mobility and transit usage are higher, while in the suburbs driving dominates (see Fig. 3A). Trip distances also increase notably with distance from the center, from an average of 4.0 km in mixed-use urban areas to 7.3 km in exurban areas.

We estimate the average weekday per capita emissions from urban mobility in Berlin to be 1.9 $kgCO_2$ per day, resulting in 0.5 tCO_2 per year, assuming 260 weekdays. Yet, with a standard deviation of 2.6 $kgCO_2$, they vary strongly between households for the surveyed day. Averaging the per capita emissions of each neighborhood reduces the variance to 0.7 $kgCO_2$ with emissions ranging from 1.2 $kgCO_2$ in mixed-use urban areas to 2.8 $kgCO_2$ exurban areas (see Fig. 3B). The main source of travel-related emissions is driving, with 43% of distances covered by car, accounting for 72% of all travel-related CO_2 emissions.

Shopping, leisure and work are the most important purposes for travel (see Fig. 3D and Appendix Table 8). While on average shopping trips are relatively short (3.3 km) and school and leisure trips are moderate in length (4.2 km and 5.2 km, respectively), work commuting trips are notably longer (11.1 km). They account for 36% of total travel distance, making it the trip purpose with the largest share of associated CO_2 emissions.

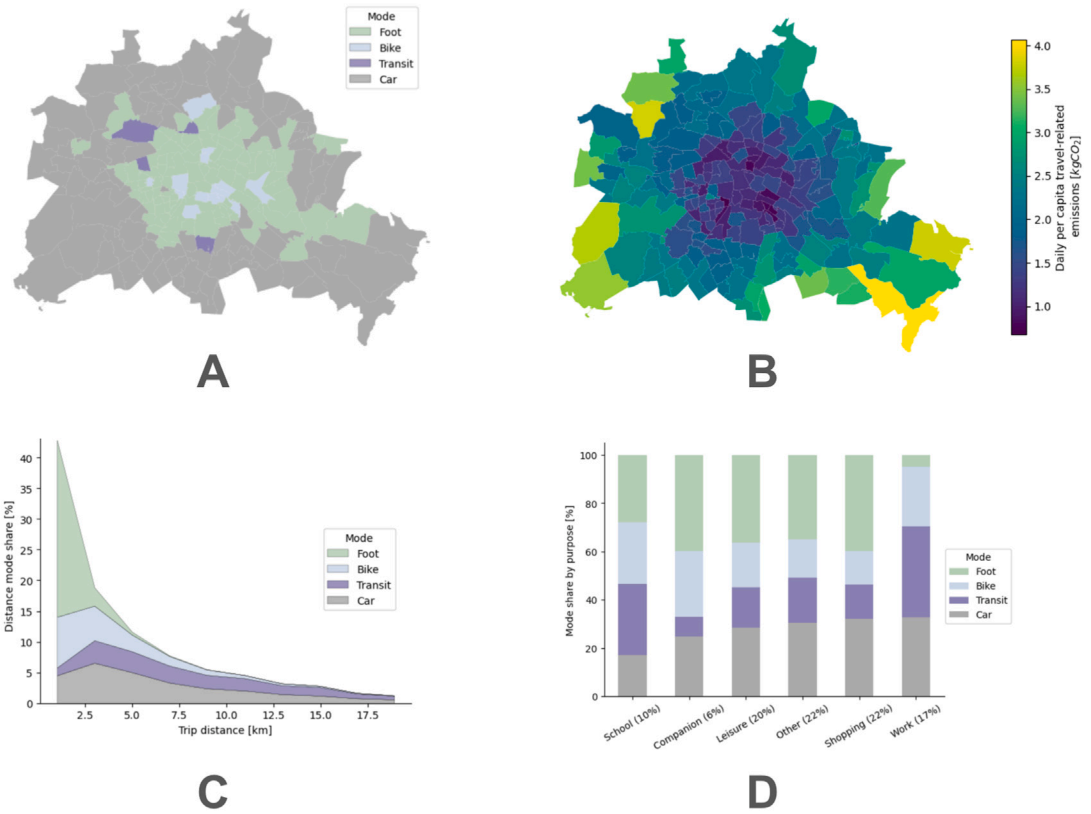

Fig. 3. Overview of trip mode, purpose, distance, and related CO_2 emissions. (A) Predominant mode of transport for each neighborhood based on trip counts. (B) Average daily travel-related CO_2 emissions per capita for each neighborhood. (C) Trip distance specific mode share. (D) Trip purpose specific mode share, ordered by increasing car share.

4.2. Clustering of neighborhoods

K-means clustering of neighborhoods, based on their urban form characteristics, yields 4 interpretable groups, which we label as exurban, suburban, urban (residential), and urban (mixed-use) (see Fig. 4). These neighborhood types show consistent differences in both their built environment and the travel behavior of their residents. Residents of exurban neighborhoods have by far the highest total daily distance traveled per capita (26 km), car use (44%), and travel-related emissions (2.8 $kgCO_2$), followed by residents of suburban neighborhoods (21 km, 37%, 2.1 $kgCO_2$ respectively). Both types of urban neighborhoods have similarly high population density and accessibility to destinations, as well as relatively low distance traveled (15.5–17.2 km), car use (17%–20%), and travel-related emissions (1.2–1.4 $kgCO_2$). But the proportion of mixed-use development, POI density, and street connectivity are substantially higher in mixed-use urban neighborhoods than in residential urban neighborhoods.

4.3. Causal inference

Treatment effect. Using DML to account for residential self-selection, we find that the built environment has a considerable effect on travel behavior and related CO_2 emissions. Compared to the city average, the built environment facilitates a 26%pt decrease in travel-related CO_2 emissions for households living in mixed-use urban areas and a 35%pt increase in exurban areas (see GATE in Table 3). In other words, if a household were to move from a mixed-use urban to an exurban area, travel-related CO_2 emissions would be expected to increase by 61%pt (see Fig. 5). When looking at individual neighborhoods, the effect can be even larger: The neighborhood-level GATE ranges from −39%pt to +78%pt, resulting in a difference of 117%pt with respect to the average travel-related CO_2 emissions. Overall, the built environment can explain 72% of the observed variability in neighborhood CO_2 emissions (see Fig. 5 and R^2 of neighborhood GATE in Table 3). The remaining residuals may partly be due to residential self-selection, noise in the data, and an insufficient characterization of the built environment. On the household level, the built environment can only explain 9% of the variability of per capita travel-related CO_2 emissions given the volatility of daily travel behavior.

Effect composition. Of the 5Ds, destination accessibility has the largest impact on travel-related CO_2 emissions. We distinguish between accessibility to local destinations, operationalized by the density of POIs in the local neighborhood, and accessibility to

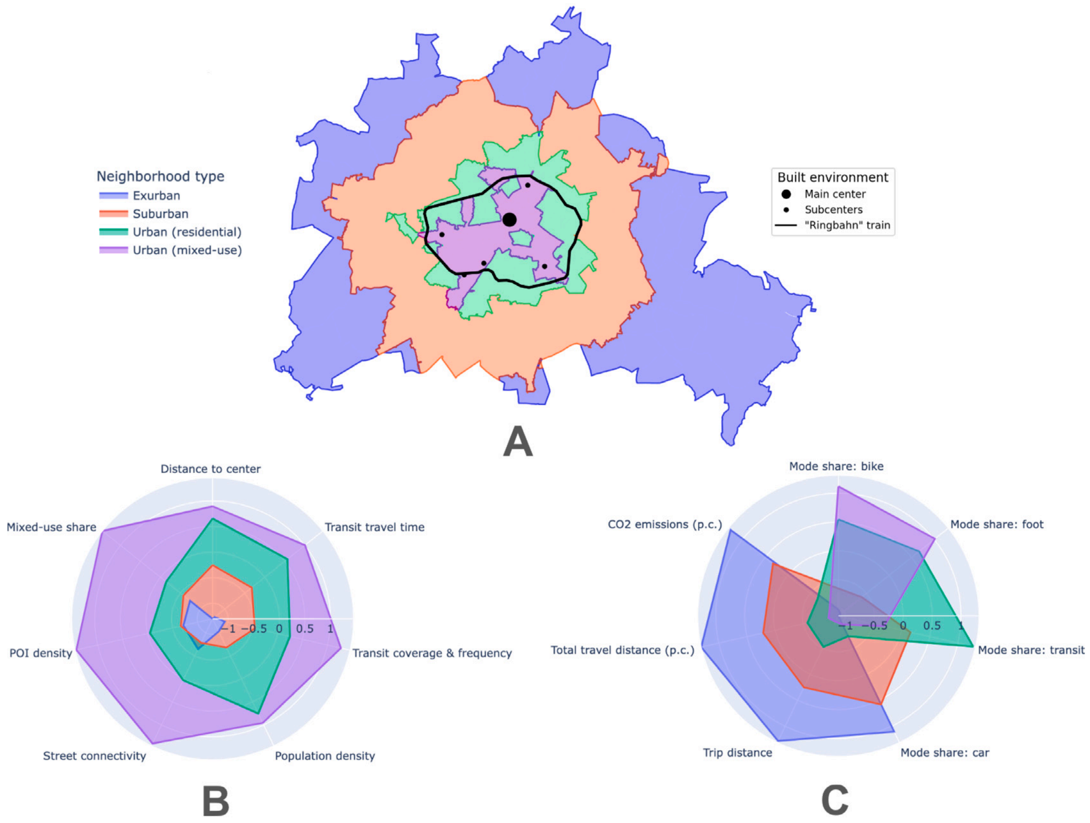

Fig. 4. Neighborhood types. (A) Map of zip code areas categorized into exurban, suburban, urban (residential), urban (mixed-use) using k-means clustering. For reference, the center, subcenters, and the "Ringbahn", an urban rail line that circles central Berlin, are indicated in black. (B) Polar plot of the normalized built environment characteristics per neighborhood type. (C) Polar plot of the normalized travel behavior per neighborhood type (p.c. refers to per capita).

city-wide destinations, operationalized by the driving distance to the center and the nearest subcenter, as well as the travel time to reach POIs by public transport. We find that city-wide destination accessibility is most important, accounting for 63.2% of the total built environment effect, while local destination accessibility, which is more relevant for short-distance active mobility, is less important for the total travel-related CO_2 emissions (see Table 4). The second most important D is density, with the adjusted population density being responsible for 12.3% of the total effect, followed by distance to transit (7.7%), street connectivity (4.9%) and land-use diversity (2.3%).

Confounding effects. The predictive performances of the first stage nuisance models indicate a low degree of confounding, with R^2 values of 0.17 for predicting the target and 0.10 for predicting the treatment from the controls (see Table 3). The most important confounding factor is car ownership, followed by education, which is important for residential self-selection, and the household type family, which particularly affect travel-related CO_2 emissions (see Table 5). Comparing the causal effect estimates with those of a model that does not account for confounding factors indicates that between 72% and 92% of the apparent effect of the built environment is truly due to the built environment, while the rest is due to residential self-selection (see BEP in Table 3). The exact proportion depends on the calculation method, 72% are due to the built environment when using the MATE as a reference, and 92% when considering the variability explained at the neighborhood level.

Nonlinearity of residential self-selection. Accounting for nonlinearities in the confounding influence by using a XGBoost model instead of a multiple linear regression model reduces the estimated impact of the built environment on travel-related CO_2 emissions. The ATE decreases by 13% and the MATE changes by 18% (see Table 3). The built environment effect of different neighborhood types is also muted. Instead, the residential self-selection bias becomes more pronounced, with the BEP decreasing by 13.4%pt on the household and by 4.7%pt on the neighborhood level. This likely results from the confounding influence being better captured by the machine learning model, with R^2 values increasing from 0.12 to 0.17 for predicting the treatment and from 0.07 to 0.10 for predicting the outcome. This suggests the presence of nonlinearities in the relationship between travel preferences and travel behavior and the built environment, which a linear model struggles to capture. These results remain consistent when using one-hot coded life stages instead of a continuous household age variable, as done in some previous studies.

