A joint demand modeling framework for ride-sourcing and dynamic ridesharing services: a geo-additive Markov random field based heterogeneous copula framework

Behram Wali[1] · Paolo Santi[1,2] · Carlo Ratti[3]

Abstract

Promoting sustainable transportation, ride-sourcing and dynamic ridesharing (DRS) services have transformative impacts on mobility, congestion, and emissions. As emerging mobility options, the demand for ride-sourcing and DRS services has rarely been simultaneously examined. This study contributes to filling this gap by jointly analyzing the demand for ride-sourcing and DRS services and examining how it varies across neighborhood-level built environment, transit accessibility and crime, behavioral, and sociodemographic factors. To achieve these objectives, unique geo-coded data containing millions of ride-sourcing and DRS trips in Chicago are spatially joined with up-to-date data on the built environment, transit accessibility, crime, active travel, and demographic factors. A novel Markov Random Field-based joint heterogeneous geo-additive copula framework is presented to simultaneously capture random, systematic, and spatial heterogeneity. Characterized by a Frank copula structure, the demand for ride-sourcing and DRS services exhibited a non-linear stochastic dependence pattern. With spatial heterogeneity and spillover effects, the stochastic dependence of ride-sourcing and DRS demand varied across time of day and was the strongest in compact and dense neighborhoods. Key aspects of the built environment related to urban design (pedestrian-oriented infrastructure), density, and land-use diversity were positively associated with ride-sourcing and DRS demand—suggesting that sustainable mobility goals can be achieved by continuing to invest in more walkable neighborhoods. Active travel and telecommuting were positively linked with ride-sourcing and DRS demand. Complementary and substitutive effects for transit accessibility were found. Results show that increasing transit accessibility in areas with low levels of accessibility (compared to those with high transit levels) could be more helpful in increasing the adoption of ride-sourcing and DRS services. Relative to ride-sourcing, the demand for DRS services appeared more responsive to improvements in pedestrian-infrastructure and transit accessibility. Quantification of non-linear associations with ceiling and overdose effects for the built environment, vehicle ownership, and transit accessibility provided deeper insights. The findings can help guide the development of policy interventions and investment decisions to further accelerate the adoption of mobility-on-demand systems.

Extended author information available on the last page of the article

Introduction

Since the widespread developments in Information and Communication Technologies (ICT), there has been a marked shift in the attitudes of end-users towards shared consumption and the ability to co-produce services. The proliferation of digital technologies has accelerated peer-to-peer based interactions enabling distributed/communal consumption. The emergence of shared ownership-based models across different sectors has enabled citizens to access goods and services without the burden of private ownership. With successful business models already developed across transportation, tourism, education, finance and food sectors, collaborative consumption is one of the ten disruptive ideas that will change the planet with a global market amounting to over $300 Billion by 2025 (Walsh 2011; Coopers 2015).

Among other sectors, transportation has benefitted and will continue to benefit, the most from the emergence of collaborative consumption and services. Coupled with the advances in ICT, the shared economy model has led to the advent of mobility-on-demand services that have the potential to contribute to reaching several United Nation's Sustainable Development Goals (SDGs) including sustainable cities and communities, responsible consumption and production, and climate action (Costanza et al. 2016, United Nations 2021). Enabled by digital technology, smart phones and connectivity, ride-sourcing, ride-splitting (or pooled rides), and car-sharing have emerged as key mobility-on-demand services and have witnessed tremendous growth in the last decade (Morency 2007; Dias et al. 2017; Stathopoulos and Sener 2017; Yu et al. 2017; Cui et al. 2020; Devaraj et al. 2020; Gomez et al. 2021; Su et al. 2021). Ride-sourcing services (e.g., Uber) match the drivers with riders using smartphone apps and sophisticated traffic assignment and routing algorithms—equipping drivers of personal vehicles to monetize their skills and resources while providing more accessible and flexible mobility options to the riders. While ride-sourcing shares many similarities with taxi service, it provides more flexible and accessible mobility optimized by sophisticated assignment/routing algorithms packaged through smartphones. Also, research has shown that taxi and ride-sourcing services differ in terms of the characteristics of wait times, the users, and the trips they serve (Rayle et al. 2016). The on-demand services have further expanded to ridesharing, also called ride-splitting and dynamic ride sharing (DRS), where multiple passengers following a similar route share the vehicle (with reduced fares) at the same time (as opposed to sharing the vehicle at different times). These emerging mobility-on-demand services have significant implications for future urban transport systems, including lower private car ownership, reduced congestion, fuel consumption, and climate change, and providing flexible accessibility in non-transit serving areas (Anderson 2014; Santi et al. 2014; Shaheen and Chan 2016; Tachet et al. 2017; Sperling 2018; Stiglic et al. 2018; Zhang and Zhang 2018; Deka and Fei 2019; Ratti and Santi 2019; Wang and Yang 2019; Tu et al. 2020; Yan et al. 2020). The concept of shared mobility and on-demand services are also considered key to the successful adoption and deployment of automated vehicles (Stocker and Shaheen 2017). We note that despite the many potential benefits mentioned above, the overall effect of on-demand mobility on urban traffic and the city is still not clear and highly debated. For instance, on-demand mobility services can

lead to a reduction in transit ridership and may impact population health by substituting short-distance trips currently undertaken through active travel modes (Chen et al. 2019). Thus, more empirical research is needed to better understand the demand for on-demand mobility services and their impacts on transport infrastructure.

Despite the widespread deployment of on-demand mobility services, empirical research on the adoption and spatial distribution of disruptive shared mobility services is in its infancy. Large gaps exist regarding our knowledge of the user and neighborhood characteristics associated with the use of ride-sourcing and DRS services—impeding investment decisions and implementation of policy/regulative measures. The built environment has remained a key policy lever to shape and provide active transportation supportive infrastructure. Can we continue to invest in compact and more walkable neighborhoods to achieve active travel and sustainable mobility-related goals? Likewise, the potential complementary and/or substitutive role of transit accessibility as it jointly relates to ride-sourcing and DRS demand is rarely examined. From a behavioral standpoint, little is known about how individuals' participation in active travel (walk, bike, and transit) correlate with the use of ride-sourcing and DRS services.[1] From a transportation and public health standpoint, examination of relationships between active travel and demand for emerging mobility services is important to understand how population-level physical activity may increase (or decrease) with technology-enabled travel options (Tribby et al. 2020). Answers to such questions can help guide the development of regulations, policies, and investment decisions to capitalize on the benefits of on-demand mobility systems while minimizing their negative societal impacts.

To answer such questions, this study harnesses a unique publicly available geo-coded data source containing millions of ride-sourcing and DRS trips in Chicago. The large-scale trip-making data are spatially joined with detailed neighborhood-level built environment, behavioral, and sociodemographic data. The study contributes by jointly examining the demand for ride-sourcing and DRS services and how it varies across neighborhood-level behavioral, built environment, and socio-demographic factors. *Methodologically*, we present a novel and comprehensive framework for examining the determinants of ride-sourcing and DRS usage simultaneously capturing multiple layers of systematic, random, and spatial heterogeneity (discussed later in detail).

The rest of the paper is structured as follows. Section "Literature review" synthesizes the relevant literature along with identifying key conceptual and methodological gaps. The methodological framework and different data streams are explained in Sect. "Methodology". Spatial distributions of ride-sourcing and DRS outcomes are discussed in Sect. "Results" along with a presentation of the results of empirical models. We discuss and synthesize the key findings in Sect. "Discussion and synthesis". Limitations and avenues for future research are discussed in Sect. "Limitations and future work", followed by conclusions of the study with a discussion of policy implications in Sect. "Closing remarks".

---

[1] Compared to walking and biking, transit is not a fully active travel mode. While transit users typically undertake more walking and biking, they also undertake some sedentary travel at least for the part where the user is on-board a bus/train. However, for simplicity, we refer to walk, bike, and transit as active travel modes throughout the paper.

Literature review

Previous studies are mostly based on individual-level stated/revealed preference survey data, interviews, and focus groups, asking travelers about their travel behavior/preferences of the use of mobility-on-demand services. More recently, a handful of studies have examined spatiotemporal data related to on-demand mobility services as public agencies in some US cities and elsewhere have publicly disseminated such data collected by service providers (e.g., Uber, Lyft). We synthesize the key literature using the two data sources next.

Stated/revealed preference-based survey data

Given the profound impacts of on-demand mobility services highlighted earlier, a broad spectrum of studies has used focus groups and stated/revealed preference-based survey data to examine the determinants of ride-sourcing and DRS travel demand (Chan and Shaheen 2012; Rayle et al. 2016; Clewlow and Mishra 2017; Dias et al. 2017; Alemi et al. 2018a, 2018b; Sikder 2019; Azimi et al. 2020; Wali 2022). Rayle et al. (2016) analyzed data from 380 intercept surveys collected in three ride-sourcing hotspots in San Francisco and found wait times for ride-sourcing to be more reliable and shorter than taxi services. A significant substitution effect of on-demand mobility vs. taxi was also observed (Rayle et al. 2016). A relatively greater substitution effect was observed for non-taxi-related trips with over half of ride-sourcing trips substituting other non-taxi modes. Contrarily, Sikder (2019) reported a complementary effect of ride-sourcing services on public transit (Sikder 2019). A host of sociodemographic factors are known to be correlated with the demand for ride-sourcing services. Early ride-sourcing adopters are reported to be high-income, more educated, and younger (Clewlow and Mishra 2017; Dias et al. 2017; Alemi et al. 2018a, 2018b). Racial differences are also observed in the adoption of ride-sourcing services—with Black individuals less likely to adopt ride-sourcing services (Sikder 2019). Attitudes and behaviors also exhibit an influence on ride-sourcing adoption. Individuals with high technological savviness and pro-environmental attitudes (such as ownership of plug-in electric vehicles and using active travel modes) tend to use ride-sourcing services more often (Alemi et al. 2018a; Wali 2022). Compared to ride-sourcing, literature on the determinants of dynamic ridesharing (DRS) is scarce. Lavieri et al. (2018) examined subjective and objective factors associated with the adoption of ride-hailing trips and found that pooled ride-hailing substituted public transit and active travel modes. Older individuals and women had a lower propensity to choose a pooled ride-hailing mode (Kang et al. 2021).

Spatiotemporal disaggregate trip data

With the availability of spatiotemporal disaggregate trip-level data, a handful of studies have developed models to understand the demand for mobility-on-demand services. Using trip-level data from Austin, demand for ride-sourcing and its variations across sociodemographic and neighborhood characteristics has been analyzed at the Traffic Analysis Zone (TAZ) (Lavieri et al. 2018) and block group level (Yu and Peng 2020). Ride sourcing demand was negatively associated with transit supply and the reverse was found for TAZs with greater retail employment density and presence of parks (Lavieri et al. 2018). Similarly, Yu and Peng (2020) found positive associations between different measures of densities (population, employment, and service job) and ride-sourcing demand (defined as an aggregated count of ride-sourcing trips per trip origin block group) (Yu and Peng

2020). Besides Austin, few studies have also examined the demand for ride-sourcing in Chicago. Marquet (2020) developed a truncated Poisson model on around 32 million ride-sourcing trips in the city of Chicago and reported a positive correlation between walkability (at trip origins and destinations) and ride-sourcing demand (Marquet 2020). Using the same Chicago dataset, Ghaffar et al. (2020) estimated a random-effects negative binomial model founding positive associations of population/employment density with ride-sourcing demand. For a comprehensive review of ride-sourcing studies using trip-level data, see Ghaffar et al. (2020).

Compared to the broad literature on ride-sourcing demand, only a few studies have examined the demand for shared rides using disaggregate trip-level data with a focus on built environment features. This could be since DRS is a newer mobility-on-demand service available only in selected cities. The Transportation Network Company (TNC) data made available by the City of Chicago is unique in that it contains millions of DRS trips. Focusing on the demand for DRS, Dean and Kockelman (2021) and Xu et al. (2021) examined built environment determinants of DRS demand—reporting conflicting results compared to the associations of the built environment with ride-sourcing demand discussed earlier. Using spatial error models, Dean and Kockelman (2021) reported negative associations between population density, pedestrian network density, and the proportion of shared trips in Chicago. Likewise, pedestrian network density was negatively correlated with the count of shared trips. With a focus on ride-splitting adoption rate, Xu et al. (2021) used random forest models for examining non-linearity in the built environment and sociodemographic factors (Xu et al. 2021). Several key built environment variables including walk score, population density, and employment were negatively correlated with shared rides but the authors noted that the relationships were rather weak.

Research gaps

Previous studies have provided valuable insights regarding the determinants of ride-sourcing or DRS demand using disaggregate trip-level data. However, important gaps remain. ***First***, no study (to our knowledge) has jointly examined the demand for ride-sourcing and DRS usage and how the demand for these two shared mobility services varies across the built environment, transit accessibility, behavioral, and sociodemographic factors. The demand for ride-sourcing and DRS can be correlated due to the presence of common observed and unobserved factors (Washington et al. 2003; Dias et al. 2017; Wali et al. 2021b). For example, older individuals or individuals with lower technological savviness could be less likely to use either of the two shared mobility options. On the contrary, individuals with sustainable attitudinal predispositions are more likely to participate in ride-sourcing and DRS services. Ride-sourcing is also routinely offered where DRS is offered. Thus, understanding the joint dependency between the two services is fundamental. ***Second***, no information exists about the nature (linear vs. non-linear) of the stochastic dependence between DRS and ride-sourcing services, and whether spillover effects exist in the spatial distribution of the stochastic dependence between the two emerging mobility options. A spatially referenced characterization of the stochastic dependence between the two on-demand services can better assist with the development of place-based regulations, policies, and future investment decisions. ***Third***, almost all previous studies have focused on linear effects of the built environment and related factors and did not examine systematic heterogeneity in the effects of behavioral, built environment, and demographic correlates as it relates to the demand for ride-sourcing and DRS services. Analysis of systematic

heterogeneity (non-linearity) intuitively and rigorously can shed light on the potential ceiling and overdose effects—enabling engineers and planners to identify strategies with greater impact to enhance ride-sourcing and DRS use.

