Transportation Research Part A 186 (2024) 104072

Contents lists available at ScienceDirect

Transportation Research Part A

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

Examining the treatment effect of teleworking on vehicle-miles driven: Applying an ordered probit selection model and incorporating the role of travel stress

Xinyi Wang[a,*], Patricia L. Mokhtarian[b]

[a] *Department of Urban Studies and Planning, Massachusetts Institute of Technology, 105 Massachusetts Avenue, Cambridge, MA 02139, United States*

[b] *School of Civil and Environmental Engineering, Georgia Institute of Technology, 790 Atlantic Drive, Atlanta, GA 30332, United States*

A R T I C L E I N F O

Keywords:
Telework
Vehicle-miles driven (VMD)
Treatment effect
Motives
Travel stress
Endogenous switching regression
Ordered probit selection model
Log-transform

A B S T R A C T

Teleworking gained considerable popularity during the pandemic, and understanding its impact on travel behavior is of critical interest for post-pandemic transportation planning given its high relevance to travel demand and related issues. We utilize ordered probit endogenous switching regression models to analyze 2021 data from two metropolitan regions, Dallas-Ft. Worth, TX and Washington, DC, consisting of a total of 1,584 observations. We identify factors that impact the adoption and frequency of teleworking (TWing), as well as the weekly vehicle-miles driven (VMD), while accounting for self-selection biases. We define three TW categories: non-TWing (NTW), non-usual TWing (NUTW, fewer than 3 days a week), and usual TWing (UTW, 3+ days a week). We further separate workers based on teleworking-related motives (specifically, travel-stressed or not) to compare results when the outcome variable (VMD) is likely congruent with the teleworking motivation versus when it is not. Based on the model results, we quantify and compare the impacts (i.e. the "treatment effects") of teleworking on VMD. We find that the *treatment effects on the treated* – i.e. the effects on NUTWers of adopting NUTWing, and the effects on UTWers of adopting UTWing – constitute significant reductions in VMD, on average, for both travel-stressed and non-travel-stressed TWers. For travel-stressed NTWers, we find that adopting NUTW teleworking at a low frequency level (i.e., less than 3 times a week) would not significantly reduce VMD, while for non-travel-stressed UTWers, adopting NUTW would significantly *increase* VMD, on average. However, adopting UTWing or increasing teleworking frequency from non-usual to usual always leads to a reduction in VMD on average, whether travel-stressed or non-travel-stressed. The ordered probit endogenous switching regression methodology used here, including visualizations of factual and counterfactual effects and back-transformation of the log-transformed outcome variable, can also be applied to numerous other research topics.

1. Introduction

Teleworking was widely adopted during the COVID-19 pandemic. As the pandemic has receded, many "pandemic teleworkers" have chosen to keep teleworking even though this work arrangement is often no longer required (or even desired) by their employers

* Corresponding author.
E-mail addresses: xinyi174@mit.edu (X. Wang), patmokh@gatech.edu (P.L. Mokhtarian).

https://doi.org/10.1016/j.tra.2024.104072
Received 16 May 2023; Received in revised form 31 January 2024; Accepted 8 April 2024
Available online 8 July 2024
0965-8564/

and/or the risk of serious illness from the coronavirus has been dramatically reduced. In a companion paper (Wang et al., 2024), we discussed different lifestyle motivations that likely relate to workers' *reasons* for teleworking. In this paper, we are interested in the *impact* of teleworking on their travel behavior, specifically, their weekly vehicle-miles driven (VMD).

Overall, the impact of teleworking on how much people travel has been a critical interest of many travel behavior studies, but the results have been debatable. From the demand management perspective, we hope teleworking can reduce VMD and thus lead to reduced traffic congestion, less fuel consumption, and a better environment. This hope may come true if teleworking only eliminates commute trips, ceteris paribus. However, the reality is much more complicated. For example, teleworking might induce mode shifts from public transit to car driving, generate new non-work trips or otherwise modify non-work trip patterns, and even lead to residential relocation farther away from the workplace – all of which could actually *increase* travel in general, or personal vehicle travel in particular. Teleworking could also *modify* travel behavior, such as through changes to departure time or route (as well as mode).