Moderating effects. The estimated effect of the built environment is moderated by socio-demographics and travel-related attitudes. Age, income, and car ownership are the most important moderators and together responsible for 81.4% of the effect heterogeneity

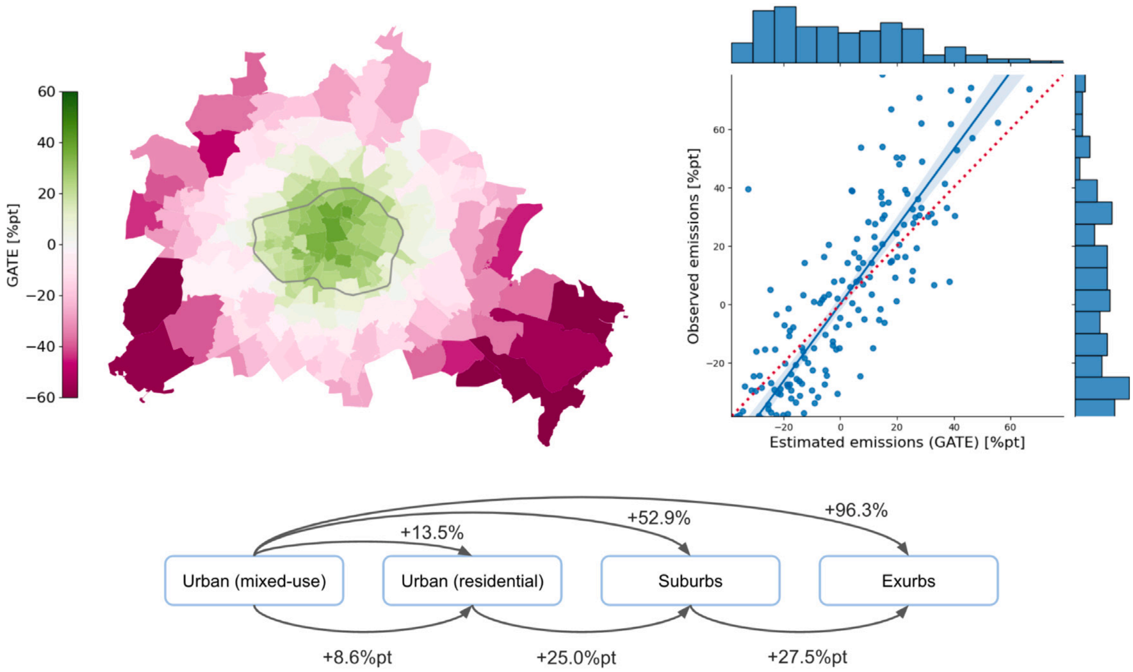

Fig. 5. Built environment significantly impacts travel behavior and related CO_2 emissions. Top left: Group average treatment effect (GATE) per neighborhood, expressed relative to the average per capita travel-related CO_2 emissions. The "Ringbahn" train is plotted in gray for orientation. Top right: Comparison between the estimated GATE of the built environment and the observed differences in CO_2 emissions between neighborhoods. For Berlin, 72% of the observed variability can be explained by the GATE. Points to the left of the dotted diagonal line indicate that the effect is smaller than the observed difference and the residual possibly due to residential self-selection. The slope of the blue line thus indicates whether the self-selection of residents has a consistent, structural effect on emissions at the neighborhood level (e.g. car lovers consistently live in more suburban areas and transit lovers in more urban areas, masking the true effect of the built environment). Bottom: Expected change in travel-related CO_2 emissions when moving from one neighborhood type to another. Change expressed in percent relative to the mixed-use urban neighborhood average and in percentage points (%pt) relative to the city-wide average. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(see Table 5). Higher income and car ownership are associated with an increasing impact of the built environment on travel-related CO_2 emissions of households. The impact is ∼10% higher for households with a monthly income of 4000 Euros and above, and ∼20% higher for households with 2 or more cars per adult (see Fig. 6). While the built environment plays a major role in facilitating travel for young and middle-aged households (∼15% above average), it is much less important for older households. From the average household age of 55 (excluding children), the impact of the built environment decreases dramatically, and by the age of 70, it is ∼40% less important than for the average household. When not using SHAP values to explain the heterogeneity, but instead examining the effect heterogeneity associated with different covariate quartiles, we find that employment status, household size, family circumstances, and, to a lesser extent, education and bike ownership are associated with notable heterogeneity (see Appendix Table 9). However, given the high correlation between covariates, we rely on SHAP values to identify the underlying drivers of heterogeneity among them.

Robustness checks. Our robustness checks confirm that the built environment's effect on travel-related CO_2 emissions is consistent across various model specifications. The DML estimator effectively handles multiple correlated treatments and noise in the target variable, with effect estimates aligning well with non-linear models. Conducting analyses at the neighborhood level yields similar results to household-level analyses, and adding environmental policy preferences does not considerably reduce the importance of the built environment, suggesting that our proxies for travel-related attitudes capture environmental attitudes to some extent. Overall, the built environment substantially influences travel-related CO_2 emissions, more so than residential self-selection. Please refer to Appendix C.2 for more details on the results.

4.4. Urban planning case study

To highlight the potential of the method for informing low-carbon residential planning, we evaluate the currently planned housing projects in terms of induced travel. We find that 15 of the 22 analyzed housing projects are expected to increase the city average of per capita travel-related CO_2 emissions due to their location (see Fig. 7), on average by 11.9%. This increase is mainly due to the large distance of most projects from the city center. For some projects, such as Wasserstadt Berlin-Oberhavel, Buckower Felder, and Tegel-Nord, the low accessibility to public transportation also substantially increases emissions. For 14 of

Table 3
Treatment effects for different model specification. Our preferred specification is highlighted in bold. The effects are expressed in percentage points (%pt) in relation to the average travel-related CO_2 emissions per capita. The ATE is calculated here as the average of the absolute individual treatment effects. See Appendix Fig. 11 for the distribution of individual treatment effects. The MATE indicates the marginal average treatment affect of one standard deviation in the direction of a more sustainable built environment (e.g. higher population density and lower distance to city center). R^2 refers to the share of the variability in observed emissions that is explained by the estimated built environment effect. The BEP refers to the proportion of the apparent built environment effect that remains when controlling for residential self-selection. On the household level, the BEP is calculated based on the MATE and on the neighborhood level based on the R^2. The ML R^2s are not part of the DML approach, but are only reported to indicate the predictiveness of the treatment and covariates included in the respective model.

Model specification	Household effect				Neighborhood GATE			Neighborhood type GATE				DML nuisance		ML	
	ATE [%pt]	MATE [%pt]	R^2	BEP [%]	Range [%pt]	R^2	BEP [%]	Urban (mixed-use) [%pt]	Urban (residential) [%pt]	Suburban [%pt]	Exurban [%pt]	R^2 $X \rightarrow T$	R^2 $X \rightarrow Y$	R^2 $T \rightarrow Y$	R^2 $T \cup X \rightarrow Y$
Household-level	21.5	−39.6	0.09	71.8	117	0.72	92.2	−26.0	−17.6	8.1	35.4	0.10	0.17	0.07	0.20
Household-level (linear)	24.3	−46.9	0.10	85.2	130	0.76	96.9	−29.6	−19.5	10.2	39.8	0.07	0.12	–	–
Household-level (wo/car ownership)	27.5	−50.4	0.10	91.3	144	0.78	99.2	−35.9	−22.2	11.6	43.2	0.06	0.09	0.07	0.15
Household-level (wo/controls)	30.0	−55.4	0.10	–	157	0.79	–	−35.8	−24.5	13.3	48.2	–	–	0.07	0.07
Household-level (categorical treatment)	18.3	–	0.06	–	66	0.59	76.9	−31.0	−22.0	–	27.1	0.36[a]	0.17	–	–
Neighborhood-level	–	–	–	–	122	0.73	84.2	−21.5	−15.8	5.5	33.1	0.48	0.65	0.77	0.77
Neighborhood-level (w/environmental pref.)	–	–	–	–	114	0.72	83.6	−19.6	−15.6	7.7	32.9	0.59	0.69	0.77	0.78

[a] The nuisance score for the categorical treatment, is given as a F_1 score.

Table 4
Destination accessibility has the strongest impact on travel-related CO_2 emissions. Decomposition of the individual treatment effects (ITEs) into the different dimensions of the built environment. Individual effect shares do not add up to aggregate effect shares because the ITEs may point in opposite directions, e.g. a high local POI density may reduce travel, while a high POI density in nearby areas (spatial lag) may increase travel.

5D's of compact development	Aggregate effect share	Feature	Individual effect share
City-wide destination accessibility	63.2%	Distance to center	37.9%
		Distance to subcenter	6.3%
		Transit travel time	10.4%
Local destination accessibility	9.4%	POI density (spatial lag)	6.6%
		POI density	6.8%
Density	12.3%	Population density	9.5%
		Population density (spatial lag)	11.6%
Diversity	2.3%	Mixed-use share	1.7%
Design	4.9%	Street connectivity	3.6%
Distance to transit	7.7%	Walking time to rail	5.5%

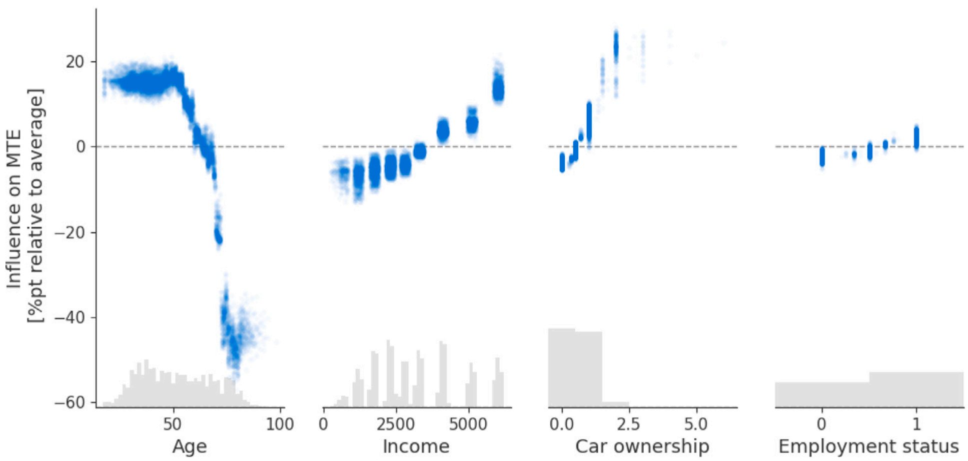

Fig. 6. Age, income, car ownership, and employment status moderate the built environment effect. Scatter plot of the influence of the most important covariates on the marginal treatment effect (MTE) for each household, according to the SHAP values of the Causal Forest. A positive influence on the MTE means that the covariate is associated with an increased effect of the built environment on travel-related CO_2 emissions.

the 22 projects, the relatively low population density of the neighborhood slightly increases emissions, especially for Schönerlinder Straße and Projektverband Karow Süd. For the Schumacher Quartier, the largest potential lies in improving public transport links to the rest of the city. In contrast, the best-located projects in terms of induced emissions, namely Michelangelostrasse and Neue Mitte Tempelhof, are benefiting primarily from the short distance to the city center and the high density of POIs.

Table 5
Moderating and confounding influence of covariates. Heterogeneity associated with covariates according to the conditional marginal treatment effect (MTE) of one standard deviation in the direction of a more sustainable built environment (e.g. higher population density and lower distance to city center). The share of heterogeneity explained by each individual covariate is calculated using the SHAP values of the final CausalForest model. The confounding influence of covariates is measured in terms of the gain-based feature importance of the XGBoost nuisance models. Relative importance above 10% is highlighted in bold.

Covariates	Moderating influence	Confounding influence	
	Share of associated heterogeneity	Importance for residential self-selection	Importance for travel-related CO_2 emissions
Age	**50.9%**	7.3%	3.8%
Income	**16.3%**	4.0%	3.3%
Car ownership	**14.2%**	**28.0%**	**34.2%**
Employment status	6.2%	5.6%	5.8%
Bike ownership	3.9%	3.7%	4.0%
Transit subscription	3.2%	6.3%	7.1%
Household type: single <65y	1.5%	7.4%	3.2%
Higher education	1.3%	**16.1%**	3.9%
Mobility constraints	1.2%	3.7%	2.7%
Household size	0.6%	6.0%	4.5%
Household type: single >65y	0.4%	3.5%	3.6%
Driving license	0.2%	4.7%	5.6%
Household type: family	0.0%	3.8%	**18.2%**

Alternative residential development strategies that focus on densifying the center and transit-oriented development may lead to significantly lower CO_2 emissions. We estimate that CO_2 emissions can be reduced by up to 42.9%pt if residential planning prioritizes the 20 neighborhoods with the most sustainable built environment according to the model (see "Optimum" strategy in Fig. 8). By targeting all neighborhoods with good transit accessibility, specifically within a 7-min walk of the nearest rail station, the CO_2 emissions of future residents are expected to be 14.6%pt below the city average and 26.5%pt below those of the planned settlements (see "TOD" strategy).

To illustrate how the approach can support low-carbon urban planning more broadly, we use it to estimate the impact of improved accessibility to public transit and increased population density in central neighborhoods. We find that such changes have the potential to reduce travel-related CO_2 emissions by up to 10%pt each relative to the city average in some neighborhoods (see Fig. 9). Not surprisingly, the emission-saving potential of improved transit access is largest in the outskirts of the city, where accessibility is currently low. According to our model, the largest potential for densification of neighborhoods still accessible by the "Ringbahn" exists for the localities of Schmargendorf, Mitte, Tiergarten, Charlottenburg-Nord, and the northeastern part of Westend.

5. Discussion

5.1. Implications of non-linear residential self-selection

Our analysis reveals that failing to account for nonlinearities when controlling for residential self-selection risks overestimating the influence of the built environment. The mechanism behind residential self-selection – the influence of travel preferences on residential choice and travel behavior – can be explained more accurately with a machine learning model than with a linear model. The respective R^2 increases from 0.07 to 0.10 for explaining residential choice and from 0.12 to 0.17 for explaining travel behavior (see Table 3), suggesting the presence of nonlinearities in the confounding influence, consistent with previous findings by Ding et al. (2024). Imposing linear constraints artificially reduces the role of residential self-selection and inflates the estimated treatment effects by 13%–18% on average and the built environment proportion by 13.4%pt in Berlin. While further studies in different regions with improved operationalization of attitudes are needed to validate these findings, it suggests that previous research may have consistently underestimated the importance of residential self-selection.