Research objective and contribution

In view of the above gaps, the present study makes both conceptual and methodological contributions. The study presents a joint analysis of the demand for ride-sourcing and DRS services and how it varies across neighborhood-level built environments, transit accessibility, behavioral, and sociodemographic factors. A tractable pathway is presented from the neighborhood environments where we live to explain systematic differences across population cohorts in the spatial distribution of ride-sourcing and DRS usage.

From a methodological perspective, the study contributes by simultaneously capturing multiple layers of systematic, random, and spatial heterogeneity in a unified modeling framework. A comprehensive and novel geo-additive Markov Random Field-based heterogeneous copula framework is presented to simultaneously address the issues outlined in the previous section. A broad suite of (survival) Archimedean and elliptical copulas is harnessed to model the complex stochastic dependence between demand for ride-sourcing and DRS services. Within the joint copula framework, a Markov Random Field spatial model is used to capture spatial spillover effects in the joint stochastic dependence pattern. Finally, to unveil systematic heterogeneity in the effects of the behavioral, built environment, and sociodemographic factors, we integrate Generalized Additive modeling within the spatial copula-based framework enabling us to examine non-linearity with sensitivity towards the potential ceiling and/or overdose effects. To our knowledge, no previous study has simultaneously accounted for the aforementioned issues in a methodologically rigorous and unified manner. To achieve the objectives, a large-scale dataset containing over 40 million ride-sourcing and DRS trips in Chicago is spatially joined with neighborhood behavioral and sociodemographic profiles. Detailed built environment data are extracted from the most recent US EPA's 2020 Smart Location Database.

Methodology

Data

The study used several data sources to compile information on the demand for ride-sourcing and dynamic ridesharing services, participation in active travel, built environment, transit accessibility and crime, and sociodemographic factors. Regarding the demand for ride-sourcing and DRS services, TNC based ride-sourcing and DRS data published by the City of Chicago was used (https://data.cityofchicago.org/Transportation). Starting from November 2018, the City of Chicago published an online portal of ride-sourcing and DRS trips made available by TNC companies every quarter. As of June 2021, the digital repository contains data on over 190 million trips. Detailed trip-level attributes are made available including start and end time, fare, duration, length, and pick-up/drop-off locations. As discussed elsewhere {Dean, 2021 #124}, the city removed Personally Identifiable Information (PII) by spatially aggregating the pick-up/drop-off locations at the census-tract level. Likewise, timestamps were rounded to the nearest quarter (15 min) interval. The spatial aggregation at the census-tract level means that we only have data on which census tracts

the pick-ups and drop-offs occurred.[2] Note that these data are spatially referenced at the census-tract level and individuals who undertook the trips cannot be identified.[3] This study collected data for six months starting from November 2018 and ending in April 2019. Trips outside the boundary of City of Chicago and those with missing locations for pick-up/drop-off, missing fare, duration/length were removed. To capture routine travel demand patterns, trip lengths greater than the 99.9th percentile value (27 miles) were removed. The final cleansed dataset contained over 33 million trips (27,427,395 ride-sourcing trips and 5,686,013 ridesharing trips) distributed over 793 census tracts in Chicago.

To obtain information on the built environment and land-use patterns, the trip-level data were spatially joined (at the census tract level) with objectively assessed environmental data extracted from the most recent 2020 Smart Location Database (SLD) by U.S. Environmental Protection Agency (http://urbandesign4health.com/projects/sld-update). Compared to the last SLD version (assembled in 2014) used in all relevant studies on this topic, the 2020 SLD exhibits significantly greater nationwide transit coverage and is based on an improved assessment of the built environment and land-use variables. By using the most recent SLD version, the present study captures an up-to-date profile of built environment, land-use, and transit accessibility in Chicago. Neighborhood (census-tract) level data on different types of crime for the city of Chicago were collected from the Crime Open Database (CODE) (https://osf.io/zyaqn/). Three years of crime data from CODE (2016 to 2018) were used to capture the potential intermediate and lag impacts of different types of crime on ride-sourcing and DRS use.

The trip and built environment data were complemented by information on active travel-related behaviors and sociodemographic factors. Census-tract level socioeconomic variables (race, gender, income, unemployment, education, internet accessibility) were calculated from the "Detailed Tables" in the American Community Survey (ACS) using the population counts for each category of socioeconomic variables. To capture the commuting patterns in each neighborhood, responses to journey to work-related questions were used to calculate neighborhood-level participation rates in telecommuting and active travel (transit, walk, bike) (ACS 2019). Telecommuting (also referred to as teleworking) refers to working from home through the use of digital technology (ACS 2019). The journey to work-related questions captures commute mode choice including car, truck, van, taxicab, motorcycle, transit, light rail, subway/elevated rail, walk, and bicycle. Note that the ACS journey to work question does not include ride-sourcing or DRS travel modes. The most recent five-year estimates from the ACS were used to increase the statistical reliability of the neighborhood-level socioeconomic and travel behavior measures.[4]

---

[2] Census tracts are small, relatively permanent, polygons defined by the US Census Bureau covering a well-defined geographic area. Census tract is second smallest geography (beside census block group) for which consistent and uniform census data are publicly made available by the US Census Bureau. On average, census-tracts contain around 4000 inhabitants and provide greater spatial granularity with around 73,000 census tracts nationwide compared to only 43,000 ZIP codes in the United States.

[3] Note that socio-demographic information about the ride-sourcing and DRS users is not available in these data. Since census tract serves as the unit of analysis, we integrate the ride-sourcing and DRS data with publicly available most recent census tract level built environment and sociodemographic data described next.

[4] Note that the American Community Survey (ACS) is different from the Decennial Census survey that is conducted once every decade. Compared to the ACS, Decennial survey asks a limited set of questions including age, sex, race, and home-ownership status. The ACS data are used in this study since it is conducted more frequently (monthly) and provides information on a much broader set of factors described below. For differences between the ACS and the Decennial Census, see https://www.census.gov/programs-surveys/acs/about/acs-and-census.html.

The two dependent variables of interest are related to the neighborhood-level demand of ride-sourcing and DRS services. The two dependent variables are specified as the count of ride-sourcing trips per 100 inhabitants in a census tract (ride-sourcing rate) and the count of dynamic ridesharing trips per 100 inhabitants (dynamic ridesharing rate). Distributions of dependent and independent variables are shown and discussed in Sect. "Results".

Methodological framework

At a basic level, the two response outcomes can be independently modeled via two univariate linear regression models given the continuous nature of the two dependent variables. While (statistically) consistent, univariate models have several shortcomings limiting our understanding of the complex mechanisms through which built environment and sociodemographic factors may influence the demand for DRS and ride-sourcing services. First, using independent multivariate models cannot capture the joint dependence between the demand for ride-sourcing and DRS services arising due to correlations among unobserved factors (*random heterogeneity*). To gain a deeper understanding, capturing the joint dependence is important especially when the two mobility services are significantly correlated (as we demonstrate later). Second, the traditional joint model does not shed light on the nature (linear vs. non-linear) of the stochastic dependence and the extent of spillover effects in the spatial distribution of the joint dependence between DRS and ride-sourcing services *(spatial heterogeneity)*. Finally, it ignores potential non-linear impacts (with ceiling and overdose effects) of exogenous variables on the demand for ride-sourcing and DRS services *(systematic heterogeneity)*.

In this section, we present a novel and comprehensive modeling framework to jointly model the demand for ride-sourcing and DRS services and to simultaneously examine random (unobserved), spatial, and systematic (observed) heterogeneity in the context of two emerging mobility services. For brevity, we directly start with the exposition for the joint modeling framework and do not discuss the univariate models. Assume $\aleph(Y_1, Y_2|X_1, X_2)$ denote the joint cumulative distribution function (CDF) for ride-sourcing (Y_1) and DRS demand (Y_2) conditional on vectors of exogenous variables associated with ride-sourcing and DRS demand (X_1, X_2). The conditioning of the joint distribution on exogenous variables is fully general and X_1 and X_2 can contain a different set of independent variables. With a different set of independent variables impacting the marginal densities underlying the joint distribution, a richer model structure can be developed. For example, statistically insignificant exogenous factors in the specification for one of the two marginal densities can be easily dropped. Using a copula framework, the joint cumulative distribution function can be expressed as a combination of marginal cumulative distributions functions such as (Nelsen 2007; Trivedi and Zimmer 2007; Sener et al. 2010):

$$\aleph(Y_1, Y_2|X_1, X_2) = C_\vartheta(\aleph_1(Y_1|X_1), \aleph_2(Y_2|X_2); \vartheta) \tag{1}$$

where: $\aleph_1(Y_1|X_1)$ and $\aleph_2(Y_2|X_2)$ indicate the marginal CDFs of ride-sourcing (Y_1) and DRS demand (Y_2) taking values on a uniform [0,1] grid conditioned on the set of exogenous variables (X_1, X_2); C_ϑ is a uniquely defined copula device that tracks the stochastic dependence between Y_1 and Y_2; and ϑ is the copula parameter measuring the (linear/nonlinear) stochastic dependence between ride-sourcing and DRS demand. The copula-based parameterization in Eq. (1) is extremely effective because it separates the modeling of

stochastic dependence from the statistical modeling of the underlying (univariate) distributions characterizing the joint distribution (Nelsen 2007).

Copula devices for joint stochastic dependence modeling

A broad spectrum of Archimedean and elliptical copulas is employed to model the neighborhood-level stochastic dependence between ride-sourcing and DRS demand. Among the elliptical class of copulas, Gaussian and student-*t* copulas are implemented. Both copulas can capture positive as well as negative dependence structures, thus termed comprehensive copulas (Bhat and Eluru 2009), but student-*t* copula exhibits fatter tails enabling modeling of tail dependence patterns. Contrarily, the (often) undesirable property of asymptotic independence is rooted in Gaussian copula—implying that extreme tail events (low levels of ride-sourcing and DRS demand, and vice versa) are independent. Both are also symmetric copulas since the stochastic dependence in the lower and upper distribution tails tend to be equal. To explore moderate levels of stochastic dependence patterns (both positive and negative), Farlie–Gumble–Morgenstern (FGM) and Plackett copulas are implemented. While useful, the above copulas can only capture linear form of stochastic dependence patterns. Archimedean copulas (including the Clayton, Joe, Gumbel, and Frank copulas) are another popular class of mathematical structures able to model different forms of non-linear stochastic dependence patterns. For a graphical illustration of the dependency structures implied by the different copulas discussed above, its mathematical expositions, and associated generator functions, see (Bhat and Eluru 2009; Wali et al. 2018). Finally, to model inverted dependence patterns, we employ survival versions of asymmetric Archimedean copulas (Clayton, Joe, and Gumbel). With a rotation of 180 degrees, the survival asymmetric copulas allow capturing positive dependence structures but with inverted tail dependencies compared to the unrotated copulas. For example, a standard (unrotated) Clayton copula is driven by a positive dependence but with the dependence in the left tail stronger compared to the dependence in the right tail. A 180 degrees rotation of the Clayton copula allows positive dependence but with weaker left tail and stronger right tail dependence. Note that the right tail dependence structure in survival Clayton copula is stronger than an unrotated Gumbel copula. For details on survival copulas, see (Wali et al. 2018).

Marginal distribution modeling

Independent of the modeling of stochastic dependence patterns, the marginal distributions for ride-sourcing (Y_1) and DRS (Y_2) demand are characterized through cumulative and probability distribution functions denoted by $\aleph_\kappa(Y_\kappa|\mu_\kappa, \sigma_\kappa)$ and $f_\kappa(Y_\kappa|\mu_\kappa, \sigma_\kappa)$, respectively, for $\kappa = 1$ (ride-sourcing demand) and $\kappa = 2$ (DRS demand) (Nelsen 2007); μ_κ and σ_κ represent the location (mean) and shape (dispersion) parameters of the two marginal distributions. In this study, we model the marginal distributions of the two outcomes with (log) normal densities (discussed further in Sect. "Results").

Specification of exogenous variables

As demonstrated elsewhere (Wali and Khattak 2022), to capture systematic and spatial heterogeneity patterns, we jointly model the (location) parameters of the marginal

distributions (ride-sourcing and DRS demand) as well as the copula dependence parameter (C_ϑ) as a function of exogenous variables in an additive fashion (explained next).

Systematic heterogeneity treatment To capture non-linear effects of exogenous variables (reflecting systematic heterogeneity), we jointly model the demand $f_\kappa(Y_\kappa|\mu_\kappa, \sigma_\kappa)$ for ride-sourcing (Y_1) and DRS (Y_2) as a function of exogenous variables with linear-in-parameter as well as smooth function effects in a generalized additive modeling (GAM) framework (Wood 2006):

$$f_1\left(Y_1|X_1\right) = \beta_o + \beta_1 X_{Q1} + \sum_{i=1}^{Q} f_{i1}\left(W_{Q1}\right) \tag{2}$$

$$f_2\left(Y_2|X_2\right) = \alpha_o + \alpha_1 X_{Q2} + \sum_{i=1}^{Q} f_{i2}\left(W_{Q2}\right) \tag{3}$$

where: Q is an index for the number of observations (census tracts); $[\beta_o, \alpha_o]$ are intercept terms, X_{Q1} and X_{Q2} are the vectors of exogenous variables with a linear influence on demand for ride-sourcing and DRS services; $[\beta_1, \alpha_1]$ are vectors of estimable parameters associated with X_{Q1} and X_{Q2}; W_{Q1} and W_{Q2} are vectors of exogenous variables with a non-linear influence on the demands for two emerging mobility options; $[f_{i1}, f_{i2}]$ are non-parametric smooth functions to capture the contours of relationships between exogenous factors in W_{Q1} and W_{Q2} with Y_1 and Y_2. Potential systematic heterogeneity in the stochastic dependence patterns is modeled similarly (Wali and Khattak 2022):

$$\vartheta = \gamma_o + \gamma_1 Z_Q + \sum_{i=1}^{Q} f_i\left(\beth_Q\right) \tag{4}$$

where: γ_o is an intercept term for the geo-additive nested model for the copula dependence parameter, Z_Q and $\beth_Q$ are matrices of exogenous variables *linearly* and *non-linearly* influencing the joint dependence between ride-sourcing and DRS demand. The combination of parametric and non-parametric terms in Eqs. 2 through 4 offers a parsimonious structure to capture systematic heterogeneity in the effects of exogenous variables. For estimating the non-parametric smoothers, low rank think plate regression splines are used—allowing the creation of basis expansions underlying each smoother in a data-driven manner (compared to subjective/user-defined identification of knot locations for nonlinear modeling or automatic identification of kernels using black-box machine learning methods) (Wood et al. 2015b). For more details about thin plate regression splines and smooth parameters estimation, see (Wood et al. 2015b) (Wali et al. 2021a). The trade-off between penalizing smooth wiggliness and badness of fit is achieved by implementing an unbiased risk estimator-based regularization scheme. For more details, see (Wood et al. 2015a).