In the literature, we find evidence that teleworking may have substitution (reduction), complementarity (generation), or mixed (modification, neutrality, or hybrid) effects on travel demand (Table 1). From a substitution perspective, Choo et al. (2005) concluded that teleworking by 12% of the workforce in 1998 reduced total annual vehicle-miles traveled (VMT) in the U.S. by around 0.8%. A more recent study by Elldér (2020) found that full-day teleworkers have significantly lower *personal*-kilometers traveled (PKT) than those who do not telework. Part-day teleworking, on the other hand, is associated with significantly higher PKT. The mixed results indicate a heterogeneous travel demand among teleworkers. However, the Elldér study (2020) did not find a significant impact of teleworking on *vehicle*-kilometers traveled (VKT) for either full-day or part-day teleworkers.

Overall in the literature of the past decade, we find more evidence showing that teleworking has a complementarity effect on travel than evidence supporting other effects (Table 1). For example, Zhu (2012) found that teleworking is positively related to trip distance, duration, and frequency for both work and non-work trips. Chakrabarti (2018) found that teleworkers tend to have higher VMD than non-teleworkers, where occasional teleworkers are even more likely to travel more than frequent teleworkers are. Su et al. (2021) found that teleworkers have more complex schedules and visit more locations than commuters. Based on respondents' expectations for their post-pandemic commute travel, Currie et al. (2021) indicated that the commute trip reduction due to teleworking may not offset the mode shift from public transit to car driving, which will lead to a net increase in car use.

Section 2 mentions a number of possible reasons for the mixed evidence. The present study focuses on two in particular: self-selection, and heterogeneity of teleworking motives. Self-selection bias may contribute to the inconsistent findings if the studies directly compare *observed* travel patterns of teleworkers and non-teleworkers. Specifically, some *unobserved* traits of people who choose to telework may relate to how much they travel. In other words, teleworkers may differ from "observationally equivalent" non-teleworkers (in ways that caused them to travel more) even before starting to telework. Addressing self-selection biases is paramount for obtaining accurate results regarding the travel impacts of teleworking itself, as well as the impacts of prospective policy interventions.

Heterogeneity of teleworking motives may also contribute to the inconsistent findings in the literature regarding the impact of teleworking on VMD. As we have found in the companion paper (Wang et al., 2024), heterogeneous telework-related motives (such as travel-dominant versus family-dominant) reflect different benefit expectations once adopting teleworking. For teleworkers with different motives, teleworking may induce different (or even reverse) impacts on travel patterns or lifestyles. Mixing all teleworkers together and examining the net impact of teleworking may shrink or mask such impacts. Separating teleworkers based on their motives may help us identify clearer relationships between teleworking adoption and its outcomes (e.g., VMD, work-life-balance, work productivity). Unraveling the heterogeneity accordingly improves the ability of planners and policymakers to predict the travel-related impacts of teleworking, and potentially to tailor more effective interventions.

From the foregoing discussion, it is clear that understanding the post-pandemic teleworking impact on travel patterns is complicated, but this is a critical question to answer as it will directly influence traffic volumes (and thus congestion) and energy consumption. The objective of this study is to quantify and compare the impact of teleworking on VMD for different types of teleworkers (i.e., by teleworking-related motives), accounting for self-selection biases. The analysis uses a dataset that was collected in Spring 2021, which captures both new (pandemic-prompted) and pre-existing teleworkers. We applied ordered probit endogenous switching regression models to account for the self-selection bias, where our method improved on the model available in Stata at the time of this analysis (Chiburis & Lokshin, 2007) by enabling the full-information maximum likelihood estimation of alternative-specific outcome equations.[1] We separated teleworkers based on their travel-stress level to uncover potential heterogeneity when analyzing the relationship between teleworking and VMD changes. Our analysis revealed markedly different results between the two groups which were both logical and meaningful, representing a significant contribution to the literature on this topic. To our knowledge, this is the first study to shed light on this important distinction, providing novel insights into the contradictory results found in previous studies. In addition, we introduced the first application (as far as we are aware) to an ordered probit switching regression model of the Yen & Rosinski (2008) approach to back-transforming a log-transformed dependent variable. In general, our mathematical and visual approach to calculating and analyzing treatment effects can be applied to other contexts in which selectivity bias is an issue, making the study relevant to a wide range of research domains.