While accounting for nonlinearities comparatively increases the importance of residential self-selection, its confounding influence in Berlin remains limited in absolute terms. Depending on the calculation method, the proportion of the total apparent effect that is actually due to resident self-selection is between 8% and 28%, while the remaining 72% to 92% is due to the built environment, which is consistent with previous studies that estimated values between 34% (Cao and Fan, 2012) and 98% (Cao et al., 2010). Residential self-selection is mainly driven by car ownership, which strongly facilitates households' decisions on where to live and how to travel (see Table 5). In addition, education and family circumstances are important factors to control for. In contrast to previous work (Bhat and Guo, 2007), household income is less important for both residential choice and travel-related CO_2 emissions in Berlin (if car ownership is controlled for separately). Given the lack of travel-related attitudes at the household level, these estimates are subject to some uncertainty. However, our analysis of including environmental policy preferences suggests that these attitudes may not be as important when analyzing the average effect at the neighborhood level, as the built environment proportion (BEP) decreases by only 0.6%pt (see Table 3).

5.2. Who is most affected by urban planning? a moderation analysis of age, income, and car ownership

The built environment's influence on travel-related CO_2 emissions varies across population groups, impacting some households more than others. This dimension is crucial for equitable urban planning, but has received little attention in previous studies. It

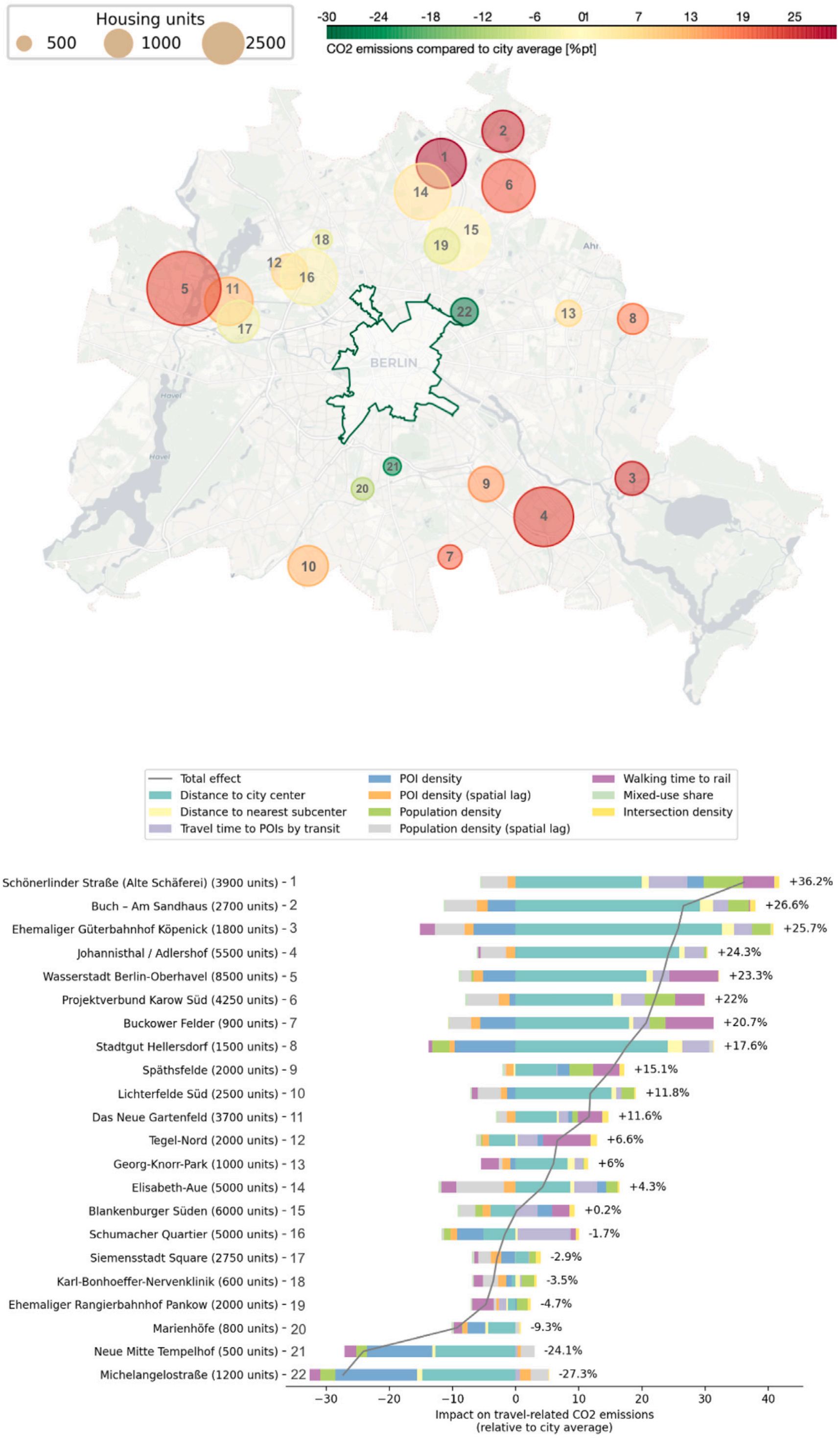

Fig. 7. Large distance to city center causes high travel-related CO_2 emissions for most planned settlements. Top: Locations of planned settlements are marked by circles, with the circle color indicating the expected average travel-related CO_2 emissions of future residents and the circle size indicating the number of planned housing units. Neighborhoods with the largest reducing effect on CO_2 emissions, targeted by the "Optimum" policy, are located within the green border. Bottom: Decomposition of the built environment effect for each planned settlement. The gray line indicates the total effect relative to the city average travel-related CO_2 emissions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

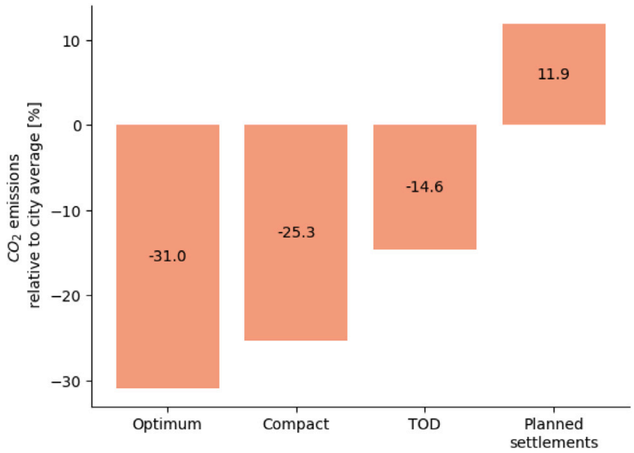

Fig. 8. Travel-related CO_2 emissions for planned settlements 62% (42.9%pt) above the optimal compact development alternative. Comparison of average travel-related CO_2 emissions per capita for different residential planning strategies in Berlin relative to the city average. "Optimum" refers to a densification of the 20 neighborhoods with the lowest CO_2 emissions. Compact development refers to a densification within the so-called "Ringbahn", an urban rail that circles the center. Transit-oriented development (TOD) refers to residential development in neighborhoods where the average walking time to rail-based transit is below 7 min.

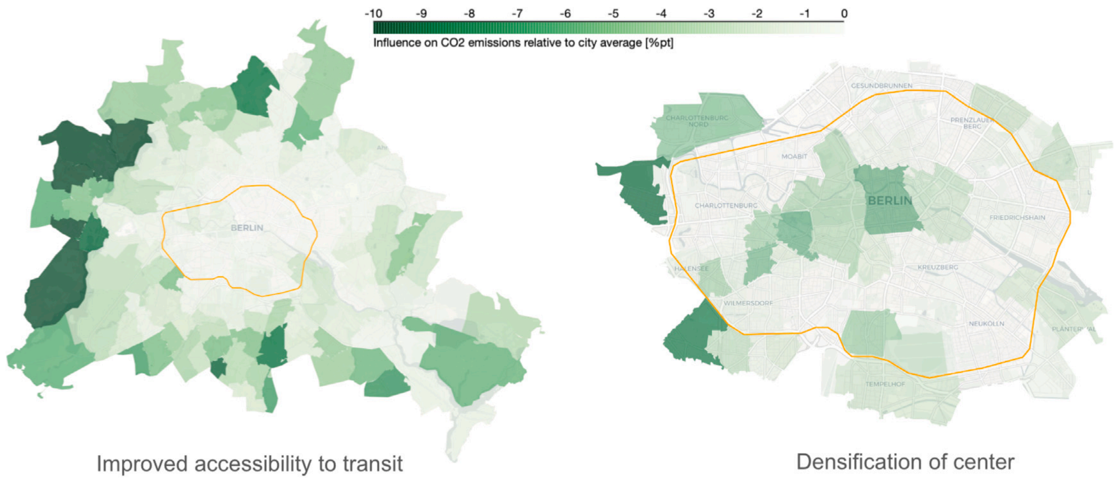

Fig. 9. Predictive modeling of changes in the built environment. Exemplary visualization of how the method can be used to prospectively model the effect of urban planning scenarios on travel-related CO_2 emissions. Left: Improving the accessibility of public transportation everywhere in Berlin to a maximum walking time of 5 min to the nearest rail station. Right: Densification of neighborhoods within or in close proximity to the "Ringbahn" (marked in orange) so that the population density is at the level of Moabit, Neukölln or Prenzlauerberg. The effect of an expected subsequent increase in POI is not taken into account.

has only been considered in depth from a health perspective for walking (e.g. Owen et al., 2007; Manaugh and El-Geneidy, 2011; Richardson et al., 2017). Our analysis reveals that non-senior, high-income, car-owning, working households are most affected, while retirement-age households are least affected. Specifically, after the age of 55, the importance of the built environment declines sharply in a non-linear fashion (see Fig. 6). This raises the important question of how policy and urban planning can enable older households to reduce their travel-related CO_2 emissions while satisfying their mobility needs if built environment interventions are less impactful.

The second and third most important moderating factor are income and car ownership. The influence of the built environment is lower for low-income households, likely because they are financially more constrained in their travel choices, consistent with previous studies on active travel (e.g. Christiansen et al., 2014; Manaugh and El-Geneidy, 2011). Together, age, income, and car ownership explain 81.4% of the effect heterogeneity (see Table 5). While the mediating influence of car ownership has been well studied (e.g. Ding et al., 2017; van Acker and Witlox, 2010), the moderating influence has received less attention, despite its clear impact on households' sensitivity to changes in the built environment. For example, improved access to public transport may have a smaller immediate effect for car-owning households, while parking management may have a larger effect. Only its moderating influence on walking has been studied (e.g. Manaugh and El-Geneidy, 2011; Eriksson et al., 2012), but to our knowledge not that on travel-related CO_2 emissions. For Berlin, we find that the impact of the built environment is larger for car-owning households in terms of travel-related CO_2 emissions, highlighting the importance of a sustainable built environment in reducing emissions despite the lock-in effects of car ownership. We call for future research to validate this finding in other regions and with other methodologies. Future quantitative research should also take into account the non-linear threshold effects, especially of age, by using non-parametric methods when modeling moderating effects or by categorizing age into life stages.

Overall, we can attribute 72% of the variation of neighborhood CO_2 emissions to the built environment. Consistent with the majority of recent case studies on travel-related CO_2 emissions (Liu et al., 2022; Shao et al., 2023; Ao et al., 2019; Zhang et al., 2022), we find that city-wide destination accessibility is the most important dimension of the built environment in Berlin (see Table 4). Comparing its relative importance of 63.2% with other studies is challenging due to differences in local settings and model specifications (Næss, 2022). It tends to be lower in other locations, e.g. 22% in the Minneapolis-St. Paul region, USA (Wu et al., 2019), 31% in Jinan, China (Liu et al., 2022), and 40% in Zhongshan, China (Shao et al., 2023), but higher in Berlin when only driving distances are considered (91%) (Wagner et al., 2022). Considering that distance to the center is generally more important for driving distance than for travel-related CO_2 emissions, as shown in two meta-analyses (Stevens, 2017; Ewing and Cervero, 2010), the 63.2% seems plausible. The second most important built environment dimension is density, which accounts for 12.3% of induced emissions, similar to previous machine learning approaches (Liu et al., 2022; Wu et al., 2019; Shao et al., 2023). However, in contrast to Wu et al. (2019) and Cui et al. (2023), we find the influence of diversity to be much smaller (2.3%), possibly because we also include local POI density, which is closely related to diversity, but consider it as a measure of local destination accessibility. Although common in the literature, we recommend that comparisons of effect composition across the 5Ds be made with caution because, due to the high collinearity of the built environment variables (see Appendix Table 6), the relative importance of each variable depends largely on which others are included and how the model detects and assigns their contributions.

5.3. Enabling evidence-based low-carbon residential planning

To reduce CO_2 emissions and prevent urban sprawl, growing cities face the urgent question of where to locate new residents to minimize travel, associated CO_2 emissions, and other externalities of a car-dependent transport system such as pollution, noise, congestion, and traffic accidents (Parry et al., 2007). Having spatially explicit estimates of the built environment's impact on transport CO_2 emissions at potential new housing sites is therefore fundamental to evidence-based low-carbon residential planning. Here, DML has great potential to support residential planning by utilizing comprehensive travel surveys.