Spatial heterogeneity treatment Related to the stochastic dependence between ride-sourcing and DRS services, the GAM-based exposition in Eq. (4) allows the joint dependence pattern to vary across observations in a linear and non-linear fashion but does not account for spatial heterogeneity. The demand for emerging mobility services is known to be spatially clustered/correlated. Given the spatial similarities/differences across space, it is reasonable to expect that the joint dependence between ride-sourcing and DRS services could indeed vary across space. We expand the GAM exposition in Eq. (4) to incorporate

a Markov Random Field (MRF) based spatial model to capture the spatial spillover effects in the stochastic dependence pattern among the two outcomes. Since the geographical area (Chicago) is split up into discrete (contiguous) census tracts, the MRF-based nested model exploits the place-based information in neighboring census tracts (Sørbye and Rue 2014). In particular, the GAM-based formulation for ϑ presented earlier is expanded as:

$$\vartheta = \gamma_o + \gamma_1 \mathbf{Z}_Q + \sum_{i=1}^{Q} f_i(\beth_Q) + \sum_{V_k=1}^{V_k} \delta_{kV_k} b_{kV_k}(\psi_{kQ}) \tag{5}$$

where: ψ_Q is the $(k \times Q)$ matrix of exogenous variables; $\delta_{kV_k} b_{kV_k}$ indicate smooth functions for the kth factor in ψ_Q represented as a linear combination of V_k basis functions $b_{kV_k}(\psi_{kQ})$ and linear regression parameters δ_{kV_k}. For a spatial Markov Random Field model based on the proximity structure of all census tracts in the sample, $\delta_{kV_k} b_{kV_k}(\psi_{kQ})$ transforms to $\delta_Q(\psi_Q)$; where ψ_Q now contains the area labels for all Q census-tracts in the sample and δ_Q is a vector of (census-tract specific) spatial random parameters $[\delta_1, \delta_2, \delta_2, \ldots, \delta_Q]'$—distributed as $N_Q(0, \sigma_e^2 \varpi^{-1})$ for a generalized inverse of $(Q \times Q)$ adjacency matrix (ϖ). Represented as $\varpi[q, m]$, for $q = 1, 2, 3, \ldots, Q$ and $m = 1, 2, 3, \ldots, Q$, the matrix ϖ contains the spatial structure of the neighborhoods (census-tracts), with elements indicated by $\varpi_{qq} = N_q$ (where N_q is the total number of adjacent census-tracts for census-tract q), $\varpi_{qm} = -1$ if census-tracts q and m are adjacent, and $\varpi_{qm} = 0$ if census-tracts q and m are non-adjacent—ensuring that the smoothing parameter reflects the spatial structure of the neighborhoods. To link an observation i with corresponding spatial random parameters in δ_Q, a proximity matrix is defined as $P[i, q] = [1$ if i belongs to census-tract q, 0 otherwise], and $q = 1, 2, 3, \ldots, Q$.

Collectively, the framework presented in this study expands previous work and the state of the art in copula modeling where only the joint stochastic dependence was modeled as a linear (Yasmin et al. 2014; Nashad et al. 2016) or generalized additive/non-linear function of exogenous factors (Wali and Khattak 2022). These studies did not model the potential spatial heterogeneity in the joint dependence between the unobserved factors underlying the response outcomes, e.g., ride-sourcing and DRS demand (as is done in the present study). In implementing the heterogeneous copula dependence structures, we use alternative parameterization forms. For comprehensive copulas (Gaussian, Frank, FGM, and AMH copulas), we use an identity link as shown in Eq. 5. For non-comprehensive (survival) copulas with dependence parameters spanning $\vartheta \in [-1, \infty)$ (such as Clayton copula), we use a log identity function $\log(\vartheta)$—re-parametrizing ϑ as $\exp(\gamma_o + \gamma_1 \mathbf{Z}_Q + \sum_{i=1}^{Q} f_i(\beth_Q) + \sum_{V_k=1}^{V_k} \delta_{kV_k} b_{kV_k}(\psi_{kQ}))$. For those with dependence parameters strictly positive (e.g., Gumbel-Hougard copula), we use a $\log(\vartheta - 1)$ parametrization.

Finally, referring to the copula-based exposition in Eq. 1, with $\aleph_1$ and $\aleph_2$ being continuous with probability densities f_1 and f_2, the joint density (f) for ride-sourcing and DRS demand can be written as:

$$f(Y_{1Q}, Y_{2Q}) = \frac{\partial^2 \aleph(Y_{1Q}, Y_{2Q})}{\partial Y_{1Q} Y_{2Q}} = \frac{\partial^2 C_\vartheta(\aleph_1(Y_{1Q}), \aleph_2(Y_{2Q}))}{\partial \aleph_1(Y_{1Q}) \partial \aleph_2(Y_{2Q})} \times \frac{\partial \aleph_1(Y_{1Q})}{\partial Y_{1Q}} \times \frac{\partial \aleph_2(Y_{2Q})}{\partial Y_{2Q}} \tag{6}$$

And, the joint log-likelihood function can be expressed as:

$$ll(\omega) = \sum_{i=1}^{Q} log\left\{ C_{\vartheta}\left(\aleph_1\left(Y_{1Q}|\mu_{1Q}, \sigma_{1Q}\right), \aleph_2\left(Y_{2Q}|\mu_{2Q}, \sigma_{2Q}\right)\right)\right\} + \sum_{i=1}^{Q}\sum_{\kappa=1}^{2} log\left\{ f_\kappa\left(Y_{\kappa Q}|\mu_{\kappa Q}, \sigma_{\kappa Q}\right)\right\}$$

(7)

where: ω is a matrix containing all the parameter estimates from the copula-based geo-additive Markov Random Field-based spatial model detailed above. In summary, the methodological framework accounts for three different layers of random, systematic, and spatial heterogeneity: (1) joint stochastic dependence between the demand for ride-sourcing and DRS services arising due to common unobserved factors (modeled through a copula-based framework), (2) systematic heterogeneity in the effects of exogenous variables on the demand for ride-sourcing and DRS services and on the joint stochastic dependence between the two (modeled through a nested GAM framework), and (3) spatial heterogeneity in the joint stochastic dependence between ride-sourcing and DRS services (modeled through a nested Markov Random Field model).

Results

Study area characteristics

To contextualize the forthcoming results and help readers infer potential similarity patterns, Table 1 presents summary statistics on demographic, socioeconomic, housing and transport, internet, and computer accessibility factors for the study area (Chicago). For comparison, summary statistics are also provided for the 11 largest (most populous) US cities and the entire nation. Regarding demographics, gender and age distributions of the population are almost similar across the three categories (Chicago, 11 largest US cities, and nationwide). The percentages of white alone populations in Chicago and the 11 largest US cities (50% and 57.8%) are significantly lower than the percentage nationwide (76.3%). Compared to the nation, the percentages of foreign-born populations in Chicago and the 11 largest US cities are significantly larger. Regarding socioeconomic status, the percentage of the population living in poverty and the per capita income in Chicago are larger than nationwide trends but similar to the trends in the 11 largest US cities. The mean owner-occupied housing rates of 45% and 46.1% in Chicago and the 11 largest cities are significantly lower than the nationwide rate of 64%. Chicago is more congested with an average 35.1-min commute. Around 88% and 79% of the households in Chicago have access to computers and broadband internet, respectively (compared to around 90% and 82% nationwide). In terms of education, the percentage of the population in Chicago having at least a high school degree is slightly lower than the nationwide average. As expected, Chicago is substantially denser in terms of population density compared to the entire nation. Despite Chicago being unique in certain aspects, the above statistics indicate that the demographic, socioeconomic, housing, and transportation fabric of Chicago seems to be representative of the largest US cities, and more broadly, seem representative of the large urban areas in the United States.

Spatial distribution of ride-sourcing and dynamic ridesharing demand

Figure 1 illustrates the spatial distribution of the response outcomes: ride-sourcing and dynamic ridesharing demand—revealing spatial heterogeneity in ride-sourcing and DRS

Table 1 Comparison of study area characteristics

Variables	Chicago city, Illinois	Average across 11 largest cities	US
Demographics			
Persons aged < 18 years (%)	20.90%	22.56%	22.30%
Persons aged > 65 years (%)	12.40%	11.83%	16.50%
Persons aged < 65 years with disability (%)	7.00%	7.45%	8.60%
Female (%)	51.40%	50.60%	50.80%
White alone (%)	50.00%	57.82%	76.30%
Black alone (%)	29.60%	16.65%	13.40%
Asian alone (%)	6.60%	10.59%	5.90%
Two or more races (%)	2.80%	3.54%	2.80%
Foreign born persons (%)	20.60%	25.53%	13.60%
Income and poverty			
Median household income (in 2019 USD)	$58,247	$64,180.73	$62,843
Per capital income (past 12 months) (in 2019 USD)	$37,103	$35,737.00	$34,103
Persons in poverty (%)	18.40%	17.10%	11.40%
Housing and transportation			
Owner-occupied housing unit rate	45.00%	46.15%	64.00%
Household size (persons per household)	2.48	2.70	2.62
Mean travel time to work (mins)	35.1	29.80	26.9
Internet and computer accessibility			
Households with a computer (%)	87.90%	90.73%	90.30%
Households with a broadband internet (%)	78.80%	82.52%	82.70%
Education			
High school graduates or higher (%)	85.10%	82.94%	88.00%
Bachelor's degree or higher (%)	39.50%	36.72%	32.10%
Population density			
Population / sq. mi (2010)	11,841.80	7550.45	87.4

The statistics are based on publicly available data provided by the US Census QuickFacts (US Census 2021). The 11 largest cities include New York City (NY), Los Angeles (CA), Chicago (IL), Houston (TX), Phoenix (AZ), Philadelphia (PA), San Antonio (TX), San Diego (CA), Dallas (TX), San Jose (CA), and Austin (TX)

rates across the census tracts in Chicago (Fig. 1a). The demand for ride-sourcing and DRS services is highest in the downtown Chicago area which is unsurprising since central business districts (CBD) exhibit larger attractions and greater activity (Fig. 1a). To gain high-level insights into the distributions across time of day, Fig. 1b–d show the ride-sourcing and DRS rates across midday, AM, and PM periods. Compared to the overall data, the magnitude of ride-sourcing and DRS demand lowered in the midday, AM, and PM periods. The mean (standard deviation) ride-sourcing and dynamic ridesharing rates per 100 individuals (over the six-month period) are: 890.9 (2084.3) and 199.4 (365.3) [overall data]; 155.2 (302.7) and 41.4 (54.6) [AM period]; 222.1 (551.1) and 59.1 (137.9) [PM period]; and 146.5 (369.9) and 31.1 (54.8) [Midday/Noon Period]. Compared to the overall data, the ride-sourcing and DRS rates are intuitively lower during the specific periods. Some spatial variations in the demand for ride-sourcing and DRS across different periods can also be seen. For example, the demand for ride-sourcing and DRS services were more
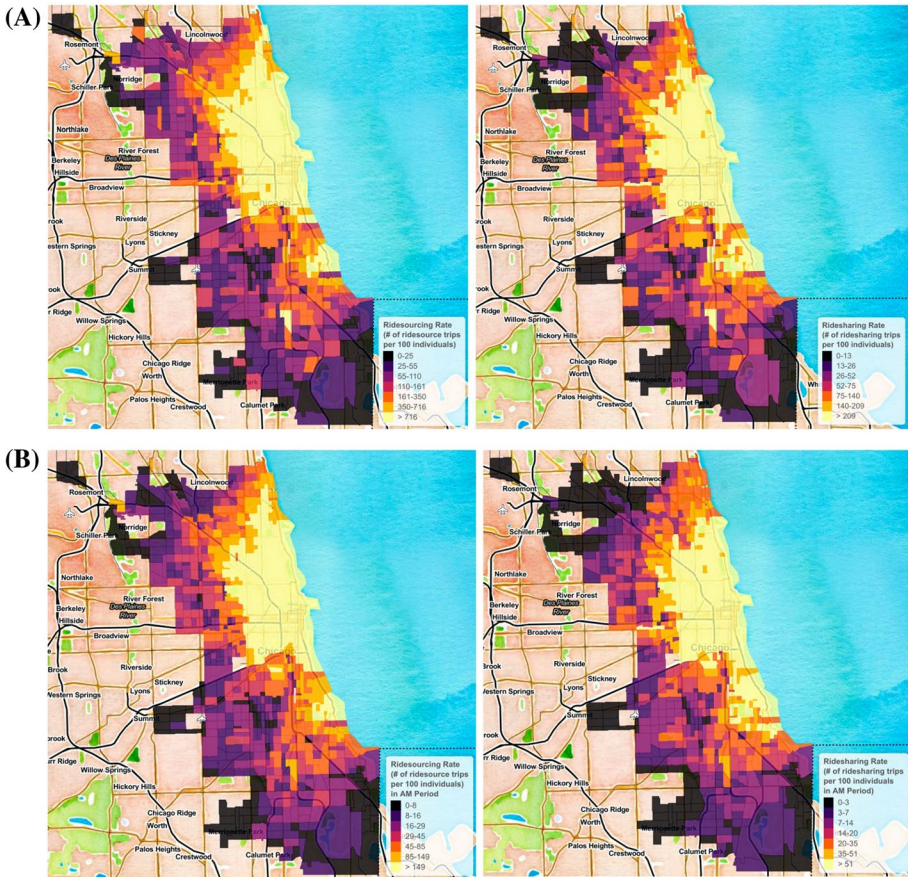
Fig. 1 Choropleth maps of ride sourcing and dynamic ridesharing demand by pick-up census tracts. **a–d** show the ride-sourcing and dynamic ridesharing demand for the overall data, AM peak [7 AM–10 AM], PM peak (2 PM–6 PM), and Midday/Noon (10 AM–2 PM), respectively. Please zoom in for better legibility

spread out throughout the city during the PM periods—with neighborhoods farther away from the downtown core producing more ride-sourcing and DRS trips (Fig. 1c).

