Data Science for Transportation (2024) 6:14
https://doi.org/10.1007/s42421-024-00100-6

RESEARCH

Multi-modal Machine Learning Investigation of Telework and Transit Connections

Deirdre Edward[1] · Jason Soria[2] · Amanda Stathopoulos[3]

Received: 23 March 2024 / Revised: 23 May 2024 / Accepted: 20 June 2024 / Published online: 3 July 2024

Abstract
Public transit in the U.S. has an unsettled future. The onset of the COVID-19 pandemic saw a dramatic decline in transit ridership, with agency operations, and user perceptions of safety changing significantly. However, one new factor beyond the control of agencies is playing an outsized role in transit ridership: the shifting employment patterns in the hybrid work era. Indeed, a lasting and widespread adoption of telework has emerged as a key determinant of individual transit behaviors. This study investigates the impact of teleworking on public transit ridership changes across the different transit services in the Chicago area during the pandemic, employing a random forest machine learning approach applied to large-scale survey data (n = 5637). The use of ensemble machine learning enables a data-driven investigation that is tailored for each of the three main transit service operators in Chicago (Chicago Transit Authority, Metra, and Pace). The analysis reveals that the number of teleworking days per week is a highly significant predictor of lapsed ridership. As a result, commuter-centric transit modes—such as Metra—saw the greatest declines in ridership during the pandemic. The study's findings highlight the need for transit agencies to adapt to the enduring trend of teleworking, considering its implications for future ridership and transportation equity. Policy recommendations include promoting non-commute transit use and addressing the needs of demographic groups less likely to telework. The study contributes to the understanding of how telework trends influence public transit usage and offers insights for transit agencies navigating the post-pandemic world.

Keywords COVID-19 pandemic · Transit ridership · Telework · Random forest analysis · Machine learning · Public transportation

Introduction

At the onset of the COVID-19 pandemic, public transit ridership experienced a precipitous drop, representing a 100-year low in the United States (Ziedan et al. 2023a). As COVID-19 shifted from being an emergent pandemic to a permanent fixture of our lives, transit usage has not recovered fully (Wilbur et al. 2023). The decline in transit ridership is not uniform, however, with variation depending on the type of transit service (Soria et al. 2023). Between August 2019 and August 2020, rail ridership declined by 72%, and bus ridership declined by 37% (Polzin et al. 2021). The sustained ridership loss poses a substantial financial challenge for transit operators, as well as a matter of social equity. Low-income, essential workers, and socially disadvantaged individuals are the most likely to rely on transit services remaining available (Soria et al. 2023; Griffin and Sener 2016; El-Geneidy et al. 2016). During the pandemic, groups that saw smaller long-term declines in ridership were typically composed of dependent riders that used transit for urgent and emergency healthcare, minimal maintenance activities such as grocery shopping, and transportation services meant to move essential workers and travelers to these activities (Liu et al. 2020). Instead, in the case of Chicago, ridership declined more in areas with higher percentages of white, educated, and high-income individuals (Hu and Chen 2021).

✉ Amanda Stathopoulos
a-stathopoulos@northwestern.edu

[1] Department of Engineering Sciences and Applied Mathematics, Northwestern University, 2145 Sheridan Rd, Evanston, IL 60208, USA

[2] Department of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0315, USA

[3] Department of Civil and Environmental Engineering, Northwestern University, 2145 Sheridan Rd, Evanston, IL 60208, USA

Some of the factors determining ridership changes were under the direct control of operators. An analysis of 40 North American cities revealed that most transit agencies heavily reduced their frequency of service (DeWeese et al. 2020). Several agencies cut their services equally across service areas and found that vulnerable communities (based on low income, population that is non-white, and population without a bachelor's degree) were disproportionately affected, highlighting a concern for equitable transportation access during the pandemic. While cutting service and implementing preventative measures such as social distancing, transit agencies contended with labor shortages that caused further degradation of service quality (Freishtat 2021; Mack et al. 2021).

In Chicago, however, the Chicago Transit Authority (CTA) service was not radically reduced, and public transit remained operational (Caine 2021). Nevertheless, Chicago still saw a drastic decline in ridership at the beginning of the pandemic (Soria et al. 2023). In 2020, the week of April 5–April 11 marked the steepest drop in ridership levels, with Metra (the commuter rail service) at 3% of pre-pandemic levels and CTA (urban rail and buses) at 22% (APTA 2024). After this initial plummet, these figures grew marginally and remained relatively steady for the next year. From June 2020 until the end of February 2021 (which overlaps with the data collection phase of this study), Metra ridership hovered around 10%, and CTA ridership was at roughly 30% (APTA 2024). After the first year, these levels began to improve, but overall ridership loss has proven persistent with average transit ridership four years after the start of the pandemic hovering around 60% of pre-pandemic levels for Metra and 70% for the CTA services according to APTA tracking (APTA 2024).

Much of what determines transit ridership lies beyond the direct control of transit agencies. Among the external factors, the closure of non-essential activities severely reduced the demand for travel. Out-of-home activity participation was reduced by approximately half with lower-income groups more likely to reduce activities (Fatmi et al. 2021). Less obvious is the indirect effect of activity restrictions. For example, school closures pushed women to remain at home because of a shift in domestic responsibilities (He et al. 2022). Moreover, the health risks associated with shared spaces during the pandemic changed user motivations regarding transit (Rahimi et al. 2021).

However, the lasting and widespread adoption of telework has emerged as a pivotal determinant of individual transit behaviors going forward. While lockdown conditions, service cuts, and rider fears around sharing transit facilities have not persisted over time, telework arrangements have proven durable (Tahlyan et al. 2024). The adoption of telework practice reduces or removes the need to commute to work, and shifts individual motivations regarding transit (Ziedan et al. 2023b). As the adoption of hybrid and remote work was initially forced on many workers, experiences were mixed (Martin et al. 2022; Tahlyan et al. 2022). With greater adoption, though, many are reprioritizing their long-term objectives and increasingly seeking teleworking opportunities (Venkataramani 2021). Several years after the start of the pandemic, most people who telework do so by choice and not out of necessity (Parker et al. 2022). According to the American Time Use Study (ATUS), 33.8% of employed Americans still telework from home in 2022 (BLS 2024).

The goal of this study is to examine the relationship between the shifting employment patterns in the hybrid work era, and changes in the use of transit. Specifically, we investigate the differences in ridership modifications for different types of transit services. The study uses random forest—an ensemble learning method that increases the robustness of predictions—to demonstrate that telework, even if part-time, is a key determinant of individual public transit behaviors. As a result, we gain valuable insight into different challenges agencies face in seeking to recover ridership. Although COVID-19 is no longer perceived as an active emergency, ridership patterns have remained unstable (Lei and Ukkusuri 2022) and the endurance of telework trends warrants serious consideration by transit agencies to navigate reduced demand and shifting motivations regarding transit in a post-pandemic world. Specifically, we gain formal insight into different strategies that transit agencies need to consider as a function of different transit options with differing customer makeup.

Literature Review

Drastic declines in transit ridership were observed during the onset of the pandemic. Though there has been a steady growth in ridership across North America toward pre-pandemic levels, it is unlikely that pre-pandemic travel patterns will return. Rather, it is more likely that *new* patterns will emerge, ones that incorporate pandemic-era behavior. A key factor shaping commuting patterns is the accelerated adoption of teleworking due to the pandemic (Vickerman 2021). This literature review first focuses on the impact of teleworking on travel behavior, and highlights ridership trends for different transit modes, agencies, and customers. Then, this review discusses the use of machine learning to study the impact of telework on travel behavior.

Telework Trends Post-pandemic

Due to the pandemic, many companies were forced to adopt teleworking to continue operations. This quickly spurred the adoption of new technologies by employers and employees alike, accelerated the adoption of remote access services, and reduced transit ridership. In 2019, 5.2% of Chicago-area households reported at least one day of telework per week, which was consistent with the national average (Farmer 2022). By August 2020, this figure had grown to 43% in Chicago, and 36.3% nationwide (Farmer 2022). Analysis from 2024 indicates that these trends have persisted, with some form of hybrid work being the norm at most companies (Tahlyan et al. 2024). Along with teleworking, between 30 and 50% of survey respondents reported greater frequency of remote activities such as e-commerce and telehealth (accessing healthcare via telecommunication technology) (Abdullah et al. 2020; Beck and Hensher 2020; Mouratidis and Papagiannakis 2021).

Remote or hybrid work was significantly less widespread before the COVID-19 pandemic (Parker et al. 2020). Adoption of teleworking prior to the pandemic has been linked to being white, being highly educated, having children at home, and having relatively high incomes (Plaut 2005; Popuri and Bhat 2003). Beyond sociodemographics, attitudes regarding satisfaction with long-term objectives are linked with greater preferences towards teleworking (Mokhtarian and Salomon 1997). For example, workers who have a greater preference for teleworking also have more "family drive," which is the desire to spend more time with family (Mokhtarian and Salomon 1997).

During the pandemic, similar sociodemographic profiles are adopting telework en masse, and this disparity is in large part due to the telework capability of different job types. The predominant trend in teleworking is toward high-income, well-educated, and non-minority households (Barbour et al. 2021; Matson et al. 2021). The "drives" described by Mokhtarian and Salomon (1997) may very well have regained relevance in today's debate on teleworking and hybrid work arrangements. These new workplace arrangements carry several implications for transit ridership (see international perspectives in: Beck and Hensher 2020; Nayak and Pandit 2021; Olde Kalter et al. 2021).