In our case study of urban planning in Berlin, we find that induced transport CO_2 emissions vary considerably between potential residential planning sites. For example, per capita CO_2 emissions differ by a factor of 1.96 between mixed-use urban and exurban areas because of the built environment (see Fig. 5). According to our calculations, the induced CO_2 emissions of the currently planned 66,000 new residential units are 62% above the optimum. The main cause of the high CO_2 emissions is the large distance to the city center (see Fig. 7). For some of these dwelling units, transit accessibility must also be improved. Alternative compact or transit-oriented development strategies with an improved accessibility of destinations would lead to substantially lower CO_2 emissions. Furthermore, in a stressed housing market many people are not able to realize their residential preferences. Providing affordable housing in urban neighborhoods with good transit accessibility enables dissonant suburbanites to realize their residential preference and substantially reduce their car use and travel-related CO_2 emissions (Næss, 2014; Wolday et al., 2018; Guan et al., 2020). Consequently, emissions savings from residential development in central neighborhoods can be expected to be greater than the mere impact of the built environment. While the provision of affordable housing at such locations might require public subsidies, Wolday et al. (2018) argue that the potential savings and co-benefits in terms of reduced congestion, road construction and maintenance, and social segregation may be substantial.

Beyond the evaluation of locations for low-carbon residential planning, DML allows for the prospective evaluation of urban planning scenarios by estimating the effects of changes in the built environment, as briefly demonstrated in Fig. 9. While these analyses are preliminary and only exemplary, they highlight the novel opportunity to predict the impact of infrastructure interventions with high spatial resolution while accounting for residential self-selection.

5.4. Limitations masking the true effect of the built environment

While the accounting for the non-linear influence of confounders and modeling treatment heterogeneity brings our estimates of the built environment's influence on travel behavior one step closer to the true causal effect compared to previous cross-sectional approaches, several issues remain that add uncertainty to the estimate: (1) unobserved travel-related attitudes (violation of unconfoundedness), (2) inadequate operationalization of the built environment (linearity assumptions, MAUP and violation of SUTVA), and (3) oversimplified conceptual representation of the interdependencies between the built environment, travel preferences, and travel behavior (violation of ignorability).

First and foremost, we are not able to fully account for the confounding influence of attitude-based residential self-selection because we do not directly observe travel-related attitudes. While we use the number of cars and bicycles owned as a proxy for manifested attitudes and show that these confound the relationship similarly to environmental policy preferences, there may be other attitudes that influence residential choice and travel decisions that are not manifested in transportation ownership. Also, using endogenous variables such as car ownership as proxies for travel-related attitudes is likely to underestimate the true effect of the built environment, as these variables act as mediators that are influenced by the environment itself (Næss, 2009).

Second, given the limited knowledge of precise residential location, we are constrained to characterize the built environment using arbitrarily shaped zip code areas that do not correspond to meaningful livable neighborhoods. This averaging over heterogeneous areas can artificially smooth the built environment, potentially underestimating its true effect (MAUP effect Fotheringham and Wong, 1991; Zhang and Kukadia, 2005). Additionally, we assume a linear or logarithmic relationship between the built environment and travel behavior, despite previous studies demonstrating that nonlinearities and threshold effects may exist (e.g. sufficient density to support rail-based transport) (Cao and Tao, 2023; Aghaabbasi and Chalermpong, 2023). While our comparison with a nonlinear effect estimator did not reveal threshold effects in our study area and suggested an approximately linear or logarithmic relationship (see Fig. 10), this assumption may still limit a realistic characterization of the built environment's impact. Our approach also cannot capture non-linear complementarities and interdependencies between different built environment dimensions, leading to a biased estimation as recent studies have challenged the notion of the 5Ds being independent (Næss et al., 2021; Wagner et al., 2023). To examine the potential impact of these interdependencies, we employ an alternative model specification that uses neighborhood type as a single categorical treatment. Our findings indicate only small differences in the overall effect of the built environment, suggesting that these interdependencies do not significantly bias the overall effect, but mainly the effect composition. Finally, the built environment of one area also affects travel behavior in neighboring areas, violating SUTVA. We include spatial lag variables to account for these spillover effects and mitigate the violation of SUTVA, but further research is needed to validate the robustness of this approach (e.g. Credit and Lehnert, 2023; Kolak and Anselin, 2020).

Third, travel attitudes are likely affected to some extent by the built environment and past travel behavior, and thus are interdependent rather than an exogenous predisposition (Næss, 2009; Scheiner, 2018) (see D in causal graph in Fig. 1). By not capturing this mediating effect, driven for example by the availability of travel options (van de Coevering et al., 2016; Bamberg, 2006) and transport cultures (Klinger and Lanzendorf, 2016), we violate the ignorability assumption and acknowledge a tendency to underestimate the true effect of the built environment in the long run. In our study, this is reinforced by neglecting the known impact of the built environment on car ownership (Ding et al., 2017; van Acker and Witlox, 2010) and instead treating it as a time-lagged proxy for travel attitudes. Our conceptualization of travel preferences also does not take into account that households may not be able to fully realize their residential preferences because of limited housing availability and individual constraints, such as children, income, and age (Wolday et al., 2018). The actual impact of the built environment may therefore be larger (Chatman, 2009).

In summary, we believe that although our DML estimates are subject to some uncertainty and further methodological improvements are needed to make them more robust, the approach is preferable to standard machine learning techniques by accounting for causal relationships and confounding effects, and is also preferable to previously used statistical methods such as propensity score matching and sample selection for cross-sectional data analysis because of its ability to model the non-parametric confounding and moderating influence of travel preferences. DML further has great potential beyond survey-based data on travel behavior and attitudes, as it is able to identify important confounding factors from a large dataset, such as user data from health, sports and navigation apps.

6. Conclusions

With the need to mitigate climate change and reduce transport CO_2 emissions, growing cities face the urgent question of where to locate new residents to minimize travel and related CO_2 emissions. We explore the potential of using double machine learning (DML) to estimate the influence of the built environment on travel-related CO_2 emissions for different neighborhoods, aiming to identify potential locations for new, low-carbon residential development. By using flexible machine learning estimators to account for confounding effects, we are able to control for residential self-selection without assuming a specific functional form, unlike previous work. Additionally, we are able to examine how socio-demographics and travel-related attitudes non-linearly moderate the influence of the built environment on travel behavior by using a Causal Forest to estimate individual treatment effects.

We find that the total impact of the built environment on travel-related CO_2 emissions in Berlin is substantial. Per capita CO_2 emissions differ by a factor of 1.96 between mixed-use urban and exurban areas because of the built environment. This disparity is mainly due to the different accessibility of destinations (73%), followed by differences in population density (12%). The effect of the built environment is largest for households that are non-senior, high-income, and car-owning. Together, age, income, and car ownership can explain 81% of the effect heterogeneity. Overall, the built environment can explain 72% of the observed variability in neighborhood CO_2 emissions in Berlin. We estimate residential self-selection to be responsible for 8%–28% of the apparent effect of the built environment, mainly driven by the confounding influence of car ownership. Failure to account for nonlinearities underestimates this by 13%pt and instead inflates the treatment effect of the built environment by 13%–18%. Assessing the current plans for 64,000 new residential units in Berlin in terms of induced travel, we find that per capita travel-related CO_2 emissions from new residents can be expected to be 12% above the city average and 62% above what could be achieved through spatially differentiated compact development as informed by our model.

While accounting for non-linear confounding and moderating effects advances the status quo for estimating the influence of the built environment on travel behavior from cross-sectional data, several uncertainties remain. Most importantly, the oversimplified conceptual representation of the interdependencies between the built environment, travel preferences, and travel behavior, which neglects mediating effects via transport cultures, etc., but also the limited methodological groundwork on how to account for spatial confounding and spillover effects within the DML framework. Nevertheless, we believe that DML is a promising method to facilitate and scale the estimation of travel demand induced by the built environment and provide cities with a cost-effective tool to establish a starting point for evidence-based sustainable urban planning.

CRediT authorship contribution statement

Florian Nachtigall: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Felix Wagner:** Writing – review & editing, Visualization, Software, Methodology, Conceptualization. **Peter Berrill:** Writing – review & editing, Visualization, Resources, Methodology, Data curation. **Felix Creutzig:** Writing – review & editing, Visualization, Supervision, Resources, Funding acquisition.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used "DeepL" and "ChatGPT" to improve the writing style. After using these tools, the author reviewed and edited the content as needed and take full responsibility for the content of the publication.

Acknowledgments and disclosure of funding

We thank Adam Millard-Ball, Nikola Milojevic-Dupont, and Aneeque Javaid for valuable discussions and review of our manuscript. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101056810 (CircEUlar project). The authors gratefully acknowledge the Ministry of Research, Science and Culture (MWFK) of Land Brandenburg for supporting this project by providing resources on the high performance computer system at the Potsdam Institute for Climate Impact Research.

Appendix A. Feature engineering

A.1. Built environment characteristics of neighborhoods

POI density. The Points of Interest (POI) density describes the combined densities of second places (offices, schools, universities, etc.), third places (cafes, restaurants, bars, etc.), healthcare (doctors, pharmacies, clinics, etc.) and retail (supermarkets, clothing, convenience, etc.) in the residential zip code area retrieved from OpenStreetMap (OpenStreetMap contributors, 2017). The idea was to create a composite index that captures the diversity of travel purposes beyond commuting to better model the travel behavior of the non-full-time working population. To give equal weight to each category, we normalize the densities of each category and calculate the average across them. We only consider built-up area according to Berlin's land-use data (Open Data Berlin, 2023) to avoid distorting densities by large green areas in otherwise dense neighborhoods. POI density serves as a measure of local destination accessibility, which is important for facilitating active travel and reducing travel distances.

Distance to center & subcenter. The distance to the main center is defined as the driving distance along the fastest route to the central business district. The fastest route is determined based on the street network from OpenStreetMap with inferred travel times. To account for irregularly shaped neighborhoods, we randomly select 10 starting points in the neighborhood and average the travel distance. We use the fastest route instead of the shortest route or as the crow flies, because it is the most meaningful for travel-related CO_2 emissions.

Likewise, we calculate the distance to the nearest subcenter. The subcenters are determined based on the POI density. We identify the 10 neighborhoods with the highest POI density, merge adjacent neighborhoods, and use their centroid to define the location of subcenters. This results in 5 subcenters located in the districts of Schöneberg, Neukölln, Prenzlauer Berg, Charlottenburg and Friedenau (see Fig. 2).

Transit travel time. As a measure of destination accessibility by public transit, we calculate the average travel time to a sample of 100 POIs located in the center and subcenter using the GoogleMaps API. As starting points, we randomly select 10 locations from the residential area of each neighborhood. The travel times are calculated for a Tuesday morning at 8 am.

Transit coverage & frequency. To characterize the accessibility of transit in each neighborhood, we develop an index that describes the spatial coverage and frequency of transit services, as used in a study by Berrill et al. (2024). We use openly available General Transit Feed Specification (GTFS) data on transit services from the local transit operator "Verkehrsverbund Berlin-Brandenburg" (VBB, 2023). For each neighborhood, we calculate the average accessibility across multiple locations using the centroids of Uber's H3 hexagonal grid with a resolution of 9 (0.1 km^2 per hexagon). For each location, we determine the transit lines (bus, tram, metro and urban rail) within walking distance and sum up their hourly departures. We use a gravity-based model to discount the number of departures with the distance to the station using a Gaussian decay function with a sigma of 0.5. This means, for example, that the transit accessibility score 500 m from a station will be 60% of the accessibility score right at the station.

A.2. Built environment characteristics of planned housing projects

Overall, the built environment is characterized in the same way as before. However, because we know the exact location of the planned housing projects, as opposed to the residential locations of households where we only know the zip code, we are able to characterize the built environment around the housing projects in a more localized way. Thus, we calculate the distance to the main center and nearest sub-center, the travel time to POIs by public transport, and the walking time to the nearest rail station for the exact location of the planned residential project. We use openly available GTFS data (VBB, 2023) to calculate the distance to the nearest rail station. The POI, population, and intersection densities, as well as the share of mixed-use development, are calculated around a buffer equal to the average size of a Berlin zip code neighborhood (i.e. 4.7 km^2). We use the highest resolution population density data available at the block level (on average 0.03 km^2) and calculate the weighted average based on the area intersected by the buffer. The spatial lag variables are calculated, as before, for the zip code neighborhoods around the planned site.

A.3. Spatial effects

Spatial spillovers & violation of SUTVA. The existence of spatial treatment spillovers between units generally violates the Stable Unit Treatment Value Assumption (SUTVA). This is inherent in spatial observational studies. However, the extent of bias can be limited by choosing appropriate spatial units and a model specification that accounts for such spatial effects. For example, aggregation of spatial units can help reduce the magnitude of spillovers, but at the cost of increasing the bias due to the MAUP effect (Kolak and Anselin, 2020). The ideal level of aggregation depends on the travel dimension of interest. For walking-specific treatments, using the zip code area as the unit of analysis should mitigate spillover problems because most walking trips occur within the zip code area, and therefore the built environment in other areas has a negligible effect on walking. To mitigate cycling-specific spillovers, the treatment must take into account the built environment within a few kilometers of the residential location. Here, spatial lag variables of the treatment of neighboring zip code areas are important to internalize the effect (Kolak and Anselin, 2020). This results in a model specification similar to a spatially-lagged X model (SLX) (Halleck Vega and Elhorst, 2015). For car and transit-specific treatments, accessibility to destinations across the city is important. Therefore, we do not use a local built environment characteristic here, but consider the whole city, e.g. using the average transit travel time to all POIs in the city. In summary, by capturing spatial spillovers by using spatial lag variables where appropriate, choosing an appropriate spatial unit, and using characteristics of the built environment that inherently incorporate spatial effects, we are able to mitigate the bias that results from violating the SUTVA.