To gain further insights into the spatial distribution of trip flows, we created an origin–destination matrix calculating the total number of all trips, ride-sourcing, and ridesharing trips between each pair of origin–destination census tracts (leading to a total of 200,956 flows/O–D pairs). For computational reasons and to focus only on key trip flows for visualization purposes, we used minimum and maximum trip flow thresholds of 50 and 200, respectively (leading to over 26,000 O–D pairs).

Figure 2 shows the trip flows for these O–D pairs revealing interesting patterns. First, we see a semi-monocentric pattern of the city with most flows ending around downtown (see the yellowish concentration in trip flow plasma for all trips in Fig. 2). Geographic clustering is noted in the demand for shared mobility, with the demand dominating in the downtown area and with some uptake on the outskirts of the city. Many small clusters of flows in the outer region are observed, e.g., near the two airports on the far west of the map in Fig. 2. Finally, to note here is the remarkable spread in accessibility (in terms
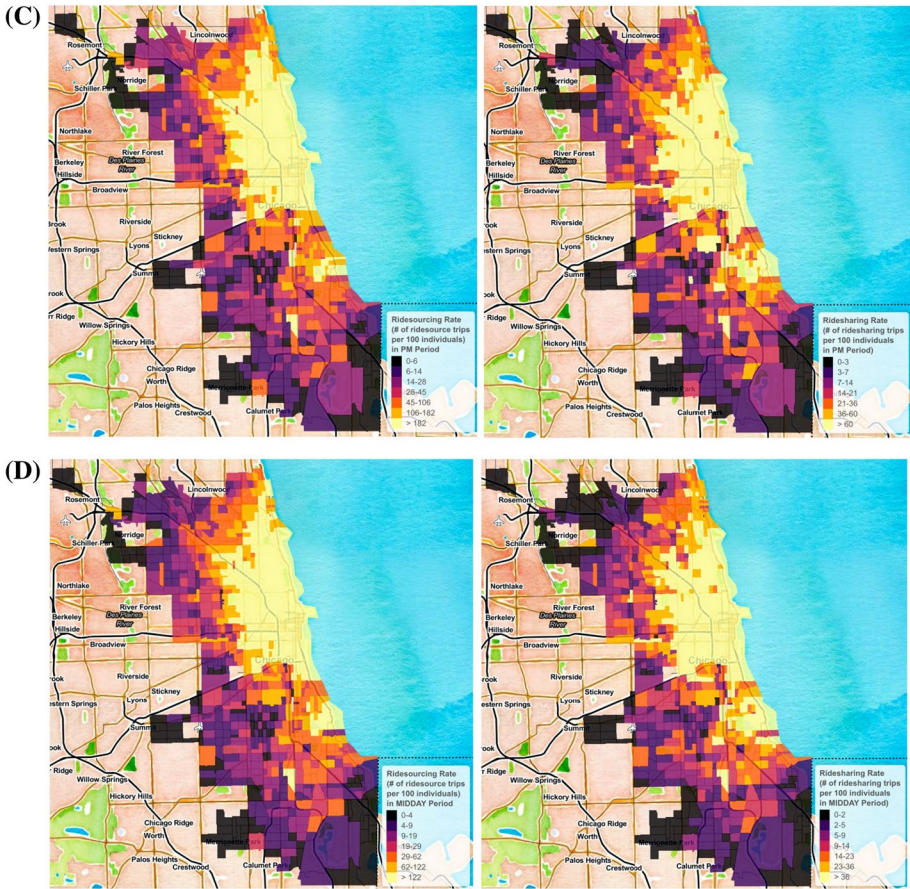
(C)

(D)
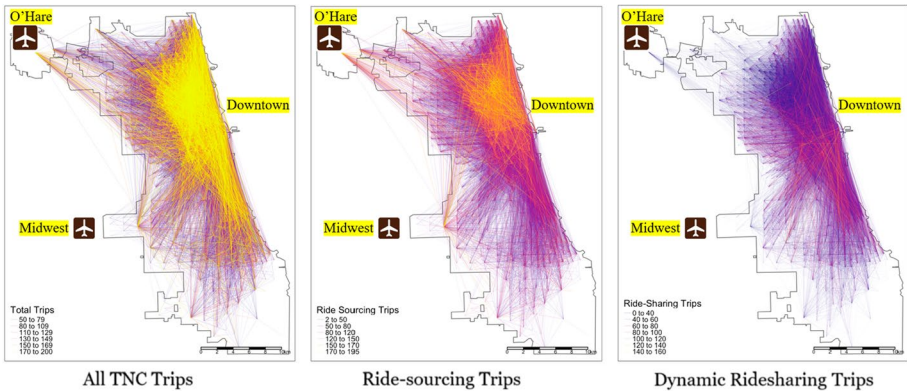

Fig. 1 (continued)

Fig. 2 Ride-sourcing and dynamic ridesharing trip flows (origin–destination Pairs) in Chicago. Since the data are census-tract specific, the origin–destination flows shown above indicate the flows between different census tracts. The starting and ending points indicate the centroids of the origin and destination census tracts. Longer flows indicate greater distances. Please zoom in for better legibility

of accessing farther locations) offered by shared mobility services. The second and third sub-plots in Fig. 2 illustrate the distribution of demand for ride-sourcing and ridesharing trips, respectively. Ride-sourcing still dominates in the city center but with sizeable clusters around the airports as well. However, for ridesharing, the cluster dominates in the city center but diminishes around the airports. This suggests that ridesharing trips are more frequent in dense areas where more non-utilitarian (leisure) travel is likely to take place.

Descriptive statistics

Table 2 shows the descriptive statistics of the two response outcomes and key exogeneous variables. Regarding the demand for the two shared mobility services, the mean ride-sourcing and dynamic ridesharing rates (trips per 100 inhabitants) were 890.92 and 199.46 over the six months, respectively (Table 2). As was shown in Fig. 1 earlier, substantial heterogeneity is observed in ride-sourcing and dynamic ridesharing rates across the census tracts (see the standard deviations in Table 2). Thus, it is of interest to examine what factors drive the observable demand for ride-sourcing and dynamic ridesharing services and how the demand for the two is jointly related.

Regarding neighborhood commute patterns, the average travel time to commute was around 35 min. Almost half of the individuals in the sampled census tracts drove alone to their work. Participation in active travel modes for commuting is also observed. Around 28%, 1.5%, and 5.2% of the individuals in the sampled census tracts took transit, walked, and biked to their work—with substantial variations across the neighborhoods. Interestingly, around 5% of the individuals participated in teleworking. By harnessing data on travel behavior, we examine the relationships between active travel and demand for emerging mobility services.

A broad spectrum of density, design, diversity, and transit measures are considered to capture the built environment and transit accessibility profiles across the sampled neighborhoods. The sampled census-tracts show remarkable variation in terms of density (residential, population, employment), diversity, and design features. The sampled census tracts exhibited greater street connectivity for pedestrian activities. Besides pedestrian-oriented intersections, we also include design measures for multi-modal and pedestrian infrastructure to better reflect the support for pedestrian and bicycle travel. The average facility miles of multi-modal and pedestrian links per square mile are 9.98 and 19.55, respectively. To examine how transit accessibility correlates with the demand for ride-sourcing and dynamic ridesharing services, several transit measures are also included in the analysis (Table 2). The average distance to the nearest transit stops from the population-weighted centroid was 265 m whereas around 21% of employment existed within 0.25 miles of a transit stop. Descriptive statistics for crime-related variables are shown in Table 2.

Regarding demographics, individuals in the sampled neighborhoods were almost evenly split between men and women and between white and non-white. The average neighborhood is young with around 37% of individuals aged between 18 and 39 years. Around 15% of Chicago's population has a graduate degree with some neighborhoods reaching 75%. As expected, the majority of residents have access to cellular Internet but there are some neighborhoods where over 40% of the residents do not possess cellular internet coverage.

Table 2 Descriptive statistics of response outcomes and key exogenous variables

Variable	Mean	SD	Min	Max	Data source
Outcome variables					
Ride-sourcing trip rate (# of ride-sourcing trips per 100 population)	890.92	2084.40	0.22	22,837.25	City of Chicago Data Portal
Ridesharing trip rate (# of ridesharing trips per 100 population)	199.46	365.30	0.02	4003.75	
Ride-sourcing trip rate (log-form)	5.31	1.79	−1.51	10.04	
Ridesharing trip rate (log-form)	4.30	1.50	0.02	8.30	
Commute (active) travel					
Work commute (mins)	35.67	5.53	17.82	53.55	American Community Survey 5-Year Estimates (2014–2018)
Drive alone (%)	50.05	15.60	10.89	85.93	
Transit (%)	28.58	12.98	3.65	63.88	
Biking (%)	1.54	2.23	0	14.70	
Walk (%)	5.27	8.23	0	62.17	
Telework (%)	4.83	3.61	0	18.09	
Built environment variables density					
Gross residential density (HU/acre)	16.41	20.07	0.82	298.63	2020 EPA smart location database
Gross population density (people/acre)	33.20	27.71	3.29	407.58	
Gross employment density (jobs/acre)	11.93	52.53	0	1102.55	
Diversity					
5-tier employment entropy [* by 100]	60.78	20.35	0	99.95	
Design					
Total road network density	30.57	6.40	9.76	69.43	
Facility miles of multi-modal links per square mile	9.98	5.46	0	26.35	
Facility miles of pedestrian-oriented links per square mile	19.55	6.77	1.24	44.79	
Street intersection density (weighted, auto-oriented intersections eliminated)	175.67	84.20	22.34	758.90	
Transit (accessibility)					
Distance from POPULATION weighted centroid to nearest transit stop (meters)	265.18	98.21	0	932.08	
Proportion of employment within ¼ mile of fixed-guideway transit stop	0.21	0.29	0	1.00	
Regional Centrality Index—Transit [* by 100]	47.68	17.11	3.33	96.14	

Table 2 (continued)

Variable	Mean	SD	Min	Max	Data source
Crime (# of offenses)					
Assault	261.77	231.30	1	1830.00	Crime open database (CODE)
Burglary	48.92	33.43	0	201.00	
Narcotic	46.73	111.48	0	1356.00	
Homicide	2.05	3.04	0	22.00	
Trip related variables					
Average trip duration (mins)	18.47	3.00	11.52	34.54	City of Chicago data portal
Average trip distance (miles)	5.48	1.49	2.79	14.10	
Average trip fare (US dollars)	9.91	1.64	7.33	25.72	
Weekend (%)	29.82	9.12	0	77.68	
Weekday AM (%)	22.71	9.85	0	59.37	
Weekday Noon (%)	10.74	3.78	0	32.69	
Weekday PM (%)	18.67	7.46	0	78.54	
Weekday Evening (%)	12.14	4.60	0	32.20	
Weekday Overnight (%)	5.91	2.98	0	41.54	
Weekend AM (%)	3.70	1.46	0	8.89	
Weekend Noon (%)	6.05	1.69	0	21.40	
Weekend PM (%)	7.25	2.45	0	25.77	
Weekend Evening (%)	7.16	3.20	0	25.85	
Weekend Overnight (%)	5.65	5.61	0	57.14	
Socio–demographic factors					

Table 2 (continued)

Variable	Mean	SD	Min	Max	Data source
Black (%)	35.90	39.57	0	100.00	American community survey 5-year estimates (2014–2018)
White (%)	46.00	32.27	0	97.49	
Male (%)	48.29	4.74	25.23	64.40	
Education: less than high school (%)	15.57	11.43	0	56.16	
Education: graduate degree (%)	14.89	13.26	0	75.73	
Low income (%)	40.62	19.21	3.86	92.05	
High income (%)	26.28	18.52	0	79.01	
Age: 18–39 years (%)	36.95	12.31	14.82	88.40	
Unemployed (%)	6.09	4.61	0	27.83	
Cellular internet access (%)	80.29	9.54	42.40	100.00	

$N = 793$ census tracts; SD is standard deviation; All variables are at the census-tract level; AM [7 AM–10 AM], Midday/Noon (10 AM–2 PM], PM (2 PM–6 PM], Evening (6 PM–10 PM], Night/Overnight (10 PM–7 AM)

Modeling results

Several models are defined next that jointly estimate ride-sourcing and DRS rates as a function of built environment, transit accessibility and crime, active travel, and sociodemographic factors while treating different layers of random, systematic, and spatial heterogeneity. To derive the models, a systematic process was followed considering variable importance (based on literature) and specification parsimony. The statistical significance criterion was set at a 90% confidence level. We used information criterion (Akaike Information Criterion) based model selection mechanisms to compare the different models with different sets of exogenous variables and treating different layers of random, systematic, and spatial heterogeneity. A difference of over 5 points between the AICs of two alternative models would favor the model with the lowest AIC. Likelihood-ratio tests were also performed to compare the alternative modeling structures.[5] To deal with distributional skewness, a natural log transformation was used for the two response outcomes.