The most direct effect on transit ridership consists of a lowered travel demand. Several researchers found that behavioral inertia such as lingering safety concerns as well as new travel patterns adopted during the pandemic, including teleworking, have affected ridership (Rothengatter et al. 2021; Vickerman 2021). Indirectly, Reuschke and Ekinsmyth explore the implications of teleworking and residential choice, where workers are faced with new possibilities to not be tied spatially to their physical workplace (Reuschke and Ekinsmyth 2021). Furthermore, an Australian survey found that most respondents who teleworked during the pandemic intended to continue this practice to a higher degree than they had before COVID-19 (Beck et al. 2020). A nationwide US study found positive intentions toward increased telework among high-income, and highly educated workers with long commutes (Mohammadi et al. 2023). There is growing evidence that employees and potential hires regard working from home as an important job benefit going forward (Rahman and Arif 2020).

Machine Learning as a Tool to Study Ridership Trends

Given the unique and unprecedented circumstances that the coronavirus pandemic presented and the abundance of variables that may be related to ridership in this context, machine learning techniques are a useful means to determine which variables are significant and quantify their influence. Historically, survey-based policy research has relied on inference-based classical statistics (Nardi 2018). Typically, these techniques will use a small number of input variables and make explicit assumptions about the relationship being studied (Karlaftis and Vlahogianni 2011; Mannering et al. 2020). However, unlike classical statistics, machine learning focuses on utilizing extensive and unwieldy data, and these algorithms do not require a priori knowledge of the underlying structure of a model (Karlaftis and Vlahogianni 2011; Bzdok et al. 2018; Van Cranenburgh et al. 2022). Within public policy, existing machine learning methods have not been adequately evaluated in practice, and many of these methods are not designed to be context specific (see Amarasinghe et al. 2023 for an overview and Noursalehi et al. 2020; Sun et al. 2024 for specific contextual applications). Combined with the question of interpretability, this set of issues has hindered the process of developing sufficient trust in these models and convincing regulators to adopt policy informed by machine learning, although post hoc evaluation approaches such as feature importance may be helpful (Amarasinghe et al. 2023). Furthermore, there is significant debate about the existence of a tradeoff between the explainability and accuracy of machine learning models (Bell et al. 2022; Rodolfa et al. 2021).

Studies like those by Wang et al. (2020) have demonstrated that machine learning can be effectively used for choice analysis, providing insights that rival traditional discrete choice models. This includes detailed analyses of choice probabilities, market shares, and the substitution patterns of alternatives, which are crucial for understanding shifts in ridership trends (Wang et al. 2020). Machine learning techniques have also been employed more broadly to explore travel changes in the disruptive aftermath of the pandemic (Mourtakos et al. 2024), and understand differences by socio-economic status (Li et al. 2023). For transit analysis, research by Sekadakis et al. (2023) has utilized these methods to analyze driving behaviors, which indirectly influence transit usage patterns. Their findings underscore the importance of considering a range of behavioral responses when predicting the future of transit (Sekadakis et al. 2023). Other recent transportation applications of machine learning include tweet-based sentiments linked to mobility (Sun et al. 2023) and on-demand microtransit use (Zhou et al. 2021).

Despite the advantages, the application of ML in public policy, particularly transportation, faces challenges related to model interpretability and the "black box" nature of certain algorithms. The underlying predictive algorithms are sometimes unintelligible to humans, or at least very difficult to interpret and contextualize. For example, deep learning algorithms tend to be black boxes because they are deeply recursive and typically too intricate to disentangle (Rudin and Radin 2019). This concern is particularly poignant when policy decisions must be explained transparently to stakeholders. Addressing this, researchers like Rudin and Radin (2019) advocate for the development of interpretable ML models that provide both high accuracy and ease of explanation, thus enhancing the trustworthiness and applicability of ML insights in public decision-making (Rudin and Radin 2019).

Compared to other machine learning algorithms, random forest (RF) has been identified as highly computationally efficient in predicting travel behavior, and it demonstrates high predictive power (Wang et al. 2021). RF algorithms have shown superior or roughly equivalent performance for a variety of big-data applications, e.g., mode detection (Efthymiou et al. 2019; Wang et al. 2021; Sadeghian et al. 2022). Furthermore, RF is more interpretable than comparably accurate methods (Yan and Shen 2022). The explainability of RF rebuts the "black box" nature of more complex models by offering a clearer understanding of how input variables affect outputs, and facilitating the adoption of findings by policy-makers who rely on comprehensible models to make informed decisions. Additionally, the ensemble nature of RF, which combines multiple decision trees, helps in handling high-dimensional and multicollinear data effectively (Yan and Shen 2022). This characteristic is crucial for analyzing complex interactions in travel data, where numerous variables might interact in non-linear ways. RF's method of using bootstrap aggregating (bagging) to sample data for each tree in the forest enhances the model's robustness against overfitting and ensures a more reliable representation of the underlying data (Brodeur et al. 2020). The ability of RF to provide consistent and accurate predictions even with the inclusion of many predictor variables and in the presence of potential noise in the data builds trust in its outputs (Yan and Shen 2022).

Literature Takeaways and Motivation for Study

Under unpredictable circumstances such as the coronavirus pandemic, machine learning techniques enable an open exploration of factors that influence transit use, such as telework. Specifically, we use random forest and experiment with teleworking data configurations to quantify the extent to which regular teleworking causes individuals to stop using transit.

Data

The data used in this study are from a survey collected by the Chicago Regional Transportation Authority (RTA). This survey contains N = 5637 responses from two waves—November 9, 2020–December 4, 2020, and January 19, 2021–February 5, 2021. The survey was conducted online, and participants were transit users recruited by email invitations. A study by RTA found that respondents who continued to use transit during the survey period were disproportionately Black, Latino, low-income, or essential workers (RTA 2021). This reflects a higher reliance on public transit of these population segments, where this difference may have been heightened due to the severity and ubiquity of the pandemic. Quota-based sampling was applied to ensure sufficient representation of each transit operator in the database. Quality control measures included screening out respondents who completed the study too quickly, or provided inconsistent answers (more details on sampling and quality control are found in RTA 2021).

Study Context

Because of the different coverage areas, service types, and rider patterns associated with the three service operators governed by RTA—CTA, Metra, and Pace—the analysis was separated by mode. Figure 1 shows the network maps of the three operators. CTA oversees both bus and rail services and has the highest level of ridership in the region historically (RTA 2024). For context, CTA focuses on downtown and radial operations with buses operating on 127 routes and trains on eight rail lines, covering 224.1 miles (CTA 2024a). CTA trains and buses typically run every 10–20 min through late evening (CTA 2024a). In terms of ridership profile, work commuting was the main purpose of travel making up 62% of trips in 2017 (RTA 2017a) down to 58% in 2022 (RTA 2022). Metra differs from CTA in its broader regional focus, extending beyond the immediate Chicago metropolitan area to cater to suburban commuters, reflecting its historical development from multiple private railroads to a consolidated commuter system (Metra 2024a). Given its focus on commuters, Metra's trains run frequently during rush hour, and every hour or two during off-peak times. Historically 93% of Metra trips were for work commuting (RTA 2017a), dropping to 77% in 2022 (RTA 2022). Metra operates 242 stations across 11 rail lines, making it one of

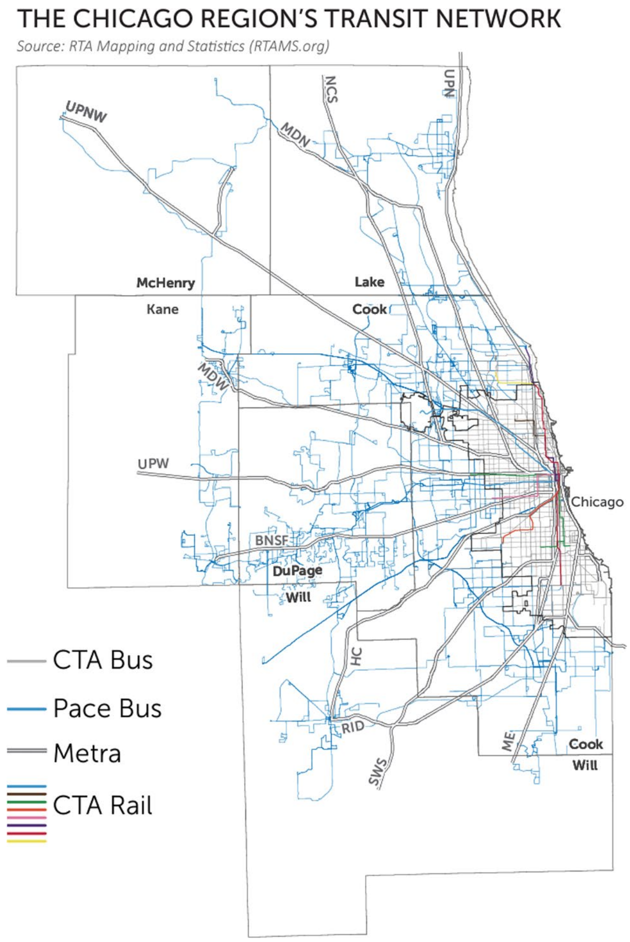


Fig. 1 Regional transit service area map (CTA, Metra, and Pace). Notes on Source: (RTA 2017b)

the busiest commuter rail systems in the United States. Pace, the Suburban Bus Division of the RTA, complements the urban and commuter rail services by focusing on suburban bus operations. Pace offers extensive coverage with fixed bus routes, vanpool services, and paratransit services catering to the needs of customers with disabilities (Pace n.d.). This makes Pace a critical component of the region's transit system, especially for areas not directly served by CTA or Metra. However, Pace service is relatively infrequent, with buses running every 30–60 min through mid-evening (CTA 2024b). The ridership profile for Pace has diversified, with work-commute trips comprising 66% of trips pre-pandemic (RTA 2017a) down to 52% of commute trips in 2022 (RTA 2022). Demographic profiles differ overall, with Metra respondents having higher income and car availability compared to the other service boards, and Pace riders are relatively more likely to identify as African American/Black (46%) and not having a car available (79%) (RTA 2022). Looking at ridership volumes, reporting shows that bus services have recovered more consistently in the years following the pandemic, while rail-only Metra lags behind in the recovery (ILEPI 2021).