Spatial lag variables. We include spatial lag variables for a selection of built environment characteristics to capture the spatial spillover effects of adjacent neighborhoods; for example, a high POI density in an adjacent neighborhood may increase cycling and shorten trips. Specifically, we include the average of the population and POI density of adjacent neighborhoods as lag variables. We do not include spatial lag variables for walkability and land use diversity in order to keep the number of treatment dimensions small, as initial experiments show that the spatial spillover effect is negligible in addition to the effects of population and POI density. No lag variables are needed for distance to center, distance to subcenters, transit travel time, and accessibility to transit because the spatial dimension is naturally accounted for in the calculation of the variables. For example, for accessibility to transit, we already consider transit stops in neighboring neighborhoods and discount their importance by the walking distance to them.

Appendix B. Orthogonalization & cross-fitting

Recall the partially linear model from Section 3.3:

$$Y = \theta_0 D_0 + g_0(X) + \epsilon \quad \mathbb{E}[\epsilon|X, D] = 0 \tag{10}$$

$$D = m_0(X) + \zeta \quad \mathbb{E}[\zeta|X] = 0 \tag{11}$$

where Y is the outcome, D is the treatment, X are the covariates, θ_0 is the causal parameter of interest, $g_0(X)$ and $m_0(X)$ are the nuisance functions, $\mathbb{E}[\cdot]$ is an empirical expectation function, and ζ and ϵ are stochastic disturbances.

Mitigating regularization bias using orthogonalization. Orthogonalization aims to make the estimation of the causal parameter insensitive to small estimation errors in the nuisance functions, allowing the use of noisy machine learning-based nuisance estimators. This robustness is achieved by constructing a so-called Neyman orthogonal score function $\psi(W; \check{\theta}_0, \hat{\eta}_0)$, where $W = (Y, D, X)$ denotes the observed data and $\eta_0 = (g_0, m_0)$ denotes the nuisance functions, that satisfies the Neyman orthogonality condition, formally given by

$$\partial_\eta \mathbb{E}[\psi(W; \theta_0, \eta_0)][\eta - \eta_0] = 0, \tag{12}$$

Moreover, for a consistent estimation, the score function must satisfy a moment condition for the parameter of interest η_0, i.e. $\mathbb{E}[\psi(W; \check{\theta}_0, \hat{\eta}_0)] = 0$. For a partially linear model, Chernozhukov et al. (2018) give an example of a Neyman orthogonal score function:

$$\psi(W; \check{\theta}_0, \hat{\eta}_0) = (D - \hat{m}_0(X))(Y - \hat{g}_0(X) - D\check{\theta}_0) \tag{13}$$

Here, robustness is achieved by partialling out the effect of X from Y and D, essentially removing the confounding influence and isolating the exogenous variation in D. The orthogonalized regressor $\tilde{D} = D - \hat{m}_0(X)$ enables the estimation of $\check{\theta}_0$ using the Generalized Method of Moments (GMM), based on the moment condition of the score function $\mathbb{E}[\psi(W; \check{\theta}_0, \hat{\eta}_0)] = 0$, for a sample I of size n evaluated using sample averages:

$$\frac{1}{n} \sum_{i \in I} \tilde{D}(Y - \hat{g}_0(X) - D\check{\theta}_0) = 0 \tag{14}$$

Expanding and rearranging this expression to solve for $\check{\theta}_0$, yields one of the double machine learning estimator from Chernozhukov et al. (2018):

$$\check{\theta}_0 = \left(\frac{1}{n} \sum_{i \in I} \tilde{D}_i D_i \right)^{-1} \frac{1}{n} \sum_{i \in I} \tilde{D}_i (Y_i - \hat{g}_0(X_i)), \tag{15}$$

Mitigating overfitting bias with cross-fitting. The idea of cross-fitting is to split the data and estimate the causal parameter θ_0 independently of the estimation of the nuisance functions g_0 and m_0 in order to avoid overfitting bias. To do so, the sample of size n is divided into K approximately equal-sized folds $I_1, I_2, \dots, I_K$. Similar to a cross-validation approach, for each fold $k \in \{1, 2, \dots, K\}$, the nuisance functions are fitted using all other folds $I_{-k} = \bigcup_{j \in \{1,2,\dots,K\} \setminus \{k\}} I_j$, yielding $\hat{g}_{0,-k}$ and $\hat{m}_{0,-k}$, to predict the nuisance values on the unseen kth fold and compute their residuals:

$$\tilde{Y}_i = Y_i - \hat{g}_{0,-k}(X_i), \quad \tilde{D}_i = D_i - \hat{m}_{0,-k}(X_i) \quad \text{for } i \in I_k. \tag{16}$$

Similar to before, the estimator of θ_0 is constructed as the solution to the moment conditions of the Neyman orthogonal score ψ, but now using the empirical expectations of the different folds:

$$\frac{1}{K} \sum_{k \in K} \mathbb{E}_k[\psi(W; \check{\theta}_0, \hat{\eta}_{0,-k})] = 0 \tag{17}$$

Appendix C. Robustness checks

C.1. Methods

To test the robustness of our results, we conduct four experiments to address the issues of low explanatory power, unobserved attitudinal confounders, treatment multicollinearity and partial linear treatment model, and compare the resulting (average) treatment effect for each neighborhood with our baseline approach.

Effect linearity. To assess whether the inability of the partial linear treatment model to capture the potentially nonlinear influence of the built environment has consequences for our effect estimates, we examine the extent to which the influence in Berlin is indeed nonlinear. We compare the neighborhood GATEs of each built environment dimension to the SHAP values of a gradient boosting decision tree, i.e. XGBoost, which is commonly used in the literature to examine the nonlinearity of the built environment's influence (Aghaabbasi and Chalermpong, 2023; Cao and Tao, 2023). This allows to validate whether the DML model missed any threshold effects or important nonlinearities.

Categorical treatment. To examine how these potential nonlinearities affect the treatment effect, we redefine the treatment as categorical and estimate the impact of 4 neighborhood types on travel-related CO_2 emissions. The types are pre-determined using k-means clustering based on the 5Ds. In this way, we can check whether the assumptions made for our baseline approach about the functional form of the influence of each built environment characteristic are reasonable and produce results similar to the categorical non-parametric estimation approach. It also serves as a robustness check that our baseline approach can handle the multicollinearity of treatments. We quantify the extent of multicollinearity using a variance inflation factor test. The disadvantage of the categorical approach is that we lose information about incremental changes in the built environment, disregard differences within clusters and fail to consider unique neighborhood types. This model specification will be referred to as "Household-level (categorical treatment)".

Neighborhood-level analysis. To address the issues of low explanatory power, which translates into a high degree of uncertainty when estimating the impact of the built environment (van Herick and Mokhtarian, 2020), we repeat the analysis at the neighborhood level, similar to Wagner et al. (2022) and Hankey and Marshall (2010). We average the travel behavior and preferences of all residents to estimate the effect of the built environment on the average travel-related CO_2 emissions per capita across all 190 neighborhoods. Averaging travel behavior has the advantage of mitigating the volatility of individual travel behavior that is due to the many unobserved, highly-personal factors that influence travel decisions on the given day. Consequently, a larger proportion of the observed differences in average travel behavior between neighborhoods can be explained with the independent variables, reducing the uncertainty of our estimates. This comes at the cost of limiting detailed insight into the confounding and moderating effects of travel preferences by obscuring the true distributions of covariates by their neighborhood means, partially violating the positivity assumption. This model specification will be referred to as "Neighborhood-level".

Environmental preferences. To gauge the impact of unobserved confounders, especially unobserved travel-related attitudes, we examine the sensitivity of our results in two ways. On the one hand, we examine the impact of including attitudes towards

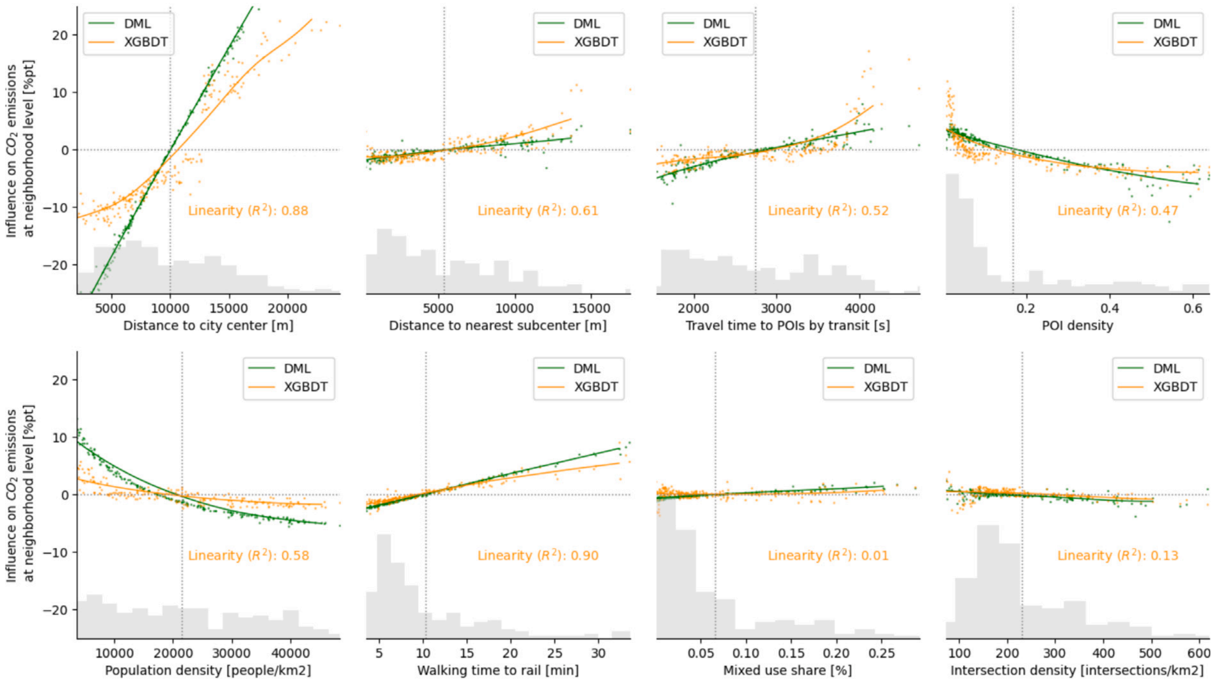

Fig. 10. Built environment effect similar in magnitude and form for DML and ML approaches. The group average treatment effect (GATE) of the Causal Forest-based DML approach is similar to the SHAP values of an extreme gradient boosting decision trees (GBDT), specifically XGBoost, averaged at the neighborhood level. The effect is given in percentage points (%pt) relative to the average per capita travel-related CO_2 emissions. The R^2 in the figure shows the fit of a linear regression for the neighborhood-level SHAP values, indicating the potential non-linear influence of the respective built environment characteristic. For illustrative purposes, a line has been fitted through the data using local tricube kernel-based regression to highlight trends in the effects. The dotted vertical line shows the mean for each characteristic (e.g. around 10 min average walking time to the nearest train station in Berlin).

environmental politics as an additional confounder for the neighborhood-level analysis and discuss whether these findings can be generalized to the household-level analysis. This model specification will be referred to as "Neighborhood-level (w/ environmental pref.)". On the other hand, we examine the impact that an unobserved confounder would need to have relative to the observed confounders in order to substantially alter or even eliminate the estimated effect. We calculate the partial R^2 of the treatment with the outcome, which, as shown by Cinelli and Hazlett (2020), corresponds to the degree to which confounders explaining all the residual outcome variation would have to be associated with the treatment to eliminate the effect.

C.2. Results

Our robustness checks suggest that the effect of the built environment at the neighborhood level is robust, as both the magnitude of the effect and the proportion of variability explained are similar across the different model specifications (e.g. see neighborhood GATE range and R^2 in Table 3 and in Fig. 12). This implies that the DML estimator can reasonably deal with multiple, partially correlated continuous treatments and a strong noise in the target variable. Also, the influence of the built environment is either approximately linear or sufficiently captured by the log transformation of the density variables and the moderating influence of the covariates.

Effect linearity. The estimated GATE at the neighborhood level is overall similar in magnitude and form to a non-linear effect estimate using extreme gradient boosting decision trees (GBDT) (see Fig. 10). The strongest differences are found for the influence of distance to city center and population density, where the DML approach estimates a larger effect. Furthermore, the emission-reducing effect of distance to the center does not decline for very short distances and approaches a threshold beyond which living closer to the center does not reduce emissions much, as would be expected and captured by the GBDT approach. For the density features, i.e. population, points of interest, and intersection density, the estimated functional form of both approaches is similar, indicating that the natural log transformation of these treatments was appropriate. Walking time to rail transit has the most linear influence according to the GBDT approach, followed by distance to city center. The influence of mixed-use land use and street connectivity is least linear and instead very noisy, showing almost no effect for both approaches.