Initially, two univariate (independent) models were estimated for the two response outcomes: modeling ride-sourcing and DRS rates as a function of exogenous variables shown in Table 2. The demand for ride-sourcing and DRS is strongly correlated with a Pearson and Kendall τ correlation coefficients of 0.92 and 0.83, respectively. To account for this correlation arising due to observed and unobserved factors, an elliptical Gaussian copula-based joint model was estimated. Accounting for the correlation between the ride-sourcing and DRS demand led to substantial improvement in log-likelihood and AIC (with one extra estimable (copula dependence) parameter—the log-likelihood for the Gaussian copula-based joint model increased by over 696 points and the AIC reduced by over 1390 points). This highlights the importance of accounting for joint stochastic dependence in estimating the demand for ride-sourcing and DRS services.

Non-linear stochastic dependence between demand for ride-sourcing and DRS services

While the Gaussian copula-based joint model led to substantial improvements, it is nonetheless characterized by a linear form of joint dependence between the unobserved factors influencing the demand for ride-sourcing and DRS services. Following the discussion in Sect. 3.2.1, a broad spectrum of comprehensive and non-comprehensive copula-based joint models was next developed to capture varying levels of non-linear stochastic dependence patterns between the demand for ride-sourcing and DRS services. To examine asymmetry in joint stochastic dependencies, asymmetric copula-based models were also estimated.

---

[5] For comparing alternative model structures, a chi-square distributed likelihood ratio (LR) test statistic is formulated as: $-2 \times \left[\mathrm{LL_R} - \mathrm{LL_U} \right]$—where $\mathrm{LL_U}$ refers to the log-likelihood of the un-restricted model and $\mathrm{LL_R}$ refers to the log-likelihood of the restricted model. In this study, the restricted models are those that restrict the stochastic dependence, assume systematic and/or spatial homogeneity (as discussed in Sect. 3.2), whereas the unrestricted models allow non-linear stochastic dependence (through copulas), systematic heterogeneity (through geo-additive framework), and/or spatial heterogeneity (through Markov Random Fields). If the more advanced (unrestricted) model is statistically superior, then the LR test statistic comparing the unrestricted and restricted models should be greater than the critical χ^2 value associated with the additional degrees of freedom in the unrestricted model (i.e., additional estimable parameters to capture stochastic dependence, systematic, and/or spatial heterogeneity). In selecting the critical χ^2 values, we use a 99.9% level of confidence.

Finally, survival variants of asymmetric Archimedean copulas were implemented to model inverted dependency patterns.

Table 3 shows the goodness-of-fit statistics for the twelve (12) elliptical and Archimedean copula-based structures for ride-sourcing and DRS demand. Among all competing copula-based models, the Frank copula-based joint model for ride-sourcing and DRS demand resulted in best fit (labeled as Model 1 in Table 3). With the same number of estimable parameters (degrees of freedom) between the Gaussian and Frank copula models, the log-likelihood for Frank copula-based joint model increased by 59 points, and the AIC reduced by over 117 points (Table 3). This finding emphasizes the presence of non-linearity in the joint stochastic dependence of unobserved factors driving the demand for ride-sourcing and DRS services. Compared to Gaussian copula, Frank copula implies a stronger dependence in the center of the bivariate distribution and weaker dependence in the tails. The copula dependence parameter for the Frank copula-based structure was 19.46 with an associated Kendall's τ of 0.81 (on a scale from -1 to 1). This shows that even after conditioning on observed exogenous variables, the demand for ride-sourcing and DRS services are strongly correlated based on latent factors and that the unobserved factors jointly increase/decrease the neighborhood-level demand for ride-sourcing and DRS services.

Systematic and spatial heterogeneity in the demand for ride-sourcing and DRS services

We developed a series of models to capture systematic and spatial heterogeneity in the effects of exogenous factors on the demand for ride-sourcing and DRS services. The best-fit Frank copula-based joint model discussed in the earlier section was expanded with a GAM-based framework—incorporating systematic heterogeneity arising from non-linearity in the effects of exogenous factors on ride-sourcing and DRS demand (Model 1A in Table 3). Accounting for non-linearity in the joint model led to a 76-points reduction in AIC for the Frank copula-based GAM model, indicating substantial improvement in model fit over the Frank copula model assuming linear covariate effects. Following the methodological discussion in Sect. 3.2., Model 1A was further expanded to account for heterogeneity in the joint dependence pattern across Chicago neighborhoods (Model 1B in Table 3). In this model, the ϑ copula dependence parameter was allowed to vary across census tracts as a function of exogenous variables. Doing so led to a further 13 unit decrease in AIC for Model 1B (compared to Model 1A), suggesting substantial improvements in model goodness of fit while elucidating the variations in the joint dependence between the two mobility services as a function of independent variables.

Finally, we expanded Model 1B to incorporate structured spatial heterogeneity (as discussed in Sect. 3.2.3.2) in the joint stochastic dependence between ride-sourcing and DRS demand (Model 1C in Table 3). In a holistic framework, Model 1C simultaneously accounted for the joint stochastic dependence between ride-sourcing and DRS demand, systematic heterogeneity in the effects of covariates on the two response outcomes, systematic and spatial heterogeneity in the joint dependence patterns across Chicago neighborhoods. Accounting for all the different sources of heterogeneity led to a 115-unit decrease in AIC for Model 1C compared to Model 1B. Likewise, the LR test statistic was 206.12 $[-2 \times ((-759.63 + -656.57)]$, which is greater than the critical χ^2 value (for a 99.9% confidence level) of 80.07 for 45 degrees of freedom (i.e., the additional estimable parameters in the spatial Markov Random Field GAM-based Frank copula model (Model 1C) compared to Model 1B (96.71–51.46)) (Table 3). This substantial improvement in the goodness of fit

Table 3 Model selection and goodness of fit statistics for joint geo-additive markov random field based heterogeneous copula models

Models	LL (β)	AIC	ϑ	τ	K	N
Univariate models						
Ride-sourcing rate	−793.83	1627.66	–	–	20	793
Dynamic ridesharing rate	−778.32	1592.64	–	–	18	793
Total	*−1572.15*	*3220.29*	*–*	*–*	*38*	*793*
Joint models						
Symmetric Copula						
Gaussian copula	−876.05	1830.10	0.93	0.75	39	793
Farlie-Gumbel-Morgenstern (FGM) copula	−1404.47	2886.95	1.00	0.22	39	793
Best-fit homogenous copula: frank copula (Model 1)	−817.07	1712.14	19.46	0.81	39	793
Student-*t* copula	−827.45	1734.89	0.95	0.81	39	793
Asymmetric copula						
Ali-Mikhail-Haq (AMH) copula	−1208.44	2494.89	1.00	0.33	39	793
Gumbel-Hougard (G-H) copula	−851.20	1780.40	4.60	0.78	39	793
Hougard copula	−851.20	1780.40	0.22	0.78	39	793
Joe copula	−956.92	1991.84	6.69	0.75	39	793
Clayton copula	−993.39	2064.77	5.27	0.73	39	793
Rotated asymmetric copula (180-degrees rotation)						
Survival G–H copula	−869.89	1817.78	4.49	0.78	39	793
Survival Joe copula	−991.46	2060.92	6.05	0.72	39	793
Survival Clayton copula	−957.91	1993.82	5.94	0.75	39	793
Heterogeneous geo-additive MRF copula models						
Generalized additive (non-linear) frank copula model (model 1A)	−767.22	1635.81	20.66	0.81	50.68	793
Generalized additive (non-linear) frank copula model with heterogeneous stochastic dependence (model 1B)	−759.63	1622.20	20.98*	0.82*	51.46	793
Generalized additive (non-linear) frank copula model with heterogeneous stochastic dependence and MRF based spatial heterogeneity (model 1C)	−656.57	1506.57	23.43*	0.83*	96.71	793

LL(β) is log-likelihood at convergence; AIC is Akaike Information Criterion; ϑ is the copula dependence parameter; τ is Kendall's Tau measure of joint dependence; K is the number of estimable parameters (degrees of freedom); N is the sample size; (*) indicates heterogeneous copula joint dependence measures (shown in Figs. 3 and 4); (–) indicates Not Applicable

for Model 1C supports our hypothesis about the existence of structured spatial heterogeneity in the joint stochastic dependence between ride-sourcing and DRS demand.

Besides substantial improvements in model goodness of fit, the more advanced Model 1C also provides key behavioral insights. We briefly discuss the behavioral insights related to joint stochastic dependence and spatial heterogeneity next. The empirical findings related to systematic heterogeneity are synthesized in the next section. Table 4 shows the estimation results for the best-fit geo-additive Markov Random Field-based joint heterogeneous copula model (Model 1C). In particular, the joint stochastic dependence implied by Frank copula varied as a function of the proportion of neighborhood trips undertaken during weekend PM periods (Table 4). With a β estimate of 0.529 (*t*-statistic of 1.82), the joint dependence between latent factors underlying the demand for ride-sourcing and DRS services increased as more trips were undertaken during weekend PM periods (Table 4).

This finding is illustrated in Fig. 3, which shows the contour plots for the variations in joint dependence patterns between ride-sourcing and DRS demand. The strong synergy between ride-sourcing and DRS services during weekend PM periods is expected since the demand for on-demand mobility services is generally higher during weekends when leisure/recreational trip-making is more likely (Rayle et al. 2016). Previous research has also revealed that social-recreational activities during PM peak are more flexible compared to recreational trip-making during off-peak and late-evening activities (Bhat 1998)—which could facilitate the use of DRS services during weekend PM periods.

In addition, the best-fit Model 1C highlights the spatially heterogeneous profile of joint dependence patterns across neighborhoods in Chicago. Figure 4 shows that the structured spatial effect is smoother for the copula dependence parameter revealing spillover effects. Precisely, the Frank copula parameter varied between 0.36 and 37.84 corresponding to Kendall's τ of 0.10 and 0.89, respectively. This demonstrates that the stochastic dependence between the two mobility options varies in space—with the strongest dependence in the downtown Chicago area and weakest in the outskirts (especially around the two airports). Such deeper insights cannot be obtained from simpler models ignoring spatially structured heterogeneity effects. The new finding that the joint stochastic dependence between ride-sourcing and DRS services is strongest in downtown has important behavioral implications. Behaviorally, this highlights the strong link between density/urban compactness and mobility-on-demand services. The Chicago Loop (being the main section of Downtown Chicago) is one of the densest places nationwide with a population density of 11,841/sq. mi. Dense and walkable places can influence individual decisions related to participation in shared mobility in several ways. For example, transportation consumption in urban environments with high-density and more mixed land use clusters is smaller due to shorter distances between housing, work, and leisure. With less reliance on motorized transportation, private (automobile) ownership could become unappealing whereas, using alternative forms of transportation (such as shared mobility) only adds a marginal cost thus becoming more attractive in denser areas. Likewise, urban environments with high-density clusters likely have limited public rights-of-ways (parking, curb space) thus encouraging the use of shared mobility services.

Discussion and synthesis

Based on the results of best-fit Model 1C (Table 4), the findings related to the joint impacts of the built environment, transit accessibility and crime, active travel, and socioeconomic factors are discussed next. The β parameter estimates from the joint model in Table 4 can be interpreted as elasticities since the two response outcomes are log-transformed. If a one-unit change is made to an exogenous variable, we would expect a 100*β, or precisely a $100 * \left(e^\beta - 1\right)$, percent change in the response outcomes. When the explanatory factors are also log-transformed, the β quantifies the percent change in Y with a one percent change in X (untransformed exogenous variables). Non-linear effects are modeled through thin-plate regression splines as discussed earlier. The smoothed spline effects cannot be represented with a single β estimate and are summarized in terms of degrees of freedom (Table 4)—where the degrees of freedom (roughly) suggest the number of piece-wise slopes tied together to approximate a non-linear relationship (Wood et al. 2015a; Wali et al. 2021a). A statistically significantly greater than one degree of freedom suggests the presence of non-linear relationships. In discussing systematic heterogeneity effects, we present

Table 4 Estimation results for best-fit geo-additive markov random field based joint heterogeneous copula model (model 1C)

Variables	Joint model of ride-sourcing and dynamic ridesharing demand					
	Ride-sourcing rate (log-form)		Dynamic ridesharing rate (log-form)		Copula dependence	
	Effective DF	Chi-square (p-value)	Effective DF	Chi-square (p-value)	Effective DF	Chi-square (p-value)
Non-parameter smoothers						
Spatial smoothers						
Spatially heterogeneous stochastic dependence (Markov Random Field smoother)	–	–	–	–	42.222	207.16 (0.000)
Thin-plate spline smoothers						
No vehicle households (%)	–	–	5.066	8.27 (0.088)	–	–
Built environment: Regional Centrality Index – Transit	5.843	79.23 (0.000)	4.892	55.57 (0.000)	–	–
Built environment: Facility miles of pedestrian-oriented links per square mile	1.121	3.517 (0.029)	2.563	6.959 (0.006)	–	–
Parametric terms	β	t-stat	β	t-stat	β	t-stat
Intercept	6.158	11.58	4.378	8.49	19.598	8.73
Variance of unobserved factors (log-form)	–0.292	–12.09	–0.280	–11.36	–	–
Built environment						
Gross residential density (log-form)	0.073	4.70	–	–	–	–
Distance from population weighted centroid to nearest transit stop (meters) (log-form)	0.158	2.28	0.141	2.21	–	–
5-tier employment entropy	0.003	2.23	0.002	1.89	–	–
Active travel						
Work Commute (mins)	–0.012	–2.08	–0.018	–2.99	–	
Commute mode: walking (%)	0.029	7.61	0.025	6.40	–	
Commute mode: transit (%)	0.017	6.39	0.017	5.84	–	
Teleworking (%)	0.008	3.77	–	–	–	
Crime						
Assault offenses	0.0014	9.09	0.0012	8.37	–	

Table 4 (continued)