[1] We thank an anonymous reviewer, whose comment led us to recognize that both Heckman's model and endogenous switching regression models are special cases of conditional mixed-process (CMP) models. In particular, CMP models comprise a broad family involving two or more equations (whose dependent variables can be of varying types, such as binary, ordinal, censored, and continuous) featuring correlated error terms, following normal distributions. Further elaboration on CMP models can be found in the work of Roodman (2011), which details the maximum likelihood estimation of the entire family of models, and provides readers with comprehensive insight into this subject.

Table 1
Summary of literature.

Authors (year)	Method	Data type	Impact on travel demand[1]
Mokhtarian et al. (1995)	Before & after analyses (review of multiple studies)	Panel surveys	S: TWing **reduces personal-miles traveled**
Choo et al. (2005)	Aggregate time series analysis	Aggregate data	S: TWing **reduces vehicle-miles traveled (VMT)**
Zhu (2012)	Ordinary least squares; Two-stage least squares, adjusted by an instrumental variable (IV)	Cross-sectional survey (CSS)	C: Teleworking is **positively** related to trip distance, duration, and frequency for both work and non-work trips
Zhu and Mason (2014)	Left-censored Tobit model, adjusted by an IV	CSS	C: TWers have **more VMT** for both daily work and non-work trips than non-TWers do
He and Hu (2015)	Poisson regression, adjusted by an IV	CSS	C: TWing has a strong **positive impact** on the number of trips
Kim et al. (2015)	Seemingly-unrelated censored regression	CSS	M: TWing **reduces commute** vehicle-kilometers traveled (VKT), but is **partially offset** by other travel demand within the household
Mishra (2017)	Two-stage predictor substitution and control function methods to account for self-selection	CSS	S: TWing has a **substitution effect** on both commute and non-work travel
Chakrabarti (2018)	Negative binomial; Binary logit	CSS	C: TWers tend to have **higher VMD** than non-teleworkers, where occasional teleworkers are even more likely to travel more than frequent teleworkers
de Abreu e Silva & Melo (2018)	Path analysis	CSS	C: TWing **increases weekly VMT**
Elldér (2020)	Tobit regression	CSS	M: TWing full-day is associated with **lower** personal-kilometers traveled (*PKT*) than for those who do not telework, while TWing part-day is associated with **higher** PKT. **No significant impact** of TWing on *VKT* was found for either full-day or part-day TWers.
Su et al. (2021)	Motif and sequence analysis	CSS	C: TWers that have at least one trip during their workday **travel more** (in VMT and trips) than their counterpart commuters
Obeid et al. (2024)	Fixed-effect and first-difference regressions	Panel surveys and point-of-interest data	M/S: TWing **increases non-commute trips**, but **total distance traveled decreases** (so, substitution in terms of distance traveled)

[1] S – substitution effect, indicating teleworking reduces travel demand; C – complementarity effect, indicating teleworking increases travel demand; M – mixed effect, including modification, neutrality, and other hybrid relationships between teleworking and travel demand.

The rest of the paper is structured as follows. We first review the literature regarding telework-related travel demand changes and endogenous switching regression models in Section 2. In Section 3, we describe the dataset used in this study. Section 4 introduces the ordered probit endogenous switching regression model. Section 5 presents the two sets of model results. In Section 6, we discuss the treatment effect of teleworking on VMD. We conclude the paper in Section 7.

2. Literature review

Teleworking inevitably brings travel behavior changes, and thus it has attracted much attention in transportation studies. For example, teleworking may reshape the market share of various transportation modes, and thus influence transit revenues, vehicle ownership, congestion, air quality, and managed lane revenues, among other effects. Teleworking eliminates (some) commute trips, which may further change teleworkers' (and potentially their households') travel patterns for the remaining commute trips and for non-work trips (Yum, 2021). Moreover, the adoption of teleworking may also prompt residential relocation to less dense areas. Recent studies have found that teleworkers have a higher tolerance for commuting distance (Ravalet & Rérat, 2019) and longer commute times (de Vos, Meijers, & van Ham, 2018). Eventually, the less dense built environment may reshape teleworkers' travel patterns in turn. In summary, frequently discussed teleworking-induced behavioral changes include those in travel demand (de Abreu e Silva & Melo, 2018), activity patterns (Su et al., 2021), mode choice (Erhardt et al., 2022), and residential location (Ory & Mokhtarian, 2006; de Abreu e Silva, 2022). In this paper, we focus on travel demand changes – do people travel more, or less, once adopting telework?