Categorical treatment. Variance Inflation Factors (VIFs) across treatment variables indicate high multicollinearity (VIF > 10) for spatial lag variables, moderately high multicollinearity (5 < VIF ≤ 10) for all destination accessibility and density characteristics, and moderately low multicollinearity (1 < VIF ≤ 5) for design, diversity, and distance to transit characteristics (see Table 6).

Table 6
Treatment multicollinearity. Analysis of multicollinearity across treatment variables using Variance Inflation Factors (VIFs). We interpret a VIF above 10 as high multicollinearity, a VIF between 5 and 10 as moderately high multicollinearity, and a VIF below 5 as moderately low multicollinearity; however, the interpretation thresholds used in the literature differ to some extent.

5D's of compact development	Feature name	VIF
Destination accessibility	Distance to center	8.1
	Distance to subcenter	7.1
	Transit travel time	9.4
	POI density	7.3
	POI density (spatial lag)	10.4
Density	Population density	8.8
	Population density (spatial lag)	13.5
Diversity	Mixed-use share	2.5
Design	Street connectivity	2.4
Distance to transit	Walking time to rail	1.3

Table 7
Transport mode emission factors. Central estimates of life-cycle greenhouse gas emissions of urban transport modes per person km according to the International Transport Forum (ITF) (Cazzola and Crist, 2020). Emissions factors are expressed as CO_2 equivalents. Life-cycle emissions include a vehicle, fuel, and infrastructure component as well as operational services. ICE refers to internal combustion engine.

Mode	Emissions [gCO_2/pkm]				
	Vehicle	Fuel	Operational	Infrastructure	Total
Foot	0.0	0.0	0.0	0.0	0.0
Bike	7.5	0.0	0.0	9.5	16.9
E-bike	12.5	11.8	0.0	9.5	33.9
E-moped	9.8	19.7	0.0	11.3	40.9
E-scooter	26.2	6.2	0.0	9.4	41.8
Shared bike	23.3	0.0	24.7	9.5	57.5
Metro	2.0	52.6	0.0	11.0	65.6
E-bus	13.7	45.6	5.1	3.9	68.3
Moped	8.0	53.7	0.0	11.4	73.1
Bus	8.0	71.8	8.0	3.6	91.4
E-car	41.6	71.5	0.0	12.2	125.4
Car	23.9	125.6	0.0	12.5	162.0
Taxi	29.0	257.2	45.8	26.5	358.5

Despite this, the average effect of the built environment for the different neighborhood types is similar to the baseline approach at the household level, differing by at most by 3.7%pt (see Table 3), suggesting that multicollinearity does not substantially bias the average effect estimates. The share of variability explained at the neighborhood level (R^2) is slightly lower, which is expected due to the lower treatment granularity of 4 discrete neighborhood types compared to the continuous built environment characteristics. Residential self-selection into these 4 neighborhood types can be observed, but is rather weak with a F_1 score ranging from 0.10 to 0.50 depending on the neighborhood type, with an average of 0.36.

Neighborhood-level analysis. Using the neighborhood as the unit of analysis yields very similar treatment effect estimates to those obtained by conducting the analysis at the household level and then aggregating the effect. The R^2 is almost identical (see Table 3), and the effect distribution (see Fig. 12), its range, and the GATEs for the different neighborhood types are also very similar. However, the nuisance scores are considerably higher when the effect is estimated at the neighborhood level, suggesting that the average population differs significantly between neighborhoods and that these differences in e.g. socio-demographics can be consistently linked to preferences in the choice of residential built environment (R^2 of 0.48). Nevertheless, overall, the built environment is much more important for travel-related CO_2 emissions than residential self-selection (BEP of 84.2%), similar to the BEP range of our baseline approach (BEP between 71.8% and 92.2%).

Environmental preferences. When environmental policy preferences are added as an additional confounder at the neighborhood level, the nuisance R^2 increases from 0.48 to 0.59, but the BEP and treatment effect estimates do not change meaningfully. This suggests that car ownership and the other factors we use as proxies for unobserved travel-related attitudes already capture attitudes well and that environmental policy preferences cannot contribute much more.

Appendix D. Additional figures

See Figs. 11 and 12 and Tables 7–9.

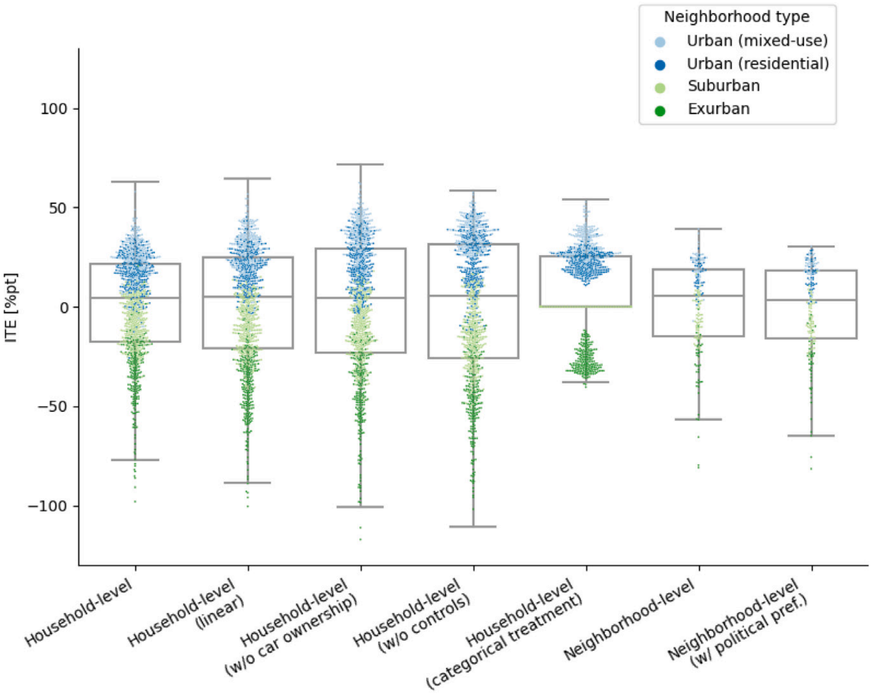

Fig. 11. Distribution of ITEs for all model specifications. The individual treatment effects (ITEs) for the built environment's influence on per capita travel-related CO_2 emissions are reported in percentage points (%pt) relative to the city average. "Household-level" is our preferred model specification. "Household-level (w/o controls)" refers to a model that does not control for residential self-selection. See Appendix C.1 for the definition of the remaining model specifications.

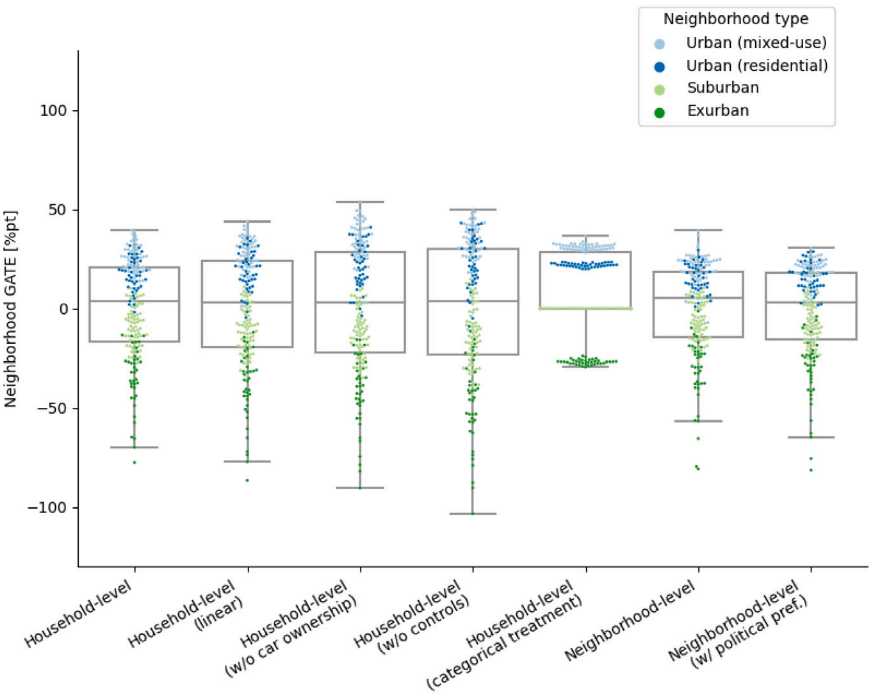

Fig. 12. Distribution of neighborhood GATEs for all model specifications. The group average treatment effects (GATEs) at the neighborhood level for the built environment's influence on per capita travel-related CO_2 emissions are reported in percentage points (%pt) relative to the city average. "Household-level" is our preferred model specification. "Household-level (w/o controls)" refers to a model that does not control for residential self-selection. See Appendix C.1 for the definition of the remaining model specifications..

Table 8
Descriptive statistics of preprocessed trip diary dataset. Mode share based on trip count and travel distance. Mean, standard deviation (std), 25%, 50%, and 75% percentiles of trip distance for different modes and trip purposes. Modes of less than 1% share were excluded. Trip modes and purposes are sorted by average travel distance.

		Share		Trip distance [km]				
		Count-based	Distance-based	Mean	Std	25%	50%	75%
Mode	Foot	0.30	0.05	0.9	0.9	0.3	0.6	1.0
	Bike	0.19	0.13	3.5	3.8	1.0	2.2	4.8
	Car	0.29	0.43	8.1	8.1	2.7	5.0	10.6
	Transit	0.21	0.38	9.9	8.0	4.0	8.0	13.5
Purpose	Companion	0.07	0.02	2.0	2.5	0.5	1.0	2.4
	Shopping	0.22	0.13	3.3	4.6	0.6	1.5	4.0
	School	0.10	0.08	4.2	5.5	1.0	2.1	5.0
	Other	0.23	0.21	4.9	6.7	0.7	2.1	6.1
	Leisure	0.21	0.20	5.2	6.7	1.0	2.9	6.5
	Work	0.18	0.36	11.1	8.7	5.0	9.0	15.0
All	–	–	–	5.4	7.0	1.0	2.8	7.0

Table 9
Moderating influence of covariates. Heterogeneity associated with covariates according to the conditional marginal average treatment effect (CMATE) of one standard deviation in the direction of a more sustainable built environment (e.g. higher population density and lower distance to city center). The share of heterogeneity explained by each individual covariate is calculated using the SHAP values of the final CausalForest model. Relative importance above 10% or with a heterogeneity above 25 percentage points (%pt) with respect to the CMATE are highlighted in bold.

Covariates	Associated heterogeneity		
	SHAP values-based heterogeneity	Covariate difference: 1st vs. 4th CMATE quartile	CMATE difference: 1st vs. 4th covariate quartile
Age	**50.9%**	−29.3	**+47.5%pt**
Income	**16.3%**	+2230	**−29.2%pt**
Car ownership	**14.2%**	+0.44	−14.0%pt
Employment status	6.2%	−0.85	**+40.9%pt**
Bike ownership	3.9%	+0.50	−18.8%pt
Transit subscription	3.2%	−0.03	+0.1%pt
Household type: single <65y	1.5%	+0.08	+1.5%pt
Higher education	1.3%	+0.24	−12.3%pt
Mobility constraints	1.2%	−0.02	+3.0%pt
Household size	0.6%	+1.24	−23.3%pt
Household type: single >65y	0.4%	−0.33	+4.0%pt
Driving license	0.2%	+0.15	−2.1%pt
Household type: family	0.0%	+0.46	−21.6%pt

Data availability

The authors do not have permission to share data.

References

van Acker, Veronique, Mokhtarian, Patricia L., Witlox, Frank, 2014. Car availability explained by the structural relationships between lifestyles, residential location, and underlying residential and travel attitudes. Transp. Policy 35, 88–99.
van Acker, Veronique, Witlox, Frank, 2010. Car ownership as a mediating variable in car travel behaviour research using a structural equation modelling approach to identify its dual relationship. J. Transp. Geogr. 18 (1), 65–74.
Aghaabbasi, Mahdi, Chalermpong, Saksith, 2023. Machine learning techniques for evaluating the nonlinear link between built-environment characteristics and travel behaviors: a systematic review. Travel Behav. Soc. 33, e100640.
Ao, Yibin, Yang, Dujuan, Chen, Chuan, Wang, Yan, 2019. Effects of rural built environment on travel-related CO2 emissions considering travel attitudes. Transp. Res. D 73, 187–204.
Aston, Laura, Currie, Graham, Delbosc, Alexa, Kamruzzaman, Md., Teller, David, 2021. Exploring built environment impacts on transit use–an updated meta-analysis. Transp. Rev. 41 (1), 73–96.
Athey, Susan, Tibshirani, Julie, Wager, Stefan, 2018. Generalized random forests. URL https://arxiv.org/abs/1610.01271.
Bagley, Michael N., Mokhtarian, Patricia L., 2002. The impact of residential neighborhood type on travel behavior: A structural equations modeling approach. Ann. Reg. Sci. 36, 279–297.
Bamberg, Sebastian, 2006. Is a residential relocation a good opportunity to change people's travel behavior? Results from a theory-driven intervention study. Environ. Behav. 38 (6), 820–840.
Barnett, David W., Barnett, Anthony, Nathan, Andrea, van Cauwenberg, Jelle, Cerin, Ester, 2017. Built environmental correlates of older adults' total physical activity and walking: a systematic review and meta-analysis. Int. J. Behav. Nutr. Phys. Activity 14 (1), 1–24.
Berrill, Peter, Nachtigall, Florian, Javaid, Aneeque, Milojevic-Dupont, Nikola, Wagner, Felix, Creutzig, Felix, 2024. Comparing urban form influences on travel distance, car ownership, and mode choice. Transp. Res. D 128, 104087.