Exploratory Analysis of Mode-Specific Rider-Cessation

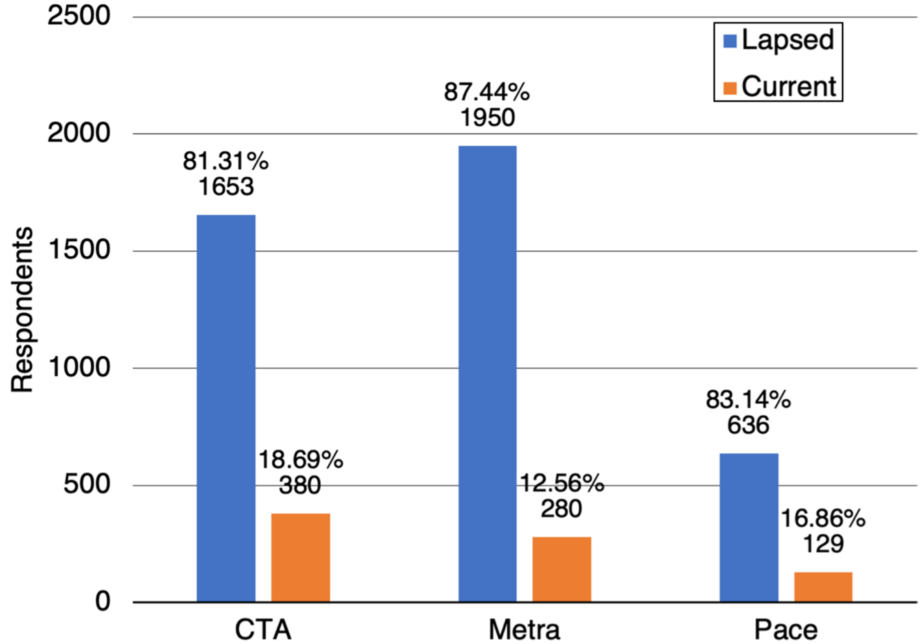
The outcome variable of interest is whether the user abandoned transit at the time of the study (shown in Fig. 2). The explanatory variables that are considered in this study are: the number of days spent teleworking per week, age range, ethnicity, annual household income, number of household vehicles, and number of people in the household (shown in Table 1). This study excludes data points that did not include responses to the telework status or sociodemographic variables used to build each model. Additionally, this study excludes respondents who were unemployed at the time of the study. Respondents were only included if they reported currently or previously using that mode. This study defines a current rider as someone who actively uses an RTA service, and a lapsed rider as someone who reported previously using that service but is no longer an active rider.

From the initial 5637 observations included by RTA, the sample sizes for each model were N = 2033 for CTA, N = 2230 for Metra, and N = 765 for Pace. These sample sizes account for respondents who may have used more than one service. In Fig. 2, we see the proportion of riders from each mode who lapsed or continued using the service. Evidently, these data are imbalanced—far more respondents lapsed than remained current. However, this is consistent with general ridership trends from that period (Soria et al. 2023).

We see in Fig. 3 that a vast majority of respondents telework full time, defined as 5 or more days. This is consistent with the distribution of lapsed riders. However, about half of current CTA and Pace riders, and three-fourths of Metra riders do not telework at all or telework less than full-time.

Table 1 describes the distribution of the input variables used. From this, we can see that respondents tend to be disproportionately older, wealthier, and less likely to identify as non-white than the Chicago average (U.S. Census Bureau 2023). However, other demographic groups are still represented in this sample. We can also see that Metra riders in this sample are more likely to telework, less likely to identify as non-white, and are wealthier on average, followed by CTA, and Pace. This is consistent with the finding that Metra is often used by workers commuting from the more racially and socioeconomically homogenous Chicago suburbs into the city. Pace riders tend to be older than Metra or CTA riders, which likely reflects the fact that Pace prioritizes ADA compliance (Pace 2024).

Fig. 2 Number of lapsed and current riders by service operator (CTA, Metra, and Pace)



Association Between Different Variables by Mode

Here, the survey data are used to explore the association among the core sociodemographic, household, vehicle, and telework status variables with lapsed ridership status during the pandemic in the winter of 2020–2021. This descriptive analysis further guides the Random Forest variable selection and feature selection process. Furthermore, it provides important insight into the correlation among explanatory variables, such as between income and number of household vehicles, to better parse later analytical findings.

As seen in Table 2, we use different measures of association due to the different variable types. For all three modes, the measures of association appear to be low for the wave number. This suggests that there is no significant difference in variable distribution between waves 1 and 2. Therefore, they will be considered jointly moving forward.

Looking in turn at each of the transit operators, we note some recurring patterns, and context-specific effects. The number of days spent teleworking, number of household vehicles, household size, and household income bracket are the variables most strongly correlated with lapsed CTA ridership. Household income is correlated with both the number of household vehicles and number of days spent teleworking. Household size is not correlated with the number of days spent teleworking, but it is correlated with income bracket and number of vehicles. The number of days spent teleworking is by far the most highly correlated variable with lapsed Metra ridership. Age range and household income bracket also appear to be significantly associated in this context. Unlike with CTA, the number of household vehicles does not seem to be correlated with Metra ridership. We do not see a high Theta value for the respondent's household size, either. Similarly to the CTA data, the Metra data indicates that household income is correlated with the number of days spent teleworking.

Finally, the number of days spent teleworking, number of household vehicles, and household income bracket are the variables most correlated with lapsed Pace ridership. The Lambda value for Pace ridership and ethnicity is marginally higher than the corresponding values for Metra and Pace, although it is still low.

Methodology

This study uses a random forest (RF) classification algorithm to predict whether a rider has lapsed transit ridership. The RF algorithm has been identified as a promising benchmark for travel choice models in a meta-analysis of 35 studies (Wang et al. 2021). Given the differences in behavior and passenger characteristics between CTA, Metra, and Pace, this model is built separately for each of the three transit service operators.

The explanatory variables are described in the data section of this paper and include the number of days spent teleworking per week. Model performance is assessed using area under the ROC curve (AUC) and out-of-bag (OOB) error rate. Given the disparity between the numbers of lapsed and current riders, random over-sampling was tested using random over-sampling examples (ROSE), which addresses imbalanced binary classification problems by artificially generating samples of the smaller class (Lunardon et al. 2014). However,

Table 1 Description of all explanatory data used for CTA, Metra, and Pace

Variable	Value	Percent, CTA N = 2033	Percent, Metra N = 2230	Percent, Pace N = 765
Age	Under 18	0.000%	0.000%	0.000%
	18–24	3.935%	1.794%	1.830%
	25–34	28.578%	15.695%	14.118%
	35–44	24.545%	21.704%	21.699%
	45–54	17.511%	24.664%	24.314%
	55–64	19.233%	27.578%	28.366%
	65–74	5.755%	8.206%	8.758%
	75 and older	0.443%	0.359%	0.915%
Ethnicity	Mixed race	4.919%	3.543%	4.052%
	American Indian or Alaskan Native	0.197%	0.135%	0.131%
	Asian	5.755%	4.395%	6.928%
	Black or African American	9.493%	8.161%	11.765%
	Hispanic, Latino, or Spanish	6.690%	3.498%	5.229%
	Middle Eastern or North African	0.492%	0.673%	0.654%
	Native Hawaiian or other Pacific Islander	0.197%	0.224%	0.261%
	White	72.258%	79.372%	70.980%
Annual household income	Under $25,000	2.656%	1.525%	4.183%
	$25,000–$49,999	9.297%	4.574%	10.458%
	$50,000–$74,999	16.921%	13.991%	16.993%
	$75,000–$99,999	17.216%	17.085%	20.654%
	$100,000–$199,999	37.777%	42.377%	35.556%
	$200,000 or more	16.134%	20.448%	12.157%
Number of household vehicles	0	17.855%	10.045%	16.209%
	1	43.483%	32.960%	36.993%
	2	29.169%	39.596%	32.026%
	3	6.394%	11.659%	9.935%
	4 or more vehicles	3.099%	5.740%	4.837%
Size of household	0 (I am living alone)	21.741%	17.578%	18.301%
	1 person	43.483%	40.493%	38.562%
	2 people	16.478%	17.892%	19.739%
	3 people	11.805%	15.785%	13.856%
	4 people	4.230%	6.009%	4.967%
	5 or more people	2.263%	2.108%	4.444%
Telework days per week	0	8.313%	6.188%	12.157%
	1	2.361%	2.108%	3.268%
	2	2.755%	2.780%	3.660%
	3	4.083%	3.946%	4.575%
	4	5.558%	6.547%	5.229%
	5	53.763%	57.175%	50.458%
	6	4.771%	4.978%	6.405%
	7	18.396%	16.278%	14.248%

this procedure did not improve AUC by more than 0.05, and therefore was not used in building the final models. Finally, variable significance is calculated to find determinants of lapsed ridership. This is done using permutation importance.