As mentioned earlier, the findings regarding whether teleworking reduces or increases travel are mixed. The inconsistent findings may result in part from different measures of travel demand. For example, commonly used measures include commute distance (Zhu, 2013), travel duration (Stiles & Smart, 2021), number of trips (He & Hu, 2015), and vehicle-miles driven/traveled (VMD/VMT, Zhu & Mason, 2014). Different measures reflect different aspects of travel demand: total vs. for a specific travel purpose; travel frequency vs. travel distance vs. travel time. Inconsistent findings may also result from a different breadth of focus. For example, individuals may reduce their own travel due to teleworking, but some of those eliminated trips may transfer to other household members (Kim et al., 2015). Teleworking may reduce commute trips, but the saved time may be used for non-work trips (Zhu, 2012). Lastly, different study contexts (e.g., different countries, urban vs. rural) may also contribute to different teleworking impacts on travel demand. In this study, we focus on the weekly VMD of workers living in two U.S. urban regions. Weekly VMD will cover all types of driving trips, including both commute and non-work travel.

Aside from differences in the study specifications, two other factors may lead to fundamental differences when analyzing teleworking impacts on travel demand: sample selection bias and the selection of an appropriate measure of the outcome of teleworking. The remainder of this section addresses those two factors.

2.1. Sample selection bias

Studies of the travel impacts of teleworking often draw on the program evaluation literature, in which a program (called a "treatment") of some kind (e.g., a new policy, technology, medical treatment, or agricultural practice) is adopted (selected) by some entities (the "treated") but not others (the "untreated"), and the goal is to estimate the effect of the treatment on a logical outcome variable. A widely adopted measure of the treatment effect is to calculate the "difference in differences" using longitudinal data, which compares the difference between "after treatment" and "before treatment" outcome measures for the treated versus untreated groups (Cameron & Trivedi, 2005). Early studies such as those reported by Mokhtarian, Handy, and Salomon (1995) collected panel data before and after teleworking adoption, and results suggested travel reductions after adopting teleworking. On the other hand, most recent studies have used cross-sectional data to analyze teleworking impacts, and many of them have found that teleworking apparently *increases* travel (Zhu & Mason, 2014; He & Hu, 2015; Kim et al., 2015; see Table 1). Significantly, one recent exception to the trend of finding complementarity is a *panel* study rather than a *cross-sectional* one: Obeid et al. (2024) found that on net, people traveled less on teleworking days than on non-teleworking days, although some (but far from all) of the commute reduction savings was partly counteracted by new trip generation. Also significantly, another recent exception (Mishra, 2017) involved cross-sectional data but accounted for selectivity bias, a subject to which we now turn.

Cross-sectional data may yield biased estimates of the treatment effect if the selectivity bias is not handled appropriately. Selectivity bias refers to occasions where unobserved factors influencing the treatment (i.e., teleworking adoption, in our case) also influence the outcome (i.e., VMD). Direct comparisons of the outcomes of treated and untreated groups observed in cross-sectional studies are subject to selectivity bias. Fig. 1 illustrates the difference in treatment effect estimates (TEs) that can occur with longitudinal versus cross-sectional data. If individuals who plan to adopt teleworking have higher *untreated* VMD (at Time 0) than those who do not plan to adopt it, they may still have higher *treated* VMD as teleworkers than observationally equivalent non-teleworkers do (at Time 1, cross-sectional TE), even if their VMD has reduced after teleworking (longitudinal TE). To correct for the potential selection bias in cross-sectional data, various methods have been applied, such as instrumental variables,[2] matching propensity scores, regression-discontinuity design, and the endogenous switching regression model (ESRM, Cameron & Trivedi, 2015). The latter, however, which is the method of the present paper, is particularly well-suited to correct for selection on *unobserved* characteristics (unlike, say, propensity score methods, which can only correct for "selection on observables").