Bhat, Chandra R., Guo, Jessica Y., 2007. A comprehensive analysis of built environment characteristics on household residential choice and auto ownership levels. Transp. Res. B 41 (5), 506–526.

Boeing, Geoff, 2017. OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. Comput. Environ. Urban Syst. 65, 126–139.

Cao, Xinyu, Fan, Yingling, 2012. Exploring the influences of density on travel behavior using propensity score matching. Environ. Plan. B: Plann. Des. 39 (3), 459–470.

Cao, Xinyu, Mokhtarian, Patricia L., Handy, Susan L., 2007a. Do changes in neighborhood characteristics lead to changes in travel behavior? A structural equations modeling approach. Transportation 34, 535–556.

Cao, Xinyu, Mokhtarian, Patricia L., Handy, Susan L., 2007b. Cross-sectional and quasi-panel explorations of the connection between the built environment and auto ownership. Environ. Plan. A 39 (4), 830–847.

Cao, Xinyu (Jason), Mokhtarian, Patricia L., Handy, Susan L., 2009. Examining the impacts of residential self-selection on travel behaviour: A focus on empirical findings. Transp. Rev. (ISSN: 0144-1647) 29 (3), 359–395. http://dx.doi.org/10.1080/01441640802539195.

Cao, Jason, Tao, Tao, 2023. Using machine-learning models to understand nonlinear relationships between land use and travel. Transp. Res. D 123, 103930.

Cao, Xinyu Jason, Xu, Zhiyi, Fan, Yingling, 2010. Exploring the connections among residential location, self-selection, and driving: Propensity score matching with multiple treatments. Transp. Res. A 44 (10), 797–805.

Cazzola, P., Crist, P., 2020. Good to Go? Assessing the Environmental Performance of New Mobility. OECD Publishing.

Cerin, Ester, Nathan, Andrea, van Cauwenberg, Jelle, Barnett, David W., Barnett, Anthony, 2017. The neighbourhood physical environment and active travel in older adults: a systematic review and meta-analysis. Int. J. Behav. Nutr. Phys. Activity 14 (1), 1–23.

Cervero, Robert, Kockelman, Kara, 1997. Travel demand and the 3Ds: Density, diversity, and design. Transp. Res. D 2 (3), 199–219.

Chang, Neng-Chieh, 2020. Double/debiased machine learning for difference-in-differences models. Econom. J. 23 (2), 177–191.

Chatman, Daniel G., 2009. Residential choice, the built environment, and nonwork travel: evidence using new data and methods. Environ. Plan. A 41 (5), 1072–1089.

Chen, Tianqi, Guestrin, Carlos, 2016. Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining. pp. 785–794.

Cheng, Long, Chen, Xuewu, De Vos, Jonas, Lai, Xinjun, Witlox, Frank, 2019. Applying a random forest method approach to model travel mode choice behavior. Travel Behav. Soc. 14, 1–10.

Cheng, Long, De Vos, Jonas, Zhao, Pengjun, Yang, Min, Witlox, Frank, 2020. Examining non-linear built environment effects on elderly's walking: A random forest approach. Transp. Res. D 88, 102552.

Chernozhukov, Victor, Chetverikov, Denis, Demirer, Mert, Duflo, Esther, Hansen, Christian, Newey, Whitney, Robins, James, 2018. Double/debiased machine learning for treatment and structural parameters.

Christiansen, Lars B., Madsen, Thomas, Schipperijn, Jasper, Ersbøll, Annette K., Troelsen, Jens, 2014. Variations in active transport behavior among different neighborhoods and across adult life stages. J. Transp. Health 1 (4), 316–325.

Cinelli, Carlos, Hazlett, Chad, 2020. Making sense of sensitivity: Extending omitted variable bias. J. R. Stat. Soc. Ser. B Stat. Methodol. 82 (1), 39–67.

Credit, Kevin, Lehnert, Matthew, 2023. A structured comparison of causal machine learning methods to assess heterogeneous treatment effects in spatial data. J. Geogr. Syst. 1–28.

Creutzig, Felix, Agoston, Peter, Minx, Jan C., Canadell, Josep G., Andrew, Robbie M., Quéré, Corinne Le, Peters, Glen P., Sharifi, Ayyoob, Yamagata, Yoshiki, Dhakal, Shobhakar, 2016. Urban infrastructure choices structure climate solutions. Nature Clim. Change 6 (12), 1054–1056.

Cui, Xuezhu, Zhuang, Caigang, Jiao, Zhenzhi, Tan, Zhangzhi, Li, Shaoying, 2023. How can urban built environment (BE) influence on-road (OR) carbon emissions? A road segment scale quantification based on massive vehicle trajectory big data. J. Transp. Geogr. 111, 103669.

Dai, Fangzhou, Diao, Mi, Sing, Tien Foo, 2020. Effects of rail transit on individual travel mode shares: A two-dimensional propensity score matching approach. Transp. Res. D 89, 102601.

De Vos, Jonas, Witlox, Frank, 2016. Do people live in urban neighbourhoods because they do not like to travel? Analysing an alternative residential self-selection hypothesis. Travel Behav. Soc. 4, 29–39.

Deng, Yiling, Zhao, Pengjun, 2022. The impact of new metro on travel behavior: Panel analysis using mobile phone data. Transp. Res. A 162, 46–57.

Dieleman, Frans, Wegener, Michael, 2004. Compact city and urban sprawl. Built Environ. 30 (4), 308–323.

Ding, Chuan, Cao, Xinyu, Liu, Chao, 2019. How does the station-area built environment influence Metrorail ridership? Using gradient boosting decision trees to identify non-linear thresholds. J. Transp. Geogr. 77, 70–78.

Ding, Chuan, Cao, Xinyu Jason, Næss, Petter, 2018. Applying gradient boosting decision trees to examine non-linear effects of the built environment on driving distance in Oslo. Transp. Res. A 110, 107–117.

Ding, Chuan, Wang, Yufan, Cao, Xinyu Jason, Chen, Yulin, Jiang, Yang, Yu, Bin, 2024. Revisiting residential self-selection and travel behavior connection using a double machine learning. Transp. Res. D 128, 104089.

Ding, Chuan, Wang, Donggen, Liu, Chao, Zhang, Yi, Yang, Jiawen, 2017. Exploring the influence of built environment on travel mode choice considering the mediating effects of car ownership and travel distance. Transp. Res. A 100, 65–80.

Duncan, Mitch J., Winkler, Elisabeth, Sugiyama, Takemi, Cerin, Ester, Dutoit, Lorinne, Leslie, Eva, Owen, Neville, 2010. Relationships of land use mix with walking for transport: do land uses and geographical scale matter? J. Urban Health 87, 782–795.

Eriksson, Ulf, Arvidsson, Daniel, Gebel, Klaus, Ohlsson, Henrik, Sundquist, Kristina, 2012. Walkability parameters, active transportation and objective physical activity: moderating and mediating effects of motor vehicle ownership in a cross-sectional study. Int. J. Behav. Nutr. Phys. Activity 9, 1–10.

Ewing, Reid, Cervero, Robert, 2010. Travel and the built environment: A meta-analysis. J. Am. Plan. Assoc. 76 (3), 265–294.

Ewing, Reid, Cervero, Robert, 2017. "Does compact development make people drive less?" the answer is yes. J. Am. Plan. Assoc. 83 (1), 19–25.

Ewing, R., Greenwald, M.J., Zhang, M., Walters, J., Feldman, M., Cervero, R., Thomas, J., 2009. Measuring the Impact of Urban Form and Transit Access on Mixed Use Site Trip Generation Rates—Portland Pilot Study. US Environmental Protection Agency, Washington, DC.

Ewing, Reid, Hamidi, Shima, Tian, Guang, Proffitt, David, Tonin, Stefania, Fregolent, Laura, 2018. Testing Newman and Kenworthy's theory of density and automobile dependence. J. Plann. Educ. Res. 38 (2), 167–182.

Fotheringham, A. Stewart, Wong, David W.S., 1991. The modifiable areal unit problem in multivariate statistical analysis. Environ. Plan. A 23 (7), 1025–1044.

Galster, George C., 2018. Nonlinear and threshold effects related to neighborhood: Implications for planning and policy. J. Plan. Lit. 33 (4), 492–508.

Gerike, Regine, Hubrich, Stefan, Ließke, Frank, Wittig, Sebastian, Wittwer, Rico, 2020. Sonderauswertung zum Forschungsprojekt "Mobilität in Städten–SrV 2018". Städtevergleich, Dresden.

Guan, Xiaodong, Wang, Donggen, 2019. Residential self-selection in the built environment-travel behavior connection: Whose self-selection? Transp. Res. D 67, 16–32.

Guan, Xiaodong, Wang, Donggen, Jason Cao, Xinyu, 2020. The role of residential self-selection in land use-travel research: a review of recent findings. Transp. Rev. (ISSN: 0144-1647) 40 (3), 267–287. http://dx.doi.org/10.1080/01441647.2019.1692965.

Guerra, Erick, Cervero, Robert, 2011. Cost of a ride: the effects of densities on fixed-guideway transit ridership and costs. J. Am. Plan. Assoc. 77 (3), 267–290.

Halleck Vega, Solmaria, Elhorst, J. Paul, 2015. The SLX model. J. Reg. Sci. 55 (3), 339–363.

Handy, Susan, 2018. Enough with the "D's" already—Let's get back to "A". Transf. Mag..