Performance Metrics: AUC and OOB Error

The performance metrics used in this study are OOB error and AUC. These values describe two different traits of model performance: raw performance and the ability to separate classes, respectively. The OOB error rate is computed using

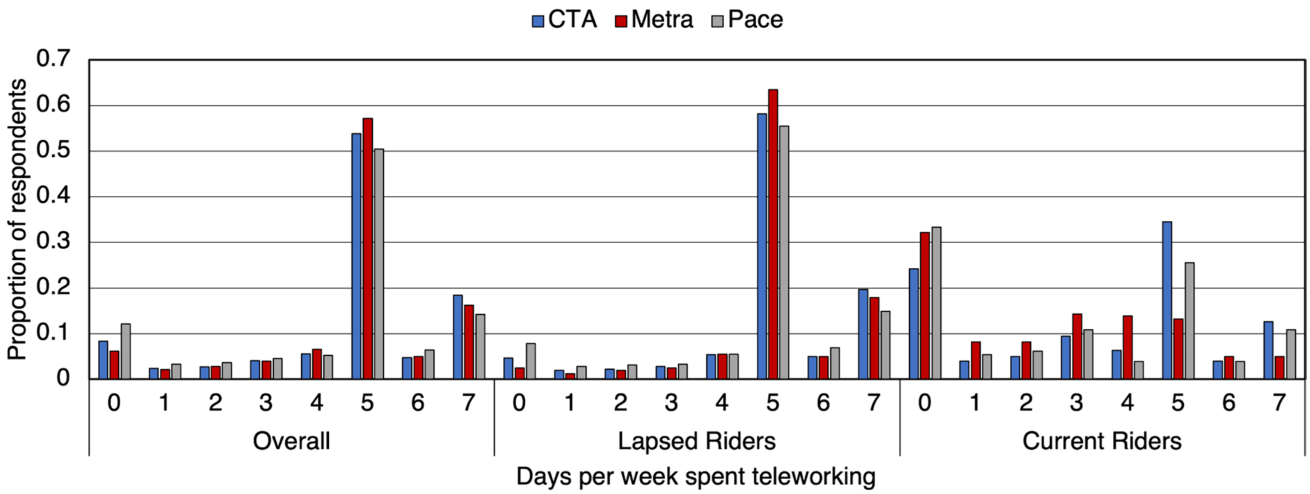


Fig. 3 Overall distribution of teleworking days per week for CTA, Metra, and Pace riders, as well as for current and lapsed riders

the testing data for each individual tree within a forest, and then these values are averaged. The OOB error rate is used to measure variable importance. To calculate AUC, we integrate the receiver operating characteristics (ROC) curve, which plots the false positive rate against the true positive rate as the RF discrimination threshold is varied. AUC represents the likelihood that a randomly selected positive value will rank above a randomly selected negative value. As such, it indicates the extent to which lapsed and current riders can be distinguished based on our explanatory variables.

Random Forest

RF algorithms defy the interpretation issues typically associated with machine learning, as they exhibit good performance and are amenable to feature importance calculations, which enables interpretation (Orlenko and Moore 2021; Wang et al. 2021). First introduced in 2001, RF is an ensemble machine learning method that aggregates the results of independent tree models (Breiman 2001). RF applies two types of bootstrap aggregation, or bagging, to reduce variance and prevent overfitting. For each tree, the first bagging method draws a random sample (with replacement) to be used as the training data. The testing data is used to compute the OOB error rate. These models vote to form the final prediction of the forest. The second method is known as feature bagging, which selects a random subset of the explanatory variables to consider at each candidate split within a tree. RFs are inherently able to handle multicollinearity due to the feature bagging used, as each feature will be left out of some trees (Tomaschek et al. 2018).

Hyperparameter Selection

The RF algorithm has several key hyperparameters that must be set before building the model. These are the number of trees to grow, the number of variables sampled in feature bagging, and the maximum depth of trees. There is no default “best” value for the number of trees or maximum depth, and typically, the default value for the number of variables sampled is $\sqrt{p}$ rounded down (Probst et al. 2019), where p is the total number of variables in our data. In this context, $\sqrt{p} = \sqrt{6} = 2.44$. Through a grid search method using 2, 4, and 6 variables, we find that 2 is indeed the optimal value across modes for this hyperparameter. Existing literature has suggested that performance—measured here by AUC—plateaus beyond a certain number of trees (Probst and Boulesteix 2018), and that the best value for this hyperparameter is the minimum number necessary to reach this threshold (Oshiro et al. 2012). The values tested for the number of trees are 50, 100, and 150, and there appears to be no significant difference in performance. As such, this hyperparameter is set to 50 for the sake of computation time. The values tested for maximum depth are 10, 5, and 2. As 10 is consistently the best-performing value, this is used for all three modes.

Model Comparison

To verify the efficacy of random forest for this context, we compare its AUC with three other models: XGBoost, decision tree, and logistic regression. The performance metrics across all three transit modes—CTA, Metra, and Pace—show that RF performs comparably to XGBoost, another advanced ensemble method, but also outperforms single

Table 2 Measures of association between explanatory variables. We use Freeman's Theta (0 to 1) to measure the association between nominal and ordinal variables, and Kendall's Tau (−1 to 1) for pairs of ordinal variables. Upon using Goodman and Kruskal's Lambda (0 to 1) for pairs of nominal variables (Khamis 2008), we found no significant association, so these results have been omitted

Mode	Freeman's Theta			Kendall's Tau		
	Nominal	Ordinal	Association	Ordinal 1	Ordinal 2	Association
CTA	Lapsed ridership	Age	0.0761	Age	Telework	−0.08579361
		Telework	0.331		Vehicles	0.1835821
		Vehicles	0.402		Household size	0.01802522
		Household size	0.163		Income	0.1336649
		Income	0.375	Telework	Vehicles	0.0155204
	Ethnicity	Age	0.192		Household size	0.005127122
		Telework	0.0812		Income	0.1262499
		Vehicles	0.073	Vehicles	Household size	0.4530741
		Household size	0.105		Income	0.2735107
		Income	0.314	Household size	Income	0.1888747
	Wave	Age	0.0119			
		Telework	0.0275			
		Vehicles	0.00182			
		Household size	0.00343			
		Income	0.0339			
Metra	Lapsed ridership	Age	0.175	Age	Telework	−0.1116095
		Telework	0.639		Vehicles	0.1947137
		Vehicles	0.0103		Household size	−0.07413817
		Household size	0.0455		Income	0.07835445
		Income	0.226	Telework	Vehicles	0.004045077
	Ethnicity	Age	0.211		Household size	0.02160146
		Telework	0.0561		Income	0.1010316
		Vehicles	0.0833	Vehicles	Household size	0.4565508
		Household size	0.0914		Income	0.3162459
		Income	0.246	Household size	Income	0.2355178
	Wave	Age	0.0198			
		Telework	0.0111			
		Vehicles	0.00187			
		Household size	0.00281			
		Income	0.032			
Pace	Lapsed ridership	Age	0.0218	Age	Telework	−0.09658438
		Telework	0.378		Vehicles	0.1196992
		Vehicles	0.144		Household size	−0.1063628
		Household size	0.0231		Income	0.09391496
		Income	0.434	Telework	Vehicles	−0.002163149
	Ethnicity	Age	0.148		Household size	−0.000803815
		Telework	0.0591		Income	0.09936922
		Vehicles	0.169	Vehicles	Household size	0.4607164
		Household size	0.148		Income	0.2936109
		Income	0.295	Household size	Income	0.1405129
	Wave	Age	0.0156			
		Telework	0.0556			
		Vehicles	0.0464			
		Household size	0.0608			
		Income	0.00156			

Table 3 Performance of XGBoost, decision tree, logistic regression, and random forest, as measured by AUC

Algorithm	XGBoost	Decision tree	Logistic regression	Random forest
CTA	0.784	0.737	0.770	0.777
Metra	0.839	0.821	0.818	0.839
Pace	0.802	0.685	0.791	0.837

models such as decision tree and logistic regression. As seen in Table 3, random forest achieved an AUC of 0.777 for CTA, closely following XGBoost's 0.784 and surpassing both decision tree (0.737) and logistic regression (0.770). Similarly, for Metra and Pace, random forest consistently shows strong performance with accuracies of 0.839 and 0.837, respectively, either matching or exceeding the performance of other models.

The choice of RF is justified not only by its competitive performance but also by its robustness and interpretability compared to other ensemble techniques, such as XGBoost. While XGBoost is known for its efficiency and effectiveness in handling a wide variety of data types and tasks, it is also more prone to overfitting, especially in scenarios where the data set is not very large or highly dimensional (Magdum et al. 2019). RF, on the other hand, is less susceptible to overfitting due to its mechanism of building multiple decision trees and averaging their results, which naturally controls the complexity of the model (Gu et al. 2021).

Additionally, RF offers greater interpretability than XGBoost without sacrificing performance. XGBoost uses an ensemble technique called boosting, which is complex and difficult to interpret (Dunn et al. 2021). RF is essentially an ensemble of Decision Trees (Gu et al. 2021). While a single decision tree is very interpretable, it often suffers from overfitting (Marcos et al. 2018). RF mitigates this by averaging multiple decision trees, thus reducing the variance without drastically reducing interpretability (Gu et al. 2021). Each tree in the forest considers a random subset of features and samples, which makes RF more robust and less likely to overfit compared to a single Decision Tree (Gu et al. 2021).