Heckman (1979) proposed a model to address the selection bias that often occurs when modeling a continuous variable that is only observed for part of the population. Specifically, Heckman's original approach contains a binary selection model and a continuous outcome model, where the outcome model is only applicable to those who are observed (selected). Selection bias results when unobserved variables that influence selection are correlated with those that influence the outcome, a situation known as "selection on unobservables". In an evaluation context, the selected cases are those receiving a "treatment" of some kind (such as a medical therapy, a change in their environment, or a policy intended to benefit them), and the goal is to properly evaluate the impact of the treatment on the outcome by correcting for the selection bias in the sample.

In many situations (including the context of the present study), however, the goal remains the same, but an outcome is observed for *all* individuals, whether treated or untreated. The approach in this case retains the selection model and the assumption of correlated selection model and outcome model error terms, but involves potentially different outcome equations for each group (treated and untreated, Cameron & Trivedi, 2005). Such an approach is known as an endogenous switching regression model (ESRM).

2.2. The selection of an appropriate outcome measure

As indicated above, the sample selection modeling approach is employed to obtain unbiased estimates of the effect of a treatment (i.e. program, policy, technology, etc.) on an outcome of interest. The original conceptual rationale of the endogenous switching regression model (ESRM) holds that "individuals participate in a program [i.e., are treated] if net utility [of participation] is positive (or nonnegative) and do not participate if net utility is negative" (p. 211, Heckman, et al., 2001). Along similar lines, Nakosteen and Zimmer (1980, p. 840) argued that individuals "choose among competing alternatives at least in part on the basis of anticipated incremental returns". However, teleworking has various "utilities" or "returns", such as better work-life balance, reduced family conflicts, reduced commute trips, and reduced costs (Thompson et al., 2021). The primary "utility" or "return" from teleworking will differ from person to person, and sometimes, the same outcome (e.g., reduced VMD, or more time at home) may constitute a utility *gain* for one person and a *loss* for another. Thus, at least in the context of teleworking, any single outcome variable will likely be mismatched to the desired gain from treatment for some individuals, which could lead to counterintuitive or at least puzzling results if not recognized. For example, travel-dominant teleworkers likely adopt teleworking to reduce their long commute trips. However, losing the buffer between work and personal life that is provided by commute trips may increase work-family conflicts. In contrast, family-dominant teleworkers may use the saved commute time for household-serving trips, which will reduce their work-family conflicts, but a side effect may be increased net VMD.

[2] Applying an instrumental variables (IV) approach, Zhu and Mason (2014) found that teleworkers have more VMT for both daily work and non-work trips than non-teleworkers. On the other hand, Mishra (Chapter 4, 2017) indirectly questions the validity of the IV used by Zhu and Mason (frequency of home internet use), and using a different IV (potential for teleworking) in a control function model that accounts for selection on unobservables, found that although teleworkers travel more than non-teleworkers, they travel substantially less than they would if not teleworking – supporting the substitution effect of teleworking.

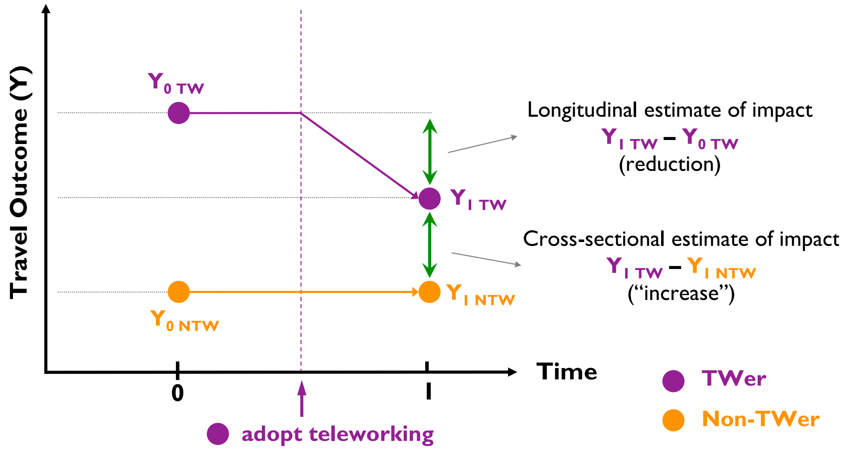

Fig. 1. Illustration of longitudinal and cross-sectional treatment effects.