Variables	Joint model of ride-sourcing and dynamic ridesharing demand					
	Ride-sourcing rate (log-form)		Dynamic ridesharing rate (log-form)		Copula dependence	
	Effective DF	Chi-square (p-value)	Effective DF	Chi-square (p-value)	Effective DF	Chi-square (p-value)
Drugs and narcotics offenses	−0.0015	−5.60	−0.0013	−5.17	–	–
Sociodemographic and trip characteristics						
% Black	−0.005	−11.58	–	–	–	–
Education: less than high school (%)	−0.013	−10.23	–	–	–	–
Education: Bachelor's degree (%)	–	–	−0.003	−2.46	–	–
Age: 18—39 years (%)	0.024	7.94	0.029	9.55	–	–
Weekend PM (%)	–	–	–	–	0.529	1.82
Weekday Noon (%)	0.020	2.52	0.044	5.86	–	–
Weekday PM (%)	0.030	6.09	–	–	–	–
Weekend Noon (%)	–	–	0.023	4.14	–	–
Average trip duration (minutes)	−0.116	−6.65	−0.064	−3.66	–	–
Average trip miles	−0.251	−5.96	−0.306	−7.20	–	–

(–) indicates not applicable or statistically insignificant terms
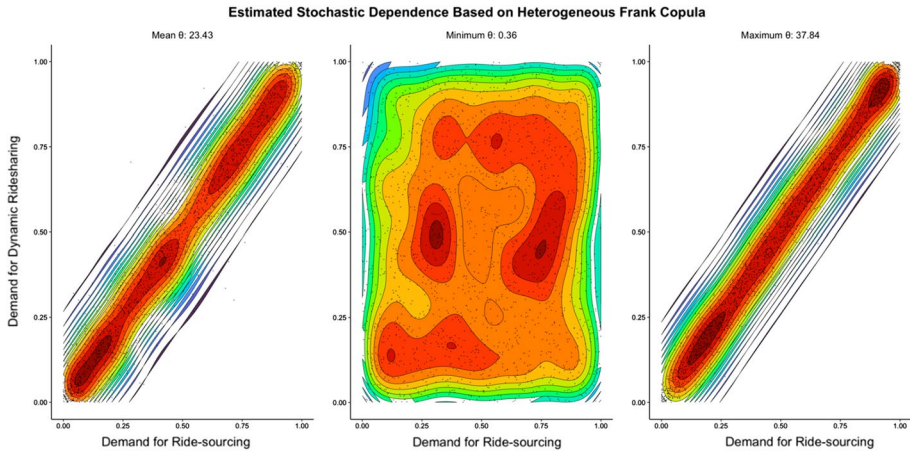
Fig. 3 Estimated heterogeneous joint dependence patterns between unobserved factors underlying the demand for ride-sourcing and dynamic ridesharing services. The contours are based on the Best-Fit Geo-Additive Markov Random Field-Based Joint Heterogeneous Copula Model–Model 1C. The intensity of joint dependence is shown by a rainbow plasma; Red color shows the portion where the joint dependence is strongest; Blue color shows the portions with weakest joint dependence; Isolines (lines indicating constant values) are tighter in regions where the joint dependence is stronger and vice versa (for blue regions). (Color figure online)

visualizations summarizing how the demand for ride-sourcing and DRS services non-linearly vary as a function of specific variables.

Built environment

The environmental features in the final model capture key aspects of the built environment related to urban design (pedestrian-oriented infrastructure), density, land-use diversity (mix of employment types), and transit accessibility. Related to urban design, we found a positive association between pedestrian-oriented infrastructure (facility miles of pedestrian-oriented links per square mile) and the demand for ride-sourcing and DRS services. In addition, greater provision of pedestrian-oriented infrastructure non-linearly influenced the demand for the two services (degrees of freedom of 1.12 and 2.56 in the equation for ride-sourcing and DRS demand) (Table 4). However, the extent of non-linearity was greater for DRS demand (see Figs. 5 and 6). In particular, when the pedestrian-oriented infrastructure density increased beyond 20 mi per sq. mi., a sharp increase in the demand for DRS services was observed (Fig. 6). These results suggest that while pedestrian-oriented connectivity can support both mobility-on-demand services, the demand for DRS services seems more responsive to changes in pedestrian-oriented connectivity. The relatively higher impact of greater pedestrian-oriented links on DRS services is intuitive. Greater pedestrian-oriented infrastructure likely generates more non-motorized traffic, helping both the demand and supply side of DRS services, i.e., supporting the use of DRS services and increasing the likelihood of matching passengers traveling to nearby destinations. Likewise, more compact, denser neighborhoods had a greater demand for ride-sourcing services. A one percent increase in residential density was associated with a 0.073% increase in ride-sourcing rate (Table 4). Regarding diversity, more mixed land use was associated
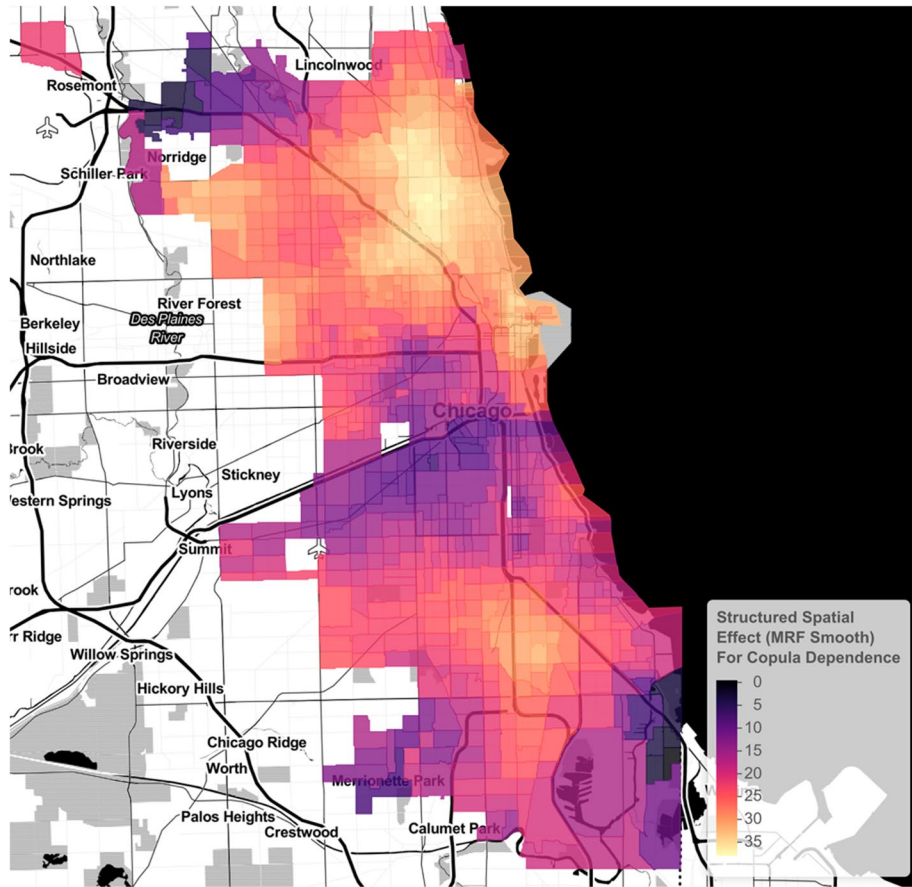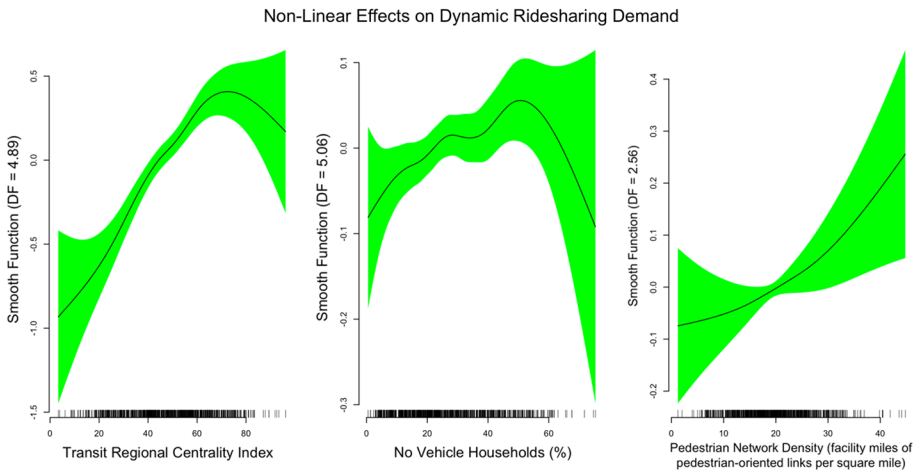
Fig. 4 Markov random field-based structured spatial heterogeneity in the joint dependence between ridesourcing and dynamic ridesharing demand. Based on Best-Fit Geo-Additive Markov Random Field-Based Joint Heterogeneous Copula Model–Model 1C

with greater demand for ride-sourcing as well as DRS services (Table 3). A one percent increase in employment entropy was correlated with a 0.003% and 0.002% increase in ride-sourcing and DRS demand, respectively. Land-use mix captures the composition of different activity types within a census tract. Greater land-use mix indicates neighborhoods (e.g., in the downtown area) with a variety of activity types (offices, restaurants, banks, shops, and other activities), thus increasing accessibility (Cervero 1989). Like density and compactness, note that land-use mix is a necessary but not sufficient condition for walkable neighborhoods. While more mixed land-use developments offer greater activity choices, these activity types must be sufficiently connected through the provision of pedestrian-oriented infrastructure to enable safer accessibility of vulnerable road users to different activity types. Collectively, these new findings highlight the effectiveness of denser, compact, connected, and more diverse neighborhoods with greater provision of pedestrian-oriented infrastructure in increasing the demand for ride-sourcing as well as DRS services. Such neighborhoods offer several benefits that are conducive to shared mobility services. For example, denser and compact neighborhoods with greater provision of pedestrian-oriented

Non-Linear Effects on Ride-sourcing Demand
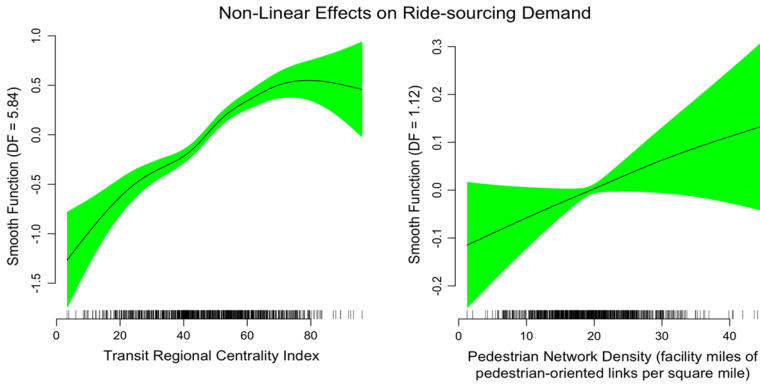
Fig. 5 Non-linear impacts on demand for ride-sourcing services. The contours are based on Best-Fit Geo-Additive Markov Random Field-Based Joint Heterogeneous Copula Model–Model 1C

Non-Linear Effects on Dynamic Ridesharing Demand

Fig. 6 Non-linear impacts on demand for dynamic ridesharing services. The contours are based on Best-Fit Geo-Additive Markov Random Field-Based Joint Heterogeneous Copula Model–Model 1C

infrastructure are more supportive of active travel and exhibit greater productions/attractions while providing safer and greater access to ride-sourcing and DRS services. Likewise, mobility-on-demand services are more attractive in walkable neighborhoods since such an urban environment provides users with accessible mobility without the burden of private vehicle ownership. Dean and Kockelman (2021) found negative associations between greater pedestrian infrastructure and the number of shared trips. Likewise, Xu et al. (2021) found negative associations (in certain ranges) of walk score, population, and employment density with ride-splitting adoption rate, whereas at certain value ranges, the contours remained flat. Our ability to detect positive and non-linear relationships discussed above can be due to the use of more recent built environment data as well as a methodologically rigorous empirical framework treating different layers of random, systematic, and spatial heterogeneity. In summary, the findings suggest that sustainable and shared mobility-related goals can be achieved by continuing to invest in compact and more walkable

neighborhoods. While more walkable and compact neighborhoods support shared mobility-related goals, an implicit assumption is that shared mobility services would not substitute trips undertaken by healthier and active travel modes (walk, bike, transit). To this end, it is important to examine the complementary effects (if any) of active travel and transit accessibility on shared mobility services (discussed next).

Transit accessibility

With regards to transit accessibility, neighborhoods with greater transit accessibility (compared to the metro region) had a higher demand for ride-sourcing and dynamic ridesharing services. Revealing significant systematic heterogeneity, the effects of transit accessibility were found to be non-linear both for ride-sourcing and DRS services (Fig. 5 and 6). A sharp increase in the demand for ride-sourcing and DRS services, in a piecewise non-linear fashion, is observed up until a transit regional index score of 60 (on a scale from 0 to 100). Importantly, ceiling and/or overdose effects are also observed. For ride-sourcing demand, a ceiling effect is observed when transit accessibility increases beyond 60 (Fig. 5). With regards to DRS demand (Fig. 6), an overdose effect is observed when transit accessibility increases beyond 60 leading to reductions in the demand for DRS services. Overall, the positive effect of increasing transit accessibility in areas with low levels of transit accessibility may be capturing the complementary role of ride-sourcing and DRS services in providing first- and last-mile connectivity. On the other hand, the ceiling and overdose effects suggest that the complementary role of shared mobility services may diminish in areas with great transit accessibility. As can be seen, the treatment of systematic heterogeneity enabled richer insights into the complex mechanisms through which transit accessibility interacts with the demand for shared mobility services.