Policy decisions require not just predictive accuracy but also a clear understanding of what drives those predictions. Given the importance of performance, generalizability, and interpretability in policy-facing applications, this paper will focus on RF analysis to determine variable importance. In doing so, this paper will identify significant variables in the context of lapsed ridership and ensure that policy interventions are targeted effectively.

Variable Importance

Variable importance describes the extent to which a model relies on a particular feature. In this study, it is measured by permutation importance, which demonstrates how the prediction error of the model increases when the values of one variable are permuted among the data points (Ou et al. 2017). This process randomizes the relationship between the variable and the outcome, thus providing insight into how much the model depends on the variable. The primary aim is to understand the contribution of each feature to the model's predictive accuracy. This metric is intended to compare different variables within a single model and cannot be used to compare across different models. Unlike significance testing, permutation importance testing does not assume any specific form of the relationship between features and the target variable. Thereby, permutation importance enables us to measure the influence of telework and other predictors without knowing the structure of the relationship between our input and transit ridership.

Results

Our results indicate that random forest can accurately distinguish between lapsed and current riders, and that telework is a highly significant variable in predicting lapsed ridership for all three modes. Based on Table 4, we see that RF achieves a reasonably high AUC for each mode. From Fig. 4, we note that the number of days spent teleworking is clearly the most significant variable in predicting lapsed ridership. This is especially pronounced for Metra, the traditionally commuter-centered rail, compared to CTA and Pace.

For predicting CTA ridership, the three most important features are the number of teleworking days, closely followed by the number of vehicles, and in third place annual household income. These are also three of the most correlated variables according to our Theta values in Table 2. Household size, age, and ethnicity showed multicollinearity with both the number of vehicles and annual household income. As such, it makes sense that these three variables are less important in this RF despite being somewhat correlated with lapsed ridership.

Table 4 Random forest performance for CTA, Metra, and Pace as measured by holdout AUC and OOB error rate

Mode	Holdout AUC	OOB error rate
CTA	0.777	16.72%
Metra	0.839	10.58%
Pace	0.837	15.82%

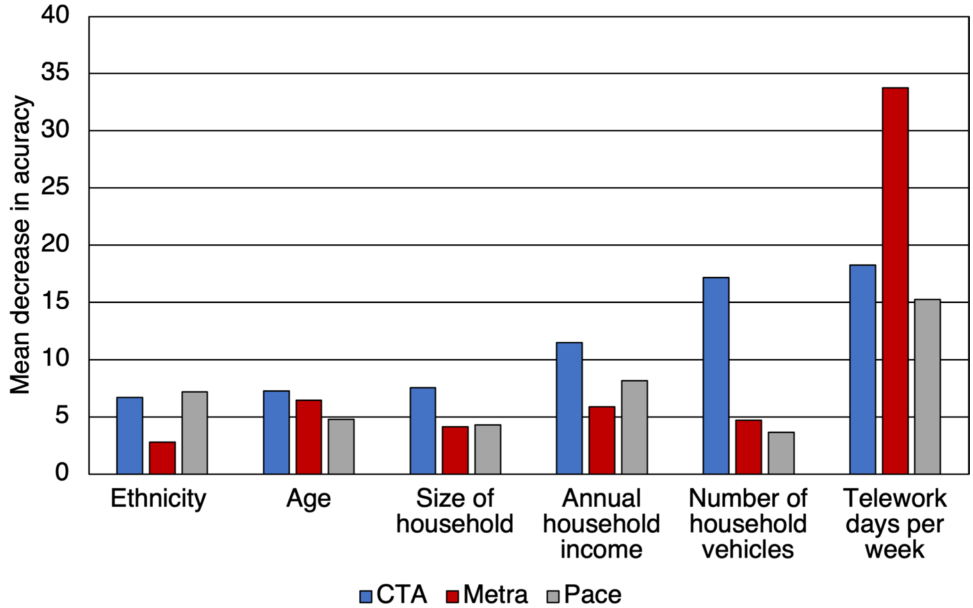


Fig. 4 Variable importance for CTA, Metra, and Pace as measured by mean decrease in OOB accuracy

Based on the association measures discussed previously, the feature importance for Metra ridership is consistent with our expectations. We observe that the number of days spent teleworking exhibits the highest mean decrease in accuracy. After this, we note that age and annual household income, two of the most highly correlated variables, are the second most important features.

The two top features in predicting Pace ridership are the number of teleworking days and household income. As seen in Table 2, this is consistent with our Theta analysis. Ethnicity is the third most important variable, despite having a low Lambda value. This may be because RF exhibits a better ability to select variables by random extraction so that all combinations of ethnicity are controlled for more thoroughly. We also see that the number of household vehicles is the least important variable, even though it showed a high Theta correlation. Again, this is probably due to multicollinearity with income and ethnicity.

The partial dependence plot seen in Fig. 5 shows the probability predicted by RF that a rider will cease to use transit, based on the number of days per week spent teleworking. This enables a scenario analysis where one can study the marginal effects of this variable. For example, a Metra rider who teleworks two days per week is about 60% likely to lapse ridership. The partial dependence of Metra ridership suggests that the likelihood of continuing to use Metra decreases as the number of days spent teleworking increases. Effectively, any teleworking amount greater than 3 days appears to make it very likely that the respondent will lapse. The partial dependence plots of CTA and Pace ridership show that people who telework 1–3 days per week are very likely to lapse. For workers with more than 3 days per

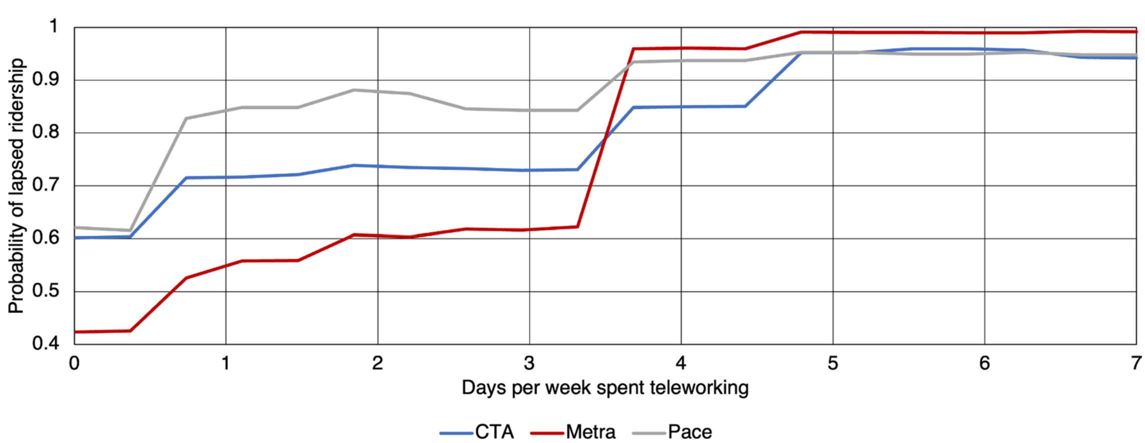


Fig. 5 Partial dependence plot for CTA, Metra, and Pace based on number of teleworking days

week of remote work, the lapsed ridership is almost certain. Compared to Metra, Pace and CTA ridership appear to be significantly less sensitive to the number of days spent teleworking. In essence, these respondents are likely to lapse with any amount of telework, whereas there is more nuance in the relationship between telework and Metra ridership.

Discussion

Interpretation of Variable Importance Results

These results indicate that RF can accurately distinguish between lapsed and current riders for each mode. Among the variables tested, the number of days spent teleworking per week was consistently the most important variable in predicting ridership status, as measured by the mean decrease in the Gini coefficient and permutation importance. While ridership was correlated with other variables, such as the number of household vehicles and income bracket, these factors exhibited multicollinearity. This is reflected in the importance rankings of the RF, where these variables didn't have as much predictive power as the number of days spent teleworking. It is also important to note that for each mode, there was a steep dropoff in both importance measures beyond the telework variable, particularly with the permutation importance. We can thus infer that telework status substantially impacts a respondent's decision to use CTA, Metra, or Pace services. This suggests that the future evolution of post-pandemic work location policies will be an essential factor to control for in transit demand studies.

Transit Ridership and the Future of Telework

When these data were collected at the beginning of the COVID-19 pandemic, numerous industries within the United States and abroad were forced to transition their workforce to full-time telework or a hybrid model (Katsabian 2022). This represented a paradigm shift in the way many people live and work. Telework in the U.S. may continue playing a significant role with almost half of workers able to work potentially remotely (Rothstein and Aughinbaugh 2022). According to a May 2020 Gallup poll, a combined 70% of remote-capable people teleworked "always" or "sometimes," and half of these remote workers wanted to continue working from home even after pandemic concerns were alleviated (Hickman and Saad 2020). This sentiment has not changed significantly since then. According to a more recent Gallup poll released in February of 2022, 81% of remote-capable workers worked remotely at least sometimes ($>10\%$ remote), 77% of them expected to continue doing so, and 91% of these workers desired a fully remote or hybrid job (Wigert 2022).