The present study takes weekly VMD as the outcome variable, and therefore implicitly presumes that reducing VMD is the desired gain leading people to self-select into telework. Knowing that such will not always be the case, we attempt to identify the group for whom it most likely *is* the case, and thus separate the sample into travel-stressed and non-travel-stressed individuals when modeling.

3. Data description[3]

This study employs survey data collected in Spring 2021 (late February through April) in two regions: the Dallas-Fort Worth-Arlington (DFA), Texas and Washington (DC)-Arlington-Alexandria (WAA) metropolitan statistical areas. The study areas were selected by the study sponsor (Cintra, a global transportation infrastructure company, which operates managed-lane concessions in those regions). The survey focuses on teleworking behavior (retrospectively) before, (contemporaneously) during, and (prospectively) after the pandemic. The "during COVID" time point refers to Spring 2021, when vaccinations were available to part of the population and some daily activities had been partially resumed. The present study mainly focuses on respondents' during-COVID teleworking adoption and weekly vehicle-miles driven (VMD). The latter variable is collected by asking respondents the following question: "considering your travel for all purposes (both commute and leisure trips), how many miles did you personally drive in a typical week recently (around February [or March or April] 2021)? *If you are a professional driver (*e.g. *bus, truck, taxi, or Uber/Lyft driver), please do* ***not*** *include the miles you cover as part of your job.*".

The survey was distributed through the Qualtrics platform to (1) a random sample of the sponsor's customer database and (2) a third-party opinion panel. After removing unemployed or self-employed individuals (at the time of taking the survey) and individuals with non-local travel patterns (i.e., one-way commute distance greater than 70 miles or weekly VMD greater than 700 miles), the working sample size for this study was 1,584.

For each study area, we developed sample weights to achieve representativeness (separately by region) on gender, age, race, ethnicity, education, household income, and employment status (full- versus part-time) based on 2019 5-year estimates from the American Community Survey (ACS, https://www.census.gov/programs-surveys/acs), and on self-employment status and teleworking frequency (non-TWer, non-usual TWer [less than three times/week], and usual TWer [3 + times/week]) using both 2019 5-year estimates and 2021 1-year estimates.[4] The weighting process is shown in Fig. 2, and readers can find more details in Wang et al., (2023). In this study, we will utilize the combined sample comprising both DFA and WAA for the analysis. Accordingly, to ensure that the sample weights accurately reflect the relative sizes of the employed populations in the two regions, as well as their spatial distribution within region, we have further rescaled the sample weights using the employment of each county in the study areas (based on the 2019 5-year estimates). The sample weights were not used in model estimation, but were applied to all descriptive statistics, including the central tendency estimates for the variable of interest (i.e., VMD) that are presented in Section 6.

Table 2 presents the weighted descriptive statistics of the full sample, travel-stressed subsample, and non-travel-stressed subsample. We distinguish travel-stressed from non-travel-stressed workers on the supposition that the treatment effect of teleworking on VMD may differ between those two groups (see Section 2.2). The classification of the two subsamples is based on the travel-stressed

[3] A similar section, including some verbatim passages, appears in Wang et al., (2024) and Wang et al., (2023).

[4] The 2020 estimates, and thence the 5-year estimates that included 2020, were distorted by difficulties in implementing standard sampling procedures during the disruption of the pandemic (Burrows et al., 2023). Accordingly, for most variables we used the 2019 5-year estimates, on the assumption that their distributions would have changed relatively little over a few years. For self-employment and teleworking status, however, we developed the sample weights to achieve representativeness on both pre- and during-pandemic distributions of those variables (using the (2019 5-year and 2021 1-year ACS estimates, respectively), so as to better reflect the corresponding changes in the population from before to during the pandemic.