Additionally, Figs. 5 and 6 reveal that relative to the demand for ride-sourcing services, the demand for DRS services appears more responsive to an increase in transit accessibility (reflected by the steeper curvature of transit regional centrality index in Fig. 6). This new finding concerning the positive but differential impacts of transit accessibility on ride-sourcing and DRS services is not unexpected. Compared to ride-sourcing, DRS services offer lower privacy, and users of DRS services are more likely to undertake transit trips as well (Zhang and Zhang 2018). Thus, increasing transit accessibility is likely to generate more demand for DRS services. On the other hand, ride-sourcing offers greater privacy and may serve as an attractive alternative for users who otherwise would have taken the trip using a taxi or personal vehicle in the absence of ride-sourcing services. Past research has shown that trips undertaken with ride-sourcing services replaced taxi trips the most, with car owners most likely to have driven themselves if ride-sourcing services were absent (Rayle et al. 2016). On the contrary, non-car owners were most likely to substitute ride-sourcing for transit services.

Collectively, compared to neighborhoods with high levels of transit accessibility, the above findings suggest that increasing transit accessibility in areas with low levels of accessibility could be more helpful in increasing the adoption of ride-sourcing and DRS services. Greater distances to transit stops are also positively correlated with the use of ride-sourcing and DRS services (Table 4). A one percent increase in the distance from the population centroid to the nearest transit stop is associated with a 0.158% and 0.141% increase in the demand for ride-sourcing and DRS services (Table 4). These findings suggest a potential substitution effect of ride-sourcing and DRS services when transit services are inconvenient or inaccessible. In line with the present study, Ghaffar et al. (2020) found

a positive relationship between greater transit supply (number of bus stops) and ride-sourcing demand.

Commuting patterns, active travel, and neighborhood crime

In terms of commuting and active travel patterns within a neighborhood, the results suggest that neighborhoods where greater percentage of individuals used transit and walking for work commute had higher rates of ride-sourcing and dynamic ridesharing. A one percent increase in the percentage of individuals who walked for work commute was associated with a 2.9% and 2.5% increase in the demand for ride-sourcing and DRS services (Table 4). In agreement with the literature (Zhang and Zhang 2018), similar positive associations were observed for transit use and ridesharing demand. The strong effect sizes reflect the synergies between active travel and shared mobility systems and, when considered together with the transit accessibility-related findings discussed earlier, highlight the potential for integrating public transit and shared mobility services for providing more accessible and multimodal transportation. At the same time, these findings could be concerning since ride-sourcing and DRS services could be replacing active travel trips. While clear complementary associations are observed, the data used in this study did not allow an examination of potential substitution effects of ride-sourcing/DRS services on active travel. Thus, future studies can benefit from identifying potential substitution effects. Neighborhoods with greater average commute distances had on-average lower ride-sourcing and DRS rates (Table 4). A one-minute increase in average commute distance was associated with a 1.24% and 1.81% reduction in ride-sourcing and DRS demand. This finding is in agreement with Dean and Kockelman (2021), who reported a negative association between work commute distance and ridesharing. Compared to ride-sourcing and DRS services, carsharing could be a more attractive alternative for commute travel since most of the commute occurs during peak/congested periods when ride-sourcing and DRS services are more expensive (Wali 2022). The relatively greater negative effect of commute distances on DRS services is intuitive since people are unlikely to use DRS services for work-related travel given the added uncertainty and travel times. Neighborhoods with a greater percentage of telecommuting workers exhibited greater ride-sourcing rates. A one percent increase in telecommuting was correlated with a 0.8% increase in ride-sourcing rate. Literature shows potential reductions in work-related trip making and VMT for telecommuters (Walls and Safirova 2004), and they may not feel the need to privately own vehicles. With an induced travel effect related to non-work or leisure travel (Kim et al. 2015), telecommuters may rely more on shared mobility services.

Related to the effects of crime, neighborhoods with greater assault offenses had a higher demand for ride-sourcing and DRS services. A possible reason for this finding is that individuals (especially those with no private vehicles) may consider shared mobility services to be safer compared to active travel or transit modes in neighborhoods with high crime rates (Ghaffar et al. 2020). On the contrary, the results show a lower demand for ride-sourcing and DRS services in neighborhoods with greater drugs and narcotics-related offenses (Table 3). This finding could be reflecting the negative effect of perceived safety risk (in terms of higher frequencies of violent/unwanted interactions) in ridesharing vehicles in neighborhoods with greater narcotics and drug use (Roughton 2020). Further research is needed on the associations between different types of assault offenses and the demand for shared mobility services. For example, there is a common feeling that on-demand mobility services are unsafe, especially for women. With several accidents reported on media, individuals have grown more

wary about safety of on-demand mobility services that are perceived as less safe than other on-demand mobility options that are better regulated (e.g., taxi).

Vehicle ownership and sociodemographic factors

Previous studies have shown a positive relationship between zero-vehicle households and demand for ride-sourcing (Ghaffar et al. 2020) or DRS services (Dean and Kockelman 2021). With an effective degree of freedom of 5.06 (chi-square value of 8.27) (Table 4), we found a non-linear interaction between zero-vehicle households and DRS demand (Fig. 6). An increase in zero-vehicle households in a neighborhood is correlated with a greater demand for ridesharing services. However, ceiling and overdose effects are observed for an increase in no-vehicle households beyond 25% (Fig. 6). The ceiling/overdose impacts of no-vehicle households may be tracing income effects—neighborhoods with a larger proportion of households having no vehicle may be comprised of low-income individuals who are relatively less willing to use ridesharing services (Bansal et al. 2020). Neighborhoods with a greater proportion of younger individuals have a higher demand for ride-sourcing/DRS services, whereas the reverse is true for neighborhoods with a greater proportion of less-educated individuals. Neighborhoods with a higher percentage of individuals with a bachelor's degree have lower rates of DRS. We note that this ACS-derived variable does not capture university students and only considers residents aged 25 or more. More educated individuals and/or university populations are reported elsewhere to have a greater willingness to use ridesharing services (Dias et al. 2017). In terms of race, neighborhoods with a greater proportion of Black individuals had lower demand for ride-sourcing services. Previous studies have also reported a negative association between race-ethnicity composition (including Black individuals) and demand for ride hailing services (Marquet 2020). Regarding temporal factors, weekday PM and weekend noon periods were associated with a greater demand for ride-sourcing and DRS services, respectively (Table 4). These findings are in agreement with previous studies suggesting a positive correlation between weekdays and ride hailing demand (Marquet 2020), and between weekends and ridesharing demand (Dean and Kockelman 2021). The positive association between weekends and DRS services demand could be due to a higher propensity of social-recreational trip-making on weekends, but also due to the decline in transit coverage on weekends. Likewise, due to a greater demand and lower supply, surge pricing is usually activated on weekends making DRS services more attractive when the fares are higher (Brown 2020).

Limitations and future work

This study has several limitations. Consistent with other studies, a limited sample of ride-sourcing and ridesharing trips was used containing over 33 million trips and spanning over six months (November 2018–April 2019). The authors could not analyze the entire data for computational reasons. The six-month study period covers Chicago's cold weather, and our results may not be generalized to other times of the year—especially when the demand for on-demand mobility services could be lower during summers (Gerte et al. 2018). We tested and controlled for potential time-of-day effects on ride-sourcing and DRS demand, and the stochastic spatial dependence between the two. As discussed, the strong spatial dependence observed between the two mobility services was robust to potential time-of-day impacts on the joint stochastic dependence of ride-sourcing and DRS services. Future studies can expand these results by separately analyzing the spatial dependency between ride-sourcing

and DRS services across different periods. Another limitation is that aggregate census-tract level demand for ride-sourcing and ridesharing services was analyzed. Common to most of the existing studies, this was due to the privacy filters put in place by the City of Chicago to protect PII data. Aggregate data mask individual-level heterogeneity and future studies can benefit from analyzing individual-level ride-sourcing and dynamic ridesharing data. However, the methodologically rigorous framework presented in the study captured the random, systematic, and spatial heterogeneity to the extent made possible by the use of aggregate data. Revealed preference-based survey data can enable individual-level analyses of ride-sourcing/carsharing use but are usually limited by sample sizes and may suffer from recall bias. In addition, due to the lack of large-scale real-world trip data, they do not provide insights into the aggregate spatial distribution of ride-sourcing and ridesharing services. The study provided new insights into the (non-linear) complementary and substitutive effects of the built environment, transit accessibility, and active travel. There is a need to examine the potential complementary and substitutive effects at the individual level (such as those available in Household Travel Surveys). Also, while built environment and active travel were both associated with the demand for ride-sourcing and DRS services, the impacts of the potential interactions between active travel and built environment were unexplored. As part of ongoing work, the authors are examining such complex effects using individual-level data. The different data streams utilized in this study do not have an exact temporal alignment. However, compared to the existing studies on this topic, the present study captures an up-to-date profile of neighborhood-built environment in Chicago by harnessing the most recent U.S. EPA Smart Location Database. Finally, despite the methodologically rigorous framework, causal inferences must be avoided given the cross-sectional nature of the study.

Closing remarks

This study contributes by presenting a joint analysis of the demand for ride-sourcing and dynamic ridesharing services and examined how this demand varies across neighborhood-built environments, transit accessibility, behavioral, and sociodemographic factors. Methodologically, a novel and comprehensive generalized geo-additive Markov Random Field-based joint heterogeneous copula framework was presented—simultaneously capturing different layers of random (unobserved), systematic (observed), and spatial heterogeneity. To achieve the objectives, unique large-scale geo-coded data containing millions of ride-sourcing and DRS trips were used. The travel behavior data were spatially joined with up-to-date neighborhood-level data on the built environment, transit accessibility and crime, active travel, and sociodemographic factors.

The analysis enhances our understanding of the demand for ride-sourcing and DRS services in conceptual and methodological ways. Key aspects of the built environment related to urban design (pedestrian-oriented infrastructure), density, and land-use diversity were positively correlated with the demand for ride-sourcing and DRS services—suggesting that sustainable and shared mobility-related goals can be achieved by continuing to invest in compact and more walkable neighborhoods. While pedestrian-oriented connectivity was positively correlated with ride-sourcing and DRS demand, the demand for DRS services was more responsive to improvements in pedestrian-oriented infrastructure. Neighborhoods with greater transit accessibility (compared to the metro region) had a higher demand for ride-sourcing and dynamic ridesharing services.

Additionally, relative to the demand for ride-sourcing services, the demand for DRS services appeared more responsive to an increase in transit accessibility. A potential substitutive effect of accessibility to transit stops on ride-sourcing and DRS services was found. Related to travel behavior, active commute travel (transit use, walking) and telecommuting were strongly and positively correlated with higher ride-sourcing and DRS demand—highlighting the synergies between active travel and shared mobility systems. Contrasting associations of assault and drug/narcotics-related offenses with ride-sourcing and DRS demand were quantified. Collectively, the findings suggest that compact and walkable environments with greater provision of pedestrian-oriented infrastructure can be useful to support sustainable development goals. When taken together with the generally positive role of transit accessibility, the positive links between active travel and shared mobility services highlight the potential for integrating public transit and on-demand mobility services for providing accessible and multimodal travel options.

New methodological insights were also obtained enhancing our understanding of the complex interactions between ride-sourcing and DRS services. First, the demand for ride-sourcing and DRS services was strongly correlated based on latent factors jointly influencing the neighborhood-level demand for ride-sourcing and DRS services. Captured by a Frank copula structure, the joint stochastic dependence was found to be non-linear. Second, systematic and spatial heterogeneity was observed in the stochastic dependence between ride-sourcing and DRS demand as well as the factors influencing the demand for the two mobility options. Besides latent factors, the stochastic dependence between ride-sourcing and DRS services was relatively stronger during the weekend PM period. In addition, characterized by a Markov Random Field structure, structured spatial heterogeneity in the stochastic dependence between the two mobility options was observed. With spillover effects, the joint dependence between ride-sourcing and DRS demand was strongest in the downtown Chicago area and weakest in the outskirts (especially around the airports). Behaviorally, this highlights the strong link between density/urban compactness and mobility-on-demand services. The spatial spillover effect was robust to the effects of time-of-day on the joint stochastic dependence of ride-sourcing and DRS services. Finally, systematic heterogeneity was observed in the effects of the built environment, vehicle ownership, and transit accessibility exhibiting ceiling and overdose effects. Complementary and substitutive effects of transit accessibility were found. Analysis of ceiling and overdose effects showed that increasing transit accessibility in areas with low levels of accessibility (compared to neighborhoods with high levels of transit) could be more helpful in increasing the adoption of ride-sourcing and DRS services. The joint analysis of the two mobility-on-demand services enabled the quantification of the differential impacts of built environment and transit accessibility on the demand for ride-sourcing and DRS services. Overall, the conceptual and methodological findings presented in this paper can help guide the development of policy interventions and investment decisions to capitalize on the benefits of on-demand mobility systems.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

ACS: Commuting (Journey to Work). American Community Survey. United States Census Bureau. https://www.census.gov/topics/employment/commuting.html (2019)

Alemi, F., Circella, G., Handy, S., Mokhtarian, P.: What influences travelers to use Uber? Exploring the factors affecting the adoption of on-demand ride services in California. Travel Behav. Soc. **13**, 88–104 (2018a)

Alemi, F., Circella, G., Mokhtarian, P., Handy, S.: Exploring the latent constructs behind the use of ridehailing in California. J. Choice Model. **29**, 47–62 (2018b)

Anderson, D.N.: "Not just a taxi"? For-profit ridesharing, driver strategies, and VMT. Transportation **41**(5), 1099–1117 (2014)

Azimi, G., Rahimi, A., Asgari, H., Jin, X.: Role of attitudes in transit and auto users' mode choice of ridesourcing. Transp. Res. Rec. **2674**(11), 1–16 (2020)

Bansal, P., Sinha, A., Dua, R., Daziano, R.A.: Eliciting preferences of TNC users and drivers: evidence from the United States. Travel Behav. Soc. **20**, 225–236 (2020)

Bhat, C.R.: Accommodating flexible substitution patterns in multi-dimensional choice modeling: formulation and application to travel mode and departure time choice. Transp. Res. B Methodol. **32**(7), 455–466 (1998)