Given the results of this study, the future of telework has important ramifications for transit ridership. For CTA and Pace services, the partial dependence plots show that lapsed ridership was almost a deterministic outcome with even a minimal amount of remote work during the COVID-19 pandemic. With Metra, there is more sensitivity to the number of days spent teleworking, but people who work from 2 to 3 days per week or more are still very likely to cut transit use. As of January 2024, many of the coronavirus-related concerns about using transit had been alleviated, but nationwide transit ridership has not recovered. Given that many remote-capable workers in the United States are still teleworking and expect to continue doing so, it may be difficult for transit ridership to recover completely.

This also has implications for transportation equity, as low-income and minority groups are particularly reliant on public transit and are likely to be disproportionately impacted by service cuts from transit agencies (Wilbur et al. 2023). The ability to telework is highly dependent on the type of industry, and according to the Bureau of Labor Statistics, white-collar industries such as financial services and information technology are far more likely to be compatible with telework (Dalton and Groen 2022). This study also revealed that these industries experienced a smaller reduction in employment (Dalton and Groen 2022). On the other hand, jobs that typically pay low wages were not likely to be compatible with telework and faced a far greater risk of unemployment (Dalton and Groen 2022). This includes service jobs as well as those in leisure and hospitality. As a result, remote work is less accessible for historically marginalized groups, such as low-income and minority workers (Dalton and Groen 2022). Furthermore, given that the budget of each RTA mode relies heavily on fare and pass revenue (RTA 2021), a reduction in ridership has consequences for the financial viability of these services.

Policy Implications

The results of this study indicate that agencies and regulators in the Chicago region may see increased ridership by promoting non-commute transit use. Strategies will differ according to the different ridership pattern shifts observed among operators. Metra rail services experienced the greatest decline in work commuting, and in travel overall. To this end, Metra has targeted increasing service levels during off-peak periods (Victory 2022), shifting from a "commute model" to a "regional rail model" with all-day transportation (Metra 2024b). Additionally, ridership levels and equity concerns may be simultaneously addressed by decreasing fares and increasing service for essential workers and other demographic groups that are typically unable to telework. In practice, this has been proven to improve ridership levels, as shown by the Fair Transit South Cook program

(Fair Transit South Cook 2024). Beginning in early 2021, this pilot reduced fares by 50% on the Electric (ME) and Rock Island (RI) Metra lines, and increased service levels in these as well as Pace Route 352 Halsted (Fair Transit South Cook 2024). These lines travel through areas in the south side of Chicago and Cook County and are typically relied on by communities that were disproportionately impacted by the pandemic and have historically faced high poverty and unemployment rates (CDPH 2024; CMAP 2023). The results of this program show that the decreased fares resulted in substantially increased Metra ridership, as ME and RI levels recovered more rapidly than the other Metra lines (Fair Transit South Cook 2023).

The significance of telework in this context demonstrates the need for consistent and thorough data collection regarding telework in the future, including emerging forms of telework (e.g., center commuting, third-place working, or part-day telework, see Stiles and Smart 2021; Okashita et al. 2023). However, given that teleworking trends are external to RTA, they cannot be directly modified by this agency or any single governmental body. This raises a question for future research: when the most important variable is not easily controlled by the regulatory entity in question, how can machine learning models be used to identify feasible policy improvements? In this context, we determined that transit agencies can focus on non-commute trips and people who don't work remotely to identify complementary but indirect solutions, since the telework variable cannot easily be controlled.

Conclusion

This study employed random forest techniques to explore the relationship between telework and public transit ridership in the Chicago area, providing insights into the post-pandemic transit landscape. Our investigation reveals that telework, which was accelerated by the pandemic, is not a transient phenomenon but a persistent determinant of transit ridership, particularly for commuter-centric services like Metra. As telework continues to shape urban transit landscapes, future studies ought to collect more detailed data regarding telework behaviors and trends to study its implications for urban mobility and transit planning.

The utilization of machine learning enables a data-driven exploration of emerging and poorly understood phenomena. The machine learning approach, particularly random forest, was effective in exploring the complex dynamics of transit ridership and teleworking behavior, showcasing its usefulness in studying transit-service behavioral differences. This offers valuable methodological insight for future transportation research. Specifically, our findings suggest that agencies and researchers need to think more critically about the role of respondent employment conditions and work-location choices in designing future data-collection and analytical approaches.

Several limitations are recognized in this study. First, data-sampling was biased toward the majority class of lapsed riders. Reliance on imbalanced data may reduce generalizability; therefore, future studies may need to verify and address the imbalance in the sampling via corrective measures like resampling. Second, the specificity of the pandemic context might limit the generalizability of our findings over time. Future studies ought to keep monitoring the return of transit ridership in tandem with labor-market trends to examine how the relationship evolves. Third, this study emphasizes mode-specific analysis within a single metropolitan area. While the work contributes new insight into mode-specific telework connections, the transferability to other urban contexts is premature. It remains unclear how post-pandemic transit usage paths vary among countries and cities. We recommend future comparative analysis to examine how the recovery of transit, and the variation across modes, is tied to local labor conditions (e.g., labor market composition, telework policies), transit policies (e.g., funding for transit, fare policy), and socio-economic trends (e.g., consumer sentiment, urban revitalization). Fourth, machine learning is a powerful approach for studying evolving transit use patterns where algorithms can automatically identify patterns that are not well understood. Yet, these tools require careful interpretation and validation to ensure applicability for policymaking in the specific context of transit funding, regulations, and rider circumstances.

In conclusion, this study not only contributes to our understanding of the telework-transit nexus but also demonstrates the value of machine learning, particularly Random Forest, in advancing transportation research. In the wake of large-scale social and physical disasters like the coronavirus pandemic, regulatory agencies must accommodate consequential and poorly understood external factors that are beyond their control. As such, leveraging sophisticated analytical tools that can handle this uncertainty, such as machine learning, will be crucial for crafting resilient, equitable, and forward-looking transit policies.

Author Contributions The authors confirm their contribution to the paper as follows: study conception and design: A. Stathopoulos; data processing: D. Edward; coding and modeling: D. Edward; analysis and interpretation of results: D. Edward, J. Soria, A. Stathopoulos; initial manuscript preparation: D. Edward; Manuscript revision: D. Edward, J. Soria, A. Stathopoulos; Supervision and funding acquisition: A. Stathopoulos. All authors reviewed the results and approved the final version of the manuscript.

Funding This research was supported by the following sources: Soria received funding from the Northwestern University Transportation Center Dissertation Year Fellowship. Partial funding is provided from the U.S. National Science Foundation (NSF) Career Grant no. 1847537 (Stathopoulos). Partial funding is provided by the U.S. Department of Transportation Tier 1 University Transportation Center on Telemobility (Stathopoulos and Edward).

Availability of Data and Materials The survey dataset that supports the findings of this study are not publicly available due to ownership and protection of human subjects by the Regional Transportation Authority.

Declarations

Conflict of interest We have no conflicts of interest to declare.

Ethics approval and consent to participate Respondents to the survey in this study consented to receiving and optionally participating in data-collection efforts by opting into email-based customer contact databases as well as online opinion panels with an invitation from the Regional Transportation Authority (RTA). Before the data collection portion of the survey, respondents were notified of the purpose of the study, the possibility of receiving randomly assigned lottery compensation, and that their answers would remain anonymous, be reported only in aggregate, and only to study lapsed transit ridership. Consent to participate was verified by asking respondents to confirm their participation in the study by explicitly clicking a button on the website to access the rest of the survey.

References

Abdullah M, Dias C, Muley D, Shahin M (2020) Exploring the impacts of COVID-19 on travel behavior and mode preferences. Transp Res Interdiscipl Perspect 8:100255

Amarasinghe K, Rodolfa KT, Lamba H, Ghani R (2023) Explainable machine learning for public policy: use cases, gaps, and research directions. Data Policy 5:e5

APTA (2024) APTA ridership trends. https://transitapp.com/apta (searching on agencies of CTA and Metra) consulted March 11th 2024

Barbour N, Menon N, Mannering F (2021) A statistical assessment of work-from-home participation during different stages of the COVID-19 pandemic. Transp Res Interdiscipl Perspect 11:100441

Beck MJ, Hensher DA (2020) Insights into the impact of COVID-19 on household travel and activities in Australia—the early days under restrictions. Transp Policy 96:76–93

Beck MJ, Hensher DA, Wei E (2020) Slowly coming out of COVID-19 restrictions in Australia: Implications for working from home and commuting trips by car and public transport. J Transp Geogr 88:102846. https://doi.org/10.1016/j.jtrangeo.2020.102846

Bell A, Solano-Kamaiko I, Nov O, Stoyanovich J (2022) It's Just not that simple: an empirical study of the accuracy-explainability trade-off in machine learning for public policy. In: Proceedings of the 2022 ACM conference on fairness, accountability, and transparency. Association for Computing Machinery, New York, NY, USA, pp 248–266

BLS (2024) Bureau of labor statistics, economic news release. American Time Use Survey. https://www.bls.gov/news.release/atus.t06.htm. Accessed 17 Apr 2024

Breiman L (2001) Random Forests. Mach Learn 45:5–32. https://doi.org/10.1023/A:1010933404324

Brodeur ZP, Herman JD, Steinschneider S (2020) Bootstrap aggregation and cross-validation methods to reduce overfitting in reservoir control policy search. Water Resour Res 56:e2020WR027184. https://doi.org/10.1029/2020WR027184

Bzdok D, Altman N, Krzywinski M (2018) Statistics versus machine learning. Nat Methods 15:233–234. https://doi.org/10.1038/nmeth.4642