Handy, Susan, 2020. Is accessibility an idea whose time has finally come? Transp. Res. D 83, 102319.
Hankey, Steve, Marshall, Julian D., 2010. Impacts of urban form on future US passenger-vehicle greenhouse gas emissions. Energy Policy 38 (9), 4880–4887.
Heres-Del-Valle, David, Niemeier, Deb, 2011. CO2 emissions: Are land-use changes enough for California to reduce VMT? Specification of a two-part model with instrumental variables. Transp. Res. B 45 (1), 150–161.
van Holle, Veerle, Deforche, Benedicte, van Cauwenberg, Jelle, Goubert, Liesbet, Maes, Lea, van de Weghe, Nico, De Bourdeaudhuij, Ilse, 2012. Relationship between the physical environment and different domains of physical activity in European adults: a systematic review. BMC Public Health 12 (1), 1–17.
Hong, Jinhyun, 2017. Non-linear influences of the built environment on transportation emissions: Focusing on densities. J. Transp. Land Use 10 (1), 229–240.
Hu, Jiawei, Wagner, Felix, Ayargarnchanakul, Eva, Nachtigall, Florian, Milojevic-Dupont, Nikola, Lu, Chenxi, Creutzig, Felix, 2024. Examining the effects of urban form on integrated use of bike-sharing services and metro. Available at SSRN 4985690 http://dx.doi.org/10.2139/ssrn.4985690.
Hubrich, Stefan, Ließke, Frank, Wittwer, Rico, Wittig, Sebastian, Gerike, Regine, 2019. Methodenbericht zum forschungsprojekt.“Mobilität in städten–SrV 2018”.
Ibraeva, Anna, de Almeida Correia, Gonçalo Homem, Silva, Cecília, Antunes, António Pais, 2020. Transit-oriented development: A review of research achievements and challenges. Transp. Res. A 132, 110–130.
ITDP, Institute for Transportation & Development Policy, 2024. The atlas of sustainable city transport - Institute for transportation and development policy. URL https://itdp.org/publication/the-atlas-of-sustainable-city-transport/. (Accessed 25 June 2024).
Jaramillo, P., Kahn Ribeiro, S., Newman, P., Dhar, S., Diemuodeke, O.E., Kajino, T., Lee, D.S., Nugroho, S.B., Ou, X., Hammer Strømman, A., Whitehead, J., 2022. Transport. In: Shukla, P.R., Skea, J., Slade, R., Khourdajie, A. Al, van Diemen, R., McCollum, D., Pathak, M., Some, S., Vyas, P., Fradera, R., Belkacemi, M., Hasija, A., Lisboa, G., Luz, S., Malley, J. (Eds.), Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK and New York, NY, USA, http://dx.doi.org/10.1017/9781009157926.012, book section 10.
Javaid, Aneeque, Creutzig, Felix, Bamberg, Sebastian, 2020. Determinants of low-carbon transport mode adoption: systematic review of reviews. Environ. Res. Lett. 15 (10), 103002.
Kamruzzaman, Md., Washington, Simon, Baker, Douglas, Brown, Wendy, Giles-Corti, Billie, Turrell, Gavin, 2016. Built environment impacts on walking for transport in Brisbane, Australia. Transportation 43, 53–77.
Kitamura, Ryuichi, Mokhtarian, Patricia L., Laidet, Laura, 1997. A micro-analysis of land use and travel in five neighborhoods in the San Francisco Bay Area. Transportation 24, 125–158.
Klinger, Thomas, Lanzendorf, Martin, 2016. Moving between mobility cultures: what affects the travel behavior of new residents? Transportation (ISSN: 1572-9435) 43 (2), 243–271. http://dx.doi.org/10.1007/s11116-014-9574-x.
Kolak, Marynia, Anselin, Luc, 2020. A spatial perspective on the econometrics of program evaluation. Int. Reg. Sci. Rev. 43 (1–2), 128–153.
Kroesen, Maarten, 2019. Residential self-selection and the reverse causation hypothesis: Assessing the endogeneity of stated reasons for residential choice. Travel Behav. Soc. 16, 108–117.
Lin, Tao, Wang, Donggen, Guan, Xiaodong, 2017. The built environment, travel attitude, and travel behavior: Residential self-selection or residential determination? J. Transp. Geogr. (ISSN: 0966-6923) 65, 111–122. http://dx.doi.org/10.1016/j.jtrangeo.2017.10.004.
Liu, Jixiang, Wang, Bo, Xiao, Longzhu, 2021. Non-linear associations between built environment and active travel for working and shopping: An extreme gradient boosting approach. J. Transp. Geogr. 92, 103034.
Liu, Qingchun, Zhao, Peixiong, Zhang, Yingying, Zhang, Zhihui, Yang, Jun, 2022. Estimating the non-linear effects of urban built environment at residence and workplace on carbon dioxide emissions from commuting. Front. Public Health 10, 1077560.
Lucas, Karen, Bates, John, Moore, José, Carrasco, Juan Antonio, 2016. Modelling the relationship between travel behaviours and social disadvantage. Transp. Res. A 85, 157–173.
Lundberg, Scott M., Lee, Su-In, 2017. A unified approach to interpreting model predictions. Adv. Neural Inf. Process. Syst. 30.
Manaugh, Kevin, El-Geneidy, Ahmed, 2011. Validating walkability indices: How do different households respond to the walkability of their neighborhood? Transp. Res. D 16 (4), 309–315.
McCormack, Gavin R., Shiell, Alan, 2011. In search of causality: a systematic review of the relationship between the built environment and physical activity among adults. Int. J. Behav. Nutr. Phys. Activity 8, 1–11.
Microsoft Research, 2019. EconML: A Python package for ML-based heterogeneous treatment effects estimation. https://github.com/microsoft/EconML. Version 0.x.
Mokhtarian, Patricia L., Cao, Xinyu, 2008. Examining the impacts of residential self-selection on travel behavior: A focus on methodologies. Transp. Res. B 42 (3), 204–228.
Mokhtarian, Patricia L., van Herick, David, 2016. Quantifying residential self-selection effects: A review of methods and findings from applications of propensity score and sample selection approaches. J. Transp. Land Use 9 (1), 9–28.
Næss, Petter, 2009. Residential self-selection and appropriate control variables in land use: Travel studies. Transp. Rev. (ISSN: 0144-1647) 29 (3), 293–324. http://dx.doi.org/10.1080/01441640802710812.
Næss, Petter, 2014. Tempest in a teapot: The exaggerated problem of transport-related residential self-selection as a source of error in empirical studies. J. Transp. Land Use (ISSN: 19387849) 7 (3), 57–79, URL http://www.jstor.org/stable/26202693.
Næss, Petter, 2022. Meta-analyses of built environment effects on travel: No new platinum standard. J. Plann. Educ. Res. 42 (2), 199–205.
Næss, Petter, Stefansdottir, Harpa, Peters, Sebastian, Czepkiewicz, Michał, Heinonen, Jukka, 2021. Residential location and travel in the Reykjavik capital region. Sustainability 13 (12), 6714.
Oakes, J. Michael, Forsyth, Ann, Schmitz, Kathryn H., 2007. The effects of neighborhood density and street connectivity on walking behavior: the Twin Cities walking study. Epidemiol. Perspect. Innov. 4 (1), 1–9.
Open Data Berlin, 2023. Offene daten Berlin | Offene daten lesbar für mensch und maschine. Das ist das ziel. URL https://daten.berlin.de/. (Accessed 27 September 2023).
OpenStreetMap contributors, 2017. Planet dump. retrieved from https://planet.osm.org. https://www.openstreetmap.org.
Owen, Neville, Cerin, Ester, Leslie, Eva, Coffee, Neil, Frank, Lawrence D., Bauman, Adrian E., Hugo, Graeme, Saelens, Brian E., Sallis, James F., et al., 2007. Neighborhood walkability and the walking behavior of Australian adults. Am. J. Prev. Med. 33 (5), 387–395.
Parry, Ian W.H., Walls, Margaret, Harrington, Winston, 2007. Automobile externalities and policies. J. Econ. Lit. 45 (2), 373–399.
Pereira, Rafael Henrique Moraes, Nadalin, Vanessa, Monasterio, Leonardo, Albuquerque, Pedro H.M., 2013. Urban centrality: a simple index. Geogr. Anal. 45 (1), 77–89.
Richardson, Andrea S., Troxel, Wendy M., Ghosh-Dastidar, Madhumita B., Beckman, Robin, Hunter, Gerald P., DeSantis, Amy S., Colabianchi, Natalie, Dubowitz, Tamara, 2017. One size doesn't fit all: cross-sectional associations between neighborhood walkability, crime and physical activity depends on age and sex of residents. BMC Public Health 17, 1–11.
Roberts, David R., Bahn, Volker, Ciuti, Simone, Boyce, Mark S., Elith, Jane, Guillera-Arroita, Gurutzeta, Hauenstein, Severin, Lahoz-Monfort, José J., Schröder, Boris, Thuiller, Wilfried, et al., 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography 40 (8), 913–929.
Saelens, Brian E., Handy, Susan L., 2008. Built environment correlates of walking: a review. Med. Sci. Sports Exerc. 40 (7 Suppl), S550.

Saelens, Brian E., Sallis, James F., Frank, Lawrence D., 2003. Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures. Ann. Behav. Med. 25 (2), 80–91.

Scheiner, Joachim, 2014. Residential self-selection in travel behavior: Towards an integration into mobility biographies. J. Transp. Land Use 7 (3), 15–29.

Scheiner, Joachim, 2018. Transport costs seen through the lens of residential self-selection and mobility biographies. Transp. Policy (ISSN: 0967-070X) 65, 126–136. http://dx.doi.org/10.1016/j.tranpol.2016.08.012.

Scheiner, Joachim, Frank, Susanne, Gerwinat, Verena, Huber, Oliver, Næss, Petter, Schimohr, Katja, van Acker, Veronique, Wismer, Annika, 2024. In search of causality in the relationship between the built environment and travel behaviour. On the challenges of planning and realising an ambitious mixed-methods panel travel survey among relocating households in Germany. Prog. Plan. 182, 100820.

SenStadt, Senatsverwaltung für Stadtentwicklung, Bauen und Wohnen Berlin, 2023. Neue stadtquartiere. URL https://www.berlin.de/sen/stadtentwicklung/neue-stadtquartiere. (Accessed 04 June 2023).

Shao, Qifan, Zhang, Wenjia, Cao, Xinyu Jason, Yang, Jiawen, 2023. Built environment interventions for emission mitigation: A machine learning analysis of travel-related CO2 in a developing city. J. Transp. Geogr. 110, 103632.

Stevens, Mark R., 2017. Does compact development make people drive less? J. Am. Plan. Assoc. 83 (1), 7–18.

Tao, Tao, Wang, Jueyu, Cao, Xinyu, 2020. Exploring the non-linear associations between spatial attributes and walking distance to transit. J. Transp. Geogr. 82, 102560.

Tertoolen, Gerard, Van Kreveld, Dik, Verstraten, Ben, 1998. Psychological resistance against attempts to reduce private car use. Transp. Res. A 32 (3), 171–181.

TomTom, 2024. TomTom traffic index. URL https://www.tomtom.com/traffic-index/ranking/. (Accessed 25 June 2024).

Transportation Research Board, Institute of Medicine, 2005. Does the Built Environment Influence Physical Activity?: Examining the Evidence – Special Report 282. The National Academies Press, Washington, DC, http://dx.doi.org/10.17226/11203, URL https://nap.nationalacademies.org/catalog/11203/does-the-built-environment-influence-physical-activity-examining-the-evidence.

Tu, Meiting, Li, Wenxiang, Orfila, Olivier, Li, Ye, Gruyer, Dominique, 2021. Exploring nonlinear effects of the built environment on ridesplitting: Evidence from Chengdu. Transp. Res. D 93, 102776.

van de Coevering, Paul, Maat, Kees, Kroesen, Maarten, van Wee, Bert, 2016. Causal effects of built environment characteristics on travel behaviour: a longitudinal approach. Eur. J. Transp. Infrastruct. Res. 16 (4).

van de Coevering, Paul, Maat, Kees, van Wee, Bert, 2015. Multi-period research designs for identifying causal effects of built environment characteristics on travel behaviour. Transp. Rev. 35 (4), 512–532.

van de Coevering, Paul, Maat, Kees, van Wee, Bert, 2018. Residential self-selection, reverse causality and residential dissonance. A latent class transition model of interactions between the built environment, travel attitudes and travel behavior. Transp. Res. A 118, 466–479.

van Herick, David, Mokhtarian, Patricia L., 2020. How much does the method matter? An empirical comparison of ways to quantify the influence of residential self-selection. Travel Behav. Soc. (ISSN: 2214-367X) 18, 68–82. http://dx.doi.org/10.1016/j.tbs.2019.09.002.

van Wee, Bert, Cao, Xinyu Jason, 2022. Residential self-selection in the relationship between the built environment and travel behavior: A literature review and research agenda. Adv. Transp. Policy Plan. 9, 75–94.

VBB, Verkehrsverbund Berlin-Brandenburg GmbH, 2023. GTFS data Berlin-Brandenburg (VBB). URL https://www.vbb.de/vbb-services/api-open-data/datensaetze/. (Accessed 27 September 2023).

Wagner, Felix, Milojevic-Dupont, Nikola, Franken, Lukas, Zekar, Aicha, Thies, Ben, Koch, Nicolas, Creutzig, Felix, 2022. Using explainable machine learning to understand how urban form shapes sustainable mobility. Transp. Res. D 111, 103442.

Wagner, Felix, Nachtigall, Florian, Franken, Lukas, Milojevic-Dupont, Nikola, Pereira, Rafael H.M., Koch, Nicolas, Runge, Jakob, Gonzalez, Marta, Creutzig, Felix, 2023. A causal discovery approach to learn how urban form shapes sustainable mobility across continents. arXiv preprint arXiv:2308.16599.

Wang, Donggen, Lin, Tao, 2019. Built environment, travel behavior, and residential self-selection: A study based on panel data from Beijing, China. Transportation 46 (1), 51–74.

Wolday, Fitwi, Cao, Jason, Næss, Petter, 2018. Examining factors that keep residents with high transit preference away from transit-rich zones and associated behavior outcomes. J. Transp. Geogr. 66, 224–234.

Wolday, Fitwi, Naess, Petter, Cao, Xinyu Jason, 2019. Travel-based residential self-selection: A qualitatively improved understanding from Norway. Cities 87, 87–102.

Wu, Xinyi, Tao, Tao, Cao, Jason, Fan, Yingling, Ramaswami, Anu, 2019. Examining threshold effects of built environment elements on travel-related carbon-dioxide emissions. Transp. Res. D 75, 1–12.

Xu, Donna, Shi, Yaxin, Tsang, Ivor W., Ong, Yew-Soon, Gong, Chen, Shen, Xiaobo, 2019. Survey on multi-output learning. IEEE Trans. Neural Netw. Learn. Syst. 31 (7), 2409–2429.

Xu, Yiming, Yan, Xiang, Liu, Xinyu, Zhao, Xilei, 2021. Identifying key factors associated with ridesplitting adoption rate and modeling their nonlinear relationships. Transp. Res. A 144, 170–188.

Yang, Shuo, Fan, Yingling, Deng, Wei, Cheng, Long, 2019. Do built environment effects on travel behavior differ between household members? A case study of Nanjing, China. Transp. Policy 81, 360–370.

Zhang, Ming, Kukadia, Nishant, 2005. Metrics of urban form and the modifiable areal unit problem. Transp. Res. Rec. 1902 (1), 71–79.

Zhang, Yiwen, Zhou, Wenzhu, Ding, Jiayi, 2022. Effects of the built environment on travel-related co2 emissions considering travel purpose: a case study of resettlement neighborhoods in Nanjing. Buildings 12 (10), 1718.

Zhou, Bin, Kockelman, Kara M., 2008. Self-selection in home choice: Use of treatment effects in evaluating relationship between built environment and travel behavior. Transp. Res. Rec. 2077 (1), 54–61.

Zhou, Meng, Wang, Donggen, Huang, Sixian, Zhou, Jun, Guo, Li, 2024. Evaluating the impact of rail transit network expansion on travel behavior in Shenzhen, China: A causal analysis across different stages of development. Transp. Res. D 132, 104246.