Bhat, C.R., Eluru, N.: A copula-based approach to accommodate residential self-selection effects in travel behavior modeling. Transp. Res. B Methodol. **43**(7), 749–765 (2009)

Brown, A.E.: Who and where rideshares? Rideshare travel and use in Los Angeles. Transp. Res. A Policy Pract. **136**, 120–134 (2020)

Cervero, R.: Land-use mixing and suburban mobility. UC Berkeley: University of California Transportation Center. https://escholarship.org/uc/item/4nf7k1v9 (1989)

Chan, N.D., Shaheen, S.A.: Ridesharing in North America: Past, present, and future. Transp. Rev. **32**(1), 93–112 (2012)

Chen, Z., Liu, X.C., Wei, R.: Agent-based approach to analyzing the effects of dynamic ridesharing in a multimodal network. Comput. Environ. Urban Syst. **74**, 126–135 (2019)

Clewlow, R. R., Mishra, G. S.: Disruptive transportation: the adoption, utilization, and impacts of ride-hailing in the United States (2017)

Coopers, P.: Sharing or paring? Growth of the sharing economy. In: Proceedings of the 17th Annual Global (2015)

Costanza, R., Fioramonti, L., Kubiszewski, I.: The UN sustainable development goals and the dynamics of well-being. Wiley, Hoboken (2016)

Cui, Y., Makhija, R.S.M.S., Chen, R.B., He, Q., Khani, A.: Understanding and modeling the social preferences for riders in rideshare matching. Transportation **1**, 1–27 (2020)

Dean, M.D., Kockelman, K.M.: Spatial variation in shared ride-hail trip demand and factors contributing to sharing: lessons from Chicago. J. Transp. Geogr. **91**, 102944 (2021)

Deka, D., Fei, D.: A comparison of the personal and neighborhood characteristics associated with ridesourcing, transit use, and driving with NHTS data. J. Transp. Geogr. **76**, 24–33 (2019)

Devaraj, A., Ambi Ramakrishnan, G., Nair, G.S., Srinivasan, K.K., Bhat, C.R., Pinjari, A.R., Ramadurai, G., Pendyala, R.M.: Joint model of application-based ride hailing adoption, intensity of use, and intermediate public transport consideration among workers in Chennai City. Transp. Res. Rec. **2674**(4), 152–164 (2020)

Dias, F.F., Lavieri, P.S., Garikapati, V.M., Astroza, S., Pendyala, R.M., Bhat, C.R.: A behavioral choice model of the use of car-sharing and ride-sourcing services. Transportation **44**(6), 1307–1323 (2017)

Gerte, R., Konduri, K.C., Eluru, N.: Is there a limit to adoption of dynamic ridesharing systems? Evidence from analysis of Uber demand data from New York City. Transp. Res. Rec. **2672**(42), 127–136 (2018)

Ghaffar, A., Mitra, S., Hyland, M.: Modeling determinants of ridesourcing usage: a census tract-level analysis of Chicago. Transp. Res. C Emerg. Technol. **119**, 102769 (2020)

Gomez, J., Aguilera-García, Á., Dias, F.F., Bhat, C.R., Vassallo, J.M.: Adoption and frequency of use of ride-hailing services in a European city: the case of Madrid. Transp. Res. C Emerg. Technol. **131**, 103359 (2021)

Kang, S., Mondal, A., Bhat, A.C., Bhat, C.R.: Pooled versus private ride-hailing: a joint revealed and stated preference analysis recognizing psycho-social factors. Transp. Res. C Emerg. Technol. **124**, 102906 (2021)

Kim, S.-N., Choo, S., Mokhtarian, P.L.: Home-based telecommuting and intra-household interactions in work and non-work travel: a seemingly unrelated censored regression approach. Transp. Res. A Policy Pract. **80**, 197–214 (2015)

Lavieri, P.S., Dias, F.F., Juri, N.R., Kuhr, J., Bhat, C.R.: A model of ridesourcing demand generation and distribution. Transp. Res. Rec. **2672**(46), 31–40 (2018)

Marquet, O.: Spatial distribution of ride-hailing trip demand and its association with walkability and neighborhood characteristics. Cities **106**, 102926 (2020)

Morency, C.: The ambivalence of ridesharing. Transportation **34**(2), 239–253 (2007)

Nashad, T., Yasmin, S., Eluru, N., Lee, J., Abdel-Aty, M.A.: Joint modeling of pedestrian and bicycle crashes: copula-based approach. Transp. Res. Rec. **2601**(1), 119–127 (2016)

Nelsen, R.B.: An introduction to copulas. Springer, New York (2007)

Ratti, C. F., Santi, P.: System for real-time optimal matching of ride sharing requests, Google Patents (2019)

Rayle, L., Dai, D., Chan, N., Cervero, R., Shaheen, S.: Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco. Transp. Policy **45**, 168–178 (2016)

Roughton, L.: A vehicle for crime: ridesharing and victimization. University of Arkansas at Little Rock (2020)

Santi, P., Resta, G., Szell, M., Sobolevsky, S., Strogatz, S.H., Ratti, C.: Quantifying the benefits of vehicle pooling with shareability networks. Proc. Natl. Acad. Sci. **111**(37), 13290–13294 (2014)

Sener, I.N., Eluru, N., Bhat, C.R.: On jointly analyzing the physical activity participation levels of individuals in a family unit using a multivariate copula framework. J. Choice Model. **3**(3), 1–38 (2010)

Shaheen, S., Chan, N.: Mobility and the sharing economy: potential to facilitate the first-and last-mile public transit connections. Built Environ. **42**(4), 573–588 (2016)

Sikder, S.: Who uses ride-hailing services in the United States? Transp. Res. Rec. **2673**(12), 40–54 (2019)

Sørbye, S.H., Rue, H.: Scaling intrinsic Gaussian Markov random field priors in spatial modelling. Spat. Stat. **8**, 39–51 (2014)

Sperling, D.: Three revolutions: steering automated, shared, and electric vehicles to a better future. Island Press, Washington (2018)

Stathopoulos, A., Sener, I.N.: Transforming mobility systems with sharing and automation. Transportation **44**(6), 1255–1259 (2017)

Stiglic, M., Agatz, N., Savelsbergh, M., Gradisar, M.: Enhancing urban mobility: integrating ride-sharing and public transit. Comput. Oper. Res. **90**, 12–21 (2018)

Stocker, A., Shaheen, S.: Shared automated vehicles: review of business models. International Transport Forum Discussion Paper (2017)

Su, D.N., Nguyen-Phuoc, D.Q., Johnson, L.W.: Effects of perceived safety, involvement and perceived service quality on loyalty intention among ride-sourcing passengers. Transportation **48**(1), 369–393 (2021)

Tachet, R., Sagarra, O., Santi, P., Resta, G., Szell, M., Strogatz, S.H., Ratti, C.: Scaling law of urban ride sharing. Sci. Rep. **7**(1), 1–6 (2017)

Tribby, C.P., Graubard, B.I., Berrigan, D.: National and metropolitan trends in public transit use, transit-related walking, and ridesharing between 2009 and 2017. J. Transp. Health **19**, 100918 (2020)

Trivedi, P.K., Zimmer, D.M.: Copula modeling: an introduction for practitioners. Now Publishers Inc, The Netherlands (2007)

Tu, W., Santi, P., He, X., Zhao, T., Liu, X., Li, Q., Wallington, T.J., Keoleian, G.A., Ratti, C.: Understanding ridesourcing mobility and the future of electrification: a comparative study in Beijing. J. Urban Technol. **1**, 1–20 (2020)

United Nations: The 17 Goals. United Nations. Department of Economic and Social Affairs. Sustainable Development. https://sdgs.un.org/goals. (2021)

US Census: QuickFacts United States. United States Census Bureau. https://www.census.gov/quickfacts/fact/table/US/PST045221 (2021)

Wali, B., Khattak, A.J.: A joint behavioral choice model for adoption of automated vehicle ride sourcing and carsharing technologies: role of built environment and sustainable travel behaviors. Transp. Res. C Emerg. Technol. **136**, 103557 (2022)

Wali, B., Greene, D.L., Khattak, A.J., Liu, J.: Analyzing within garage fuel economy gaps to support vehicle purchasing decisions–a copula-based modeling and forecasting approach. Transp. Res. D Transp. Environ. **63**, 186–208 (2018)

Wali, B., Frank, L.D., Chapman, J.E., Fox, E.H.: Developing policy thresholds for objectively measured environmental features to support active travel. Transp. Res. D Transp. Environ. **90**, 102678 (2021a)

Wali, B., Santi, P., Ratti, C.: Modeling consumer affinity towards adopting partially and fully automated vehicles: the role of preference heterogeneity at different geographic levels. Forthcom. Transp. Res. C Emerg. Technol. (2021b)

Wali, B.: (Under-review) modeling consumers' use of carsharing and ridesharing services: an integrated finite mixture copula-based behavioral choice framework. Presented at the 101st TRB Annual Meeting (2022)

Walls, M., Safirova, E.: A review of the literature on telecommuting and its implications for vehicle travel and emissions. Resources for the Future. https://ageconsearch.umn.edu/record/10492/ (2004)

Walsh, B.: Today's smart choice: Don't own. Share. Time Int. **1**(3), 49 (2011)

Wang, H., Yang, H.: Ridesourcing systems: a framework and review. Transp. Res. B Methodol. **129**, 122–155 (2019)

Washington, S.P., Karlaftis, M.G., Mannering, F.: Statistical and econometric methods for transportation data analysis. Chapman and Hall/CRC, New York (2003)

Wood, S.N.: Generalized additive models: an introduction with R. Chapman and Hall/CRC, New York (2006)

Wood, S. N., Goude, Y., Shaw, S.: Generalized additive models for large data sets. J. R. Stat. Soc. Ser. C Appl. Stat. 64(1):139–155 (2015a)

Wood, S. N., Y. Goude, S. Shaw (2015b). Generalized additive models for large data sets. J. Stat. Soc. Seri. C Appl. Stat. 64(1): 139–155.

Xu, Y., Yan, X., Liu, X., Zhao, X.: Identifying key factors associated with ridesplitting adoption rate and modeling their nonlinear relationships. Transp. Res. A Policy Pract. **144**, 170–188 (2021)

Yan, L., Luo, X., Zhu, R., Santi, P., Wang, H., Wang, D., Zhang, S., Ratti, C.: Quantifying and analyzing traffic emission reductions from ridesharing: A case study of Shanghai. Transp. Res. D Transp. Environ. **89**, 102629 (2020)

Yasmin, S., Eluru, N., Pinjari, A.R., Tay, R.: Examining driver injury severity in two vehicle crashes—a copula based approach. Accid. Anal. Prevent. **66**, 120–135 (2014)

Yu, H., Peng, Z.-R.: The impacts of built environment on ridesourcing demand: A neighbourhood level analysis in Austin, Texas. Urban Stud. **57**(1), 152–175 (2020)

Yu, B., Ma, Y., Xue, M., Tang, B., Wang, B., Yan, J., Wei, Y.-M.: Environmental benefits from ridesharing: A case of Beijing. Appl. Energy **191**, 141–152 (2017)

Zhang, Y., Zhang, Y.: Exploring the relationship between ridesharing and public transit use in the United States. Int. J. Environ. Res. Public Health **15**(8), 1763 (2018)

Behram Wali is lead research scientist at Urban Design 4 Health, Inc. His research harnesses methodological developments and cyber-physical systems to examine to role of complex human behaviors in surface transportation with relevance to proactive safety, smart cities, and urban health. He completed his PhD in transportation engineering from University of Tennessee, Knoxville and attended MIT Senseable City Lab for his postdoctoral training.

Paolo Santi is Principal Research Scientist at MIT Senseable City Lab and Research Director at the Istituto di Informatica e Telematica, CNR, Pisa, Italy. Dr. Santi holds a "Laurea" degree and the PhD in computer science from the University of Pisa, Italy. Dr. Santi is a member of the IEEE Computer Society and has recently been recognized as Distinguished Scientist by the Association for Computing Machinery. His research interest is in the modeling and analysis of complex systems ranging from wireless multi hop networks to sensor and vehicular networks and, more recently, smart mobility and intelligent transportation systems. In these fields, he has contributed more than 160 scientific papers and two books.

Carlo Ratti received the degrees in civil engineering from the Politecnico di Torino and the École nationale des ponts et chaussées and the M.Phil. and Ph.D. degrees from the University of Cambridge, U.K. He is currently the Founder and the Director of the MIT Senseable City Laboratory. As an Architect and an Engineer by training, he practices in Italy and teaches at the Massachusetts Institute of Technology. He has coauthored over 200 publications and holds several patents. His work has been exhibited worldwide at venues, such as the Venice Biennale, the Design Museum of Barcelona, Barcelona, the Science Museum, London, GAFTA, San Francisco, and The Museum of Modern Art, New York. His Digital Water Pavilion at the 2008 World Expo was hailed by Time Magazine as one of the Best Inventions of the Year. He has been included in Esquire Magazine's Best and Brightest list, Blueprint Magazine's 25 People Who Will Change the World of Design, and Forbes Magazine's Names You Need To Know in 2011. He was a Presenter at TED 2011 and is serving as a member of the World Economic Forum Global Agenda Council for Urban Management. He is a regular Contributor to the architecture magazine Domus and the Italian newspaper Il Sole 24 Ore. He has also written as an op-ed contributor for BBC, La Stampa, Scientific American, and The New York Times.

Authors and Affiliations

Behram Wali[1] · Paolo Santi[1,2] · Carlo Ratti[3]

✉ Paolo Santi
psanti@mit.edu

Behram Wali
walibehram@yahoo.com; bwali@mit.edu

Carlo Ratti
ratti@mit.edu

[1] Senseable City Lab, Massachusetts Institute of Technology, Cambridge, MA, USA

[2] Istituto di Informatica e Telematica, Pisa, Italy

[3] Senseable City Lab, Urban Technologies and Planning, Massachusetts Institute of Technology, Cambridge, MA, USA