Caine P (2021) Federal Funds Seen as Key to Reviving Public Transit. WTTW News. https://news.wttw.com/2021/04/06/federal-funds-seen-key-reviving-public-transit

CMAP: Chicago Metropolitan Agency for Planning (2023). Community Data Snapshot: Southside of Chicago. https://www.cmap.illinois.gov/documents/10180/126764/South+Chicago.pdf. Accessed 6 May 2024

CDPH: Chicago Department of Public Health (2024) "COVID-19 at a glance | COVID 19 segmented by Race and Ethnicity. https://www.chicago.gov/city/en/sites/covid-19/home/latest-data.html. Accessed 6 May 2024

CTA (2024a) CTA facts at a glance. In: CTA. https://www.transitchicago.com/facts/. Accessed 17 Apr 2024

CTA (2024b) How-to guide: making regional Chicago transit connections. In: CTA. https://www.transitchicago.com/howto/regional-connections/. Accessed 17 Apr 2024

Dalton M, Groen J (2022) Telework during the COVID-19 pandemic: estimates using the 2021 Business Response Survey : Monthly Labor Review: U.S. Bureau of Labor Statistics. https://www.bls.gov/opub/mlr/2022/article/telework-during-the-covid-19-pandemic.htm. Accessed 26 Jan 2024

DeWeese J, Hawa L, Demyk H et al (2020) A tale of 40 cities: a preliminary analysis of equity impacts of COVID-19 service adjustments across North America. Findings. https://doi.org/10.32866/001c.13395

Dunn J, Mingardi L, Zhuo YD (2021) Comparing interpretability and explainability for feature selection. arXiv preprint http://arxiv.org/abs/2105.05328

Efthymiou A, Barmpounakis EN, Efthymiou D, Vlahogianni EI (2019) Transportation mode detection from low-power smartphone sensors using tree-based ensembles. J Big Data Anal Transp 1:57–69

El-Geneidy A, Levinson D, Diab E et al (2016) The cost of equity: assessing transit accessibility and social disparity using total travel cost. Transp Res Part A Policy Pract 91:302–316. https://doi.org/10.1016/j.tra.2016.07.003

Fair Transit South Cook (2023) Pilot Program Second-Year Report

Fair Transit South Cook (2024) Fair Transit South Cook. https://www.cookcountyil.gov/service/fair-transit-south-cook. Accessed 26 Jan 2024

Farmer L (2022) Remote work during COVID-19: what can it tell us about the future of work? In: Rockefeller Institute of Government. https://rockinst.org/issue-area/remote-work-during-covid-19-what-can-it-tell-us-about-the-future-of-work/. Accessed 17 Apr 2024

Fatmi MR, Thirkell C, Hossain MS (2021) COVID-19 and travel: how our out-of-home travel activity, in-home activity, and long-distance travel have changed. Transp Res Interdiscipl Perspect 10:100350. https://doi.org/10.1016/j.trip.2021.100350

Freishtat S (2021) Waits increase for public transit [Corrected 12/11/2021]: study: CTA has been running fewer buses, trains during pandemic. Chicago Tribune

Griffin GP, Sener IN (2016) Public transit equity analysis at metropolitan and local scales: a focus on nine large cities in the US. J Public Trans 19:126–143. https://doi.org/10.5038/2375-0901.19.4.8

Gu Y, Si B, Liu B (2021) A novel hierarchical model in ensemble environment for road detection application. Remote Sens 13:1213. https://doi.org/10.3390/rs13061213

He Q, Rowangould D, Karner A et al (2022) Covid-19 pandemic impacts on essential transit riders: findings from a US survey. Transp Res Part D Transp Environ 105:103217. https://doi.org/10.1016/j.trd.2022.103217

Hickman A, Saad L (2020) reviewing remote work in the U.S. under COVID-19. In: Gallup.com. https://news.gallup.com/poll/311375/reviewing-remote-work-covid.aspx. Accessed 26 Jan 2024

Hu S, Chen P (2021) Who left riding transit? Examining socioeconomic disparities in the impact of COVID-19 on ridership. Transp Res Part D Transp Environ 90:102654

Illinois Economic Policy Institute (2021) COVID-19 and transportation funding in Illinois one year later. Tyler, M. Report. https://www.railwayage.com/wp-content/uploads/2021/03/covid-transportation-one-year-later-final.pdf. Accessed 18 Apr 2024

Karlaftis MG, Vlahogianni EI (2011) Statistical methods versus neural networks in transportation research: Differences, similarities and some insights. Transp Res Part C Emerg Technol 19(3):387–399

Katsabian T (2022) The telework virus: how COVID-19 has affected telework and exposed its implications for privacy. Berkeley Journal of Labor and Employment Law. https://doi.org/10.2139/ssrn.4092787

Khamis H (2008) Measures of association: how to choose? J Diagn Med Sonogr 3:155–162. https://doi.org/10.1177/8756479308317006

Lei Z, Ukkusuri SV (2022) Understanding the recovery of on-demand mobility services in the COVID-19 era. J Big Data Anal Transp 4(1):1–21

Li Z, Wei Z, Zhang Y, Kong X, Ma C (2023) Applying an interpretable machine learning framework to study mobility inequity in the recovery phase of COVID-19 pandemic. Travel Behav Soc 33:100621

Liu L, Miller HJ, Scheff J (2020) The impacts of COVID-19 pandemic on public transit demand in the United States. PLoS ONE 15:e0242476. https://doi.org/10.1371/journal.pone.0242476

Lunardon N, Menardi G, Torelli N (2014) ROSE: a package for binary imbalanced learning. R J 6:79. https://doi.org/10.32614/RJ-2014-008

Mack EA, Agrawal S, Wang S (2021) The impacts of the COVID-19 pandemic on transportation employment: a comparative analysis. Transp Res Interdiscipl Perspect 12:100470. https://doi.org/10.1016/j.trip.2021.100470

Magdum J, Ghorse R, Chaku C et al (2019) A computational evaluation of distributed machine learning algorithms. In: 2019 IEEE 5th international conference for convergence in technology (I2CT), pp 1–6

Mannering F, Bhat CR, Shankar V, Abdel-Aty M (2020) Big data, traditional data and the tradeoffs between prediction and causality in highway-safety analysis. Anal Methods Accid Res 25:100113. https://doi.org/10.1016/j.amar.2020.100113

Marcos R, García-Cantú O, Herranz R (2018) A machine learning approach to air traffic route choice modelling. https://arxiv.org/abs/1802.06588

Martin L, Hauret L, Fuhrer C (2022) Digitally transformed home office impacts on job satisfaction, job stress and job productivity. COVID-19 findings. PLoS ONE 17:0265131

Matson G, McElroy S, Circella G, Lee Y (2021) Telecommuting rates during the pandemic differ by job type, income, and gender. National Center for Sustainable Transportation. https://doi.org/10.7922/G2445JT6

Metra (2024a) Our history, Metra. https://metra.com/our-history. Accessed 17 Apr 2024

Metra (2024b) Metra 2023–2027 strategic plan. https://metra.com/sites/default/files/MyMetraOurFutureFinal/mobile/index.html. Accessed 24 Apr 2024

Mohammadi M, Rahimi E, Davatgari A, Javadinasr M, Mohammadian A, Bhagat-Conway MW, Salon D, Derrible S, Pendyala RM, Khoeini S (2023) Examining the persistence of telecommuting after the COVID-19 pandemic. Transp Lett 15(6):608–621

Mokhtarian PL, Salomon I (1997) Modeling the desire to telecommute: the importance of attitudinal factors in behavioral models. Transp Res Part A Policy Pract 31:35–50

Mouratidis K, Papagiannakis A (2021) COVID-19, internet, and mobility: the rise of telework, telehealth, e-learning, and e-shopping. Sustain Cities Soc 74:103182. https://doi.org/10.1016/j.scs.2021.103182

Mourtakos V, Mantouka EG, Fafoutellis P, Vlahogianni EI, Kepaptsoglou K (2024) Reconstructing mobility from smartphone data: empirical evidence of the effects of COVID-19 pandemic crisis on working and leisure. Transp Policy 146:241–254

Nardi PM (2018) Doing survey research: a guide to quantitative methods, 4th edn. Routledge, New York

Nayak S, Pandit D (2021) Potential of telecommuting for different employees in the Indian context beyond COVID-19 lockdown. Transp Policy 111:98–110

Noursalehi P, Koutsopoulos HN, Zhao J (2020) Machine-learning-augmented analysis of textual data: application in transit disruption management. IEEE Open J Intell Transp Syst 1:227–236

Okashita A, Arzate H, Kim JH (2023) What does the prevalence of telecommuting mean for urban planning?. ITS UC Davis. https://escholarship.org/uc/item/9gz1j5qv

Olde Kalter M-J, Geurs KT, Wismans L (2021) Post COVID-19 teleworking and car use intentions. Evidence from large scale GPS-tracking and survey data in the Netherlands. Transp Res Interdiscipl Perspect 12:100498. https://doi.org/10.1016/j.trip.2021.100498

Orlenko A, Moore JH (2021) A comparison of methods for interpreting random forest models of genetic association in the presence of non-additive interactions. BioData Min 14:9. https://doi.org/10.1186/s13040-021-00243-0

Oshiro TM, Perez PS, Baranauskas JA (2012) How many trees in a Random Forest? In: Perner P (ed) Machine learning and data mining in pattern recognition. Springer, Berlin, pp 154–168