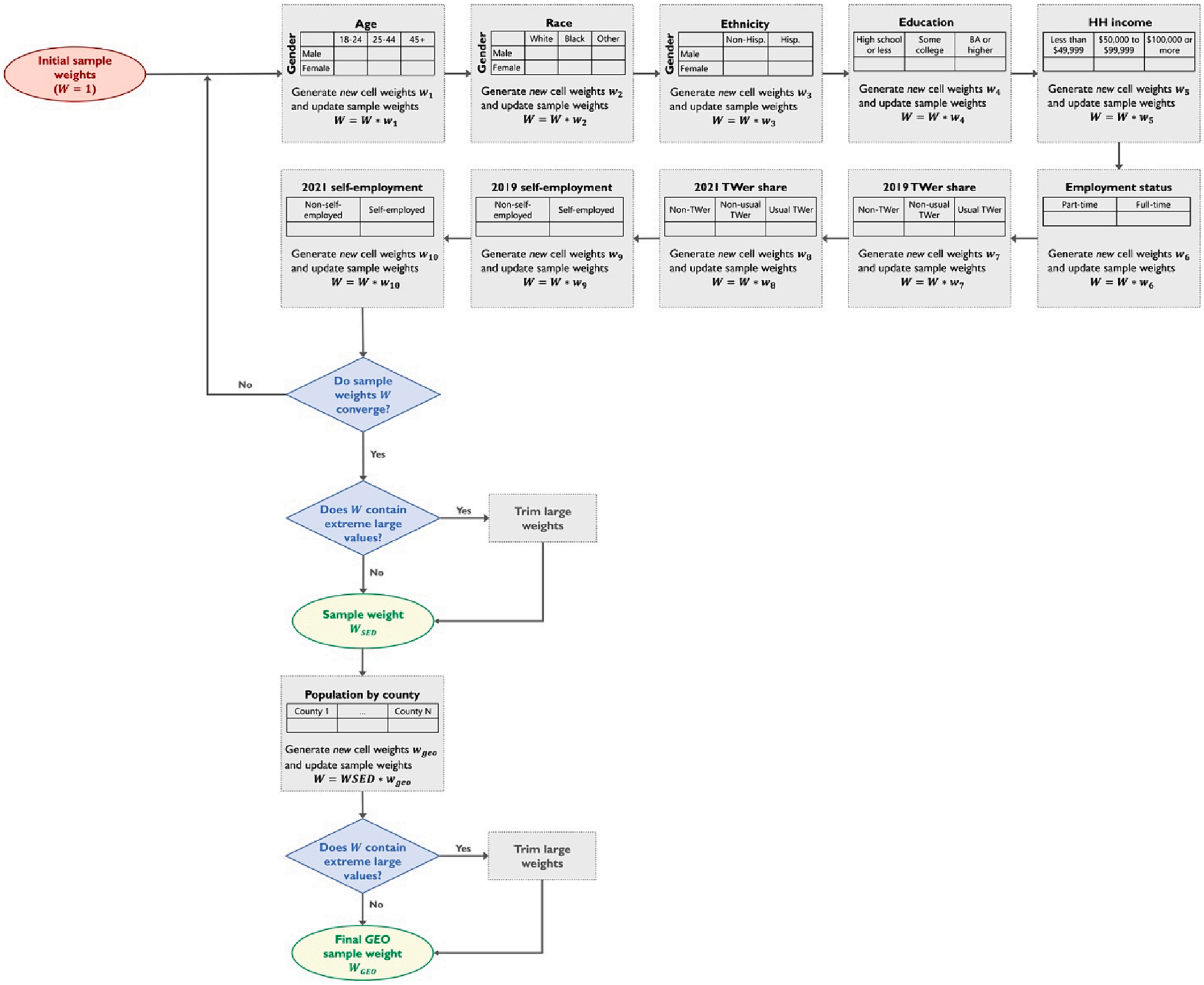

Fig. 2. Weighting flowchart.

attitudinal factor score derived from an exploratory factor analysis (EFA; see full results in Wang et al., 2023). We applied a *neutral-centered* approach to generate factor scores and identified respondents with positive factor scores as travel-stressed and the rest as non-travel-stressed.[5] Fig. 3 presents the weighted distribution of the travel-stressed attitude; its mean is –0.05, where 0 is neutral.

4. Methodology

To address the self-selection bias and quantify the treatment effect of teleworking on VMD, this study will apply endogenous switching regression models (ESRMs, Amemiya, 1985). Specifically, we will estimate three switching regression models: a model for the *full working sample*, a model for a sample that only contains *travel-stressed workers*, and one for a sample that only contains *non-travel-stressed workers*.

Each switching regression model contains one selection model and a set of outcome models. Typically, the selection variable is binary, signifying "treated" and "untreated" states