Ou J, Xia J, Wu YJ, Rao W (2017) Short-term traffic flow forecasting for urban roads using data-driven feature selection strategy and bias-corrected random forests. Transp Res Rec 2645(1):157–167

Pace (2022) Pace 2022 customer satisfaction survey. https://www.pacebus.com/sites/default/files/2023-02/Pace%202022%20CSS%20Final%20Report.pdf. Accessed 20 Apr 2024

Pace (2024) ADA Paratransit Service | Pace Suburban Bus. https://www.pacebus.com/ada. Accessed 16 July 2023

Pace (n.d.) Pace Suburban Bus. https://www.pacebus.com/about. Accessed 17 Apr 2024

Parker K, Horowitz JM, Minkin R (2020) How the coronavirus outbreak has—and hasn't—changed the way Americans Work. In: Pew Research Center's Social & Demographic Trends Project. https://www.pewresearch.org/social-trends/2020/12/09/how-the-coronavirus-outbreak-has-and-hasnt-changed-the-way-americans-work/. Accessed 26 Jan 2024

Parker K, Horowitz JM, Minkin R (2022) COVID-19 pandemic continues to reshape work in America. In: Pew Research Center's Social & Demographic Trends Project (blog), February 16

Plaut PO (2005) Non-motorized commuting in the US. Transp Res Part D Transp Environ 10:347–356

Polzin S, Choi T, United States. Department of transportation. Office of the Assistant Secretary for Research and Technology (2021)

COVID-19's effects on the future of transportation. https://rosap.ntl.bts.gov/view/dot/54292
Pouri YD, Bhat CR (2003) On modeling choice and frequency of home-based telecommuting. Transp Res Rec 1858(1):55–60
Probst P, Boulesteix A-L (2018) To tune or not to tune the number of trees in Random Forest. J Mach Learn Res 18:1–18
Probst P, Wright MN, Boulesteix A-L (2019) Hyperparameters and tuning strategies for random forest. Wires Data Min Knowl Discov 9:e1301. https://doi.org/10.1002/widm.1301
Rahimi E, Shabanpour R, Shamshiripour A, Kouros Mohammadian A (2021) Perceived risk of using shared mobility services during the COVID-19 pandemic. Transp Res Part F Traffic Psychol Behav 81:271–281. https://doi.org/10.1016/j.trf.2021.06.012
Rahman K, Arif MZU (2020) Working from home during the COVID-19 pandemic: satisfaction, challenges and productivity of employees. Int J Trade Commerce-IIARTC 9:282–294. https://doi.org/10.46333/ijtc/9/2/3
Reuschke D, Ekinsmyth C (2021) New spatialities of work in the City. Urban Stud 58:2177–2187. https://doi.org/10.1177/00420980211009174
Rodolfa KT, Lamba H, Ghani R (2021) Empirical observation of negligible fairness—accuracy trade-offs in machine learning for public policy. Nat Mach Intell 3:896–904. https://doi.org/10.1038/s42256-021-00396-x
RTA 2022 Customer Satisfaction Survey Report.https://www.rtachicago.org/uploads/files/general/Region/RTA-CSS-2022-Final-Report.pdf. Accessed 18 Apr 2024
Rothengatter W, Zhang J, Hayashi Y, Nosach A, Wang K, Oum TH (2021) Pandemic waves and the timeafter Covid-19 – Consequences for the transport sector. Transp. Policy 110:225–237
Rothstein D, Aughinbaugh A (2022) How did employment change during the COVID-19 pandemic? Evidence from a new BLS survey supplement: beyond the numbers: U.S. Bureau of Labor Statistics. https://www.bls.gov/opub/btn/volume-11/how-did-employment-change-during-the-covid-19-pandemic.htm. Accessed 26 Jan 2024
RTA (2017a) Customer satisfaction study. Resource Systems Group Report https://www.rtachicago.org/uploads/files/general/Drupal-Old/documents/plansandprograms/Regional-Transportation-Authority_-2016-Customer-Satisfaction-Study.pdf
RTA (2017b) Regional transit strategic plan 2018–2023 beginning the discussion May 2017. www.rtachicago.org/uploads/files/general/Drupal-Old/documents/strategicprograms/strategicplan/IIT_2018-23_Final/InvestInTransit_BeginDiscuss_18-23.pdf. Accessed 18 April 2024
RTA (2021) RTA COVID-19 lapsed rider survey. www.rtachicago.org/uploads/files/general/Drupal-Old/2021-04/RTA%20COVID-19%20Lapsed%20Rider%20Survey%20-%20Final%20Report.pdf. Accessed 10 June 2023
RTA (2024) Systemwide ridership (annually, in millions) dashboard, web resource. https://rtagis.maps.arcgis.com/apps/dashboards/43366b7d17e04becb2e4b10f5e6f2e50. Accessed 22 April 2024
Rudin C, Radin J (2019) Why are we using black box models in AI when we don't need to? A lesson from an explainable AI competition. Harv Data Sci Rev. https://doi.org/10.1162/99608f92.5a8a3a3d
Sadeghian P, Zhao X, Golshan A, Håkansson J (2022) A stepwise methodology for transport mode detection in GPS tracking data. Travel Behav Soc 26:159–167
Sekadakis M, Katrakazas C, Michelaraki E, Ziakopoulos A, Yannis G (2023) COVID-19 and driving behavior: which were the most crucial influencing factors? Data Sci Transp 5(3):16
Soria J, Edward D, Stathopoulos A (2023) Requiem for transit ridership? An examination of who abandoned, who will return, and who will ride more with mobility as a service. Transp Policy 134:139–154. https://doi.org/10.1016/j.tranpol.2023.02.016
Stiles J, Smart MJ (2021) Working at home and elsewhere: daily work location, telework, and travel among United States knowledge workers. Transportation 48(5):2461–2491
Sun W, Kobayashi H, Nakao S, Schmöcker JD (2023) On the relationship between crowdsourced sentiments and mobility trends during COVID-19: a case study of Kyoto. Data Sci Transp 5(3):17
Sun Y, Huang SK, Zhao X (2024) Predicting hurricane evacuation decisions with interpretable machine learning methods. Int J Disaster Risk Sci 15(1):134–148
Tahlyan D, Said M, Mahmassani H et al (2022) For whom did telework not work during the pandemic? Understanding the factors impacting telework satisfaction in the US using a multiple indicator multiple cause (MIMIC. Model Transp Res Part A Policy Pract 155:387–402
Tahlyan D, Mahmassani H, Stathopoulos A et al (2024) In-person, hybrid or remote? Employers' perspectives on the future of work post-pandemic. https://arxiv.org/abs/2402.18459
Tomaschek F, Hendrix P, Baayen RH (2018) Strategies for addressing collinearity in multivariate linguistic data. J Phon 71:249–267
U.S. Census Bureau (2023) U.S. Census Bureau QuickFacts: Chicago city, Illinois. https://www.census.gov/quickfacts/fact/table/chicagocityillinois/PST045222. Accessed 16 Jul 2023
Van Cranenburgh S, Wang S, Vij A, Pereira F, Walker J (2022) Choice modelling in the age of machine learning-discussion paper. J Choice Model 42:100340
Venkataramani S (2021) Returning employees to an office? Consider the talent risks. Gartner Insight Human Resources. https://www.gartner.com/smarterwithgartner/returning-employees-to-an-office-consider-the-talent-risks
Vickerman R (2021) Will Covid-19 put the public back in public transport? A UK perspective. Transp Policy 103:95–102. https://doi.org/10.1016/j.tranpol.2021.01.005
Victory L (2022) Metra aims to restore full pre-pandemic service by end of next year—CBS Chicago. https://www.cbsnews.com/chicago/news/metra-aims-to-restore-full-pre-pandemic-service-by-end-of-next-year/. Accessed 26 Jan 2024
Wang S, Mo B, Zhao J (2020) Deep neural networks for choice analysis: architecture design with alternative-specific utility functions. Transp Res Part C Emerg Technol 112:234–251. https://doi.org/10.1016/j.trc.2020.01.012
Wang S, Mo B, Hess S, Zhao J (2021) Comparing hundreds of machine learning classifiers and discrete choice models in predicting travel behavior: an empirical benchmark. arXiv preprint arXiv:2102.01130.
Wigert B (2022) The future of hybrid work: 5 key questions answered with data. In: Gallup.com. https://www.gallup.com/workplace/390632/future-hybrid-work-key-questions-answered-data.aspx. Accessed 26 Jan 2024
Wilbur M, Ayman A, Sivagnanam A et al (2023) Impact of COVID-19 on public transit accessibility and ridership. Transp Res Rec 2677:531–546. https://doi.org/10.1177/03611981231160531
Yan M, Shen Y (2022) Traffic accident severity prediction based on Random Forest. Sustainability 14:1729. https://doi.org/10.3390/su14031729
Zhou Y, Liu XC, Grubesic T (2021) Unravel the impact of COVID-19 on the spatio-temporal mobility patterns of microtransit. J Transp Geogr 97:103226
Ziedan A, Brakewood C, Watkins K (2023a) Will transit recover? A retrospective study of nationwide ridership in the United States during the COVID-19 pandemic. J Public Trans 25:100046. https://doi.org/10.1016/j.jpubtr.2023.100046

Ziedan A, Lima L, Brakewood C (2023b) A multiple mediation analysis to untangle the impacts of COVID-19 on nationwide bus ridership in the United States. Transp Res Part A Policy Pract 173:103718. https://doi.org/10.1016/j.tra.2023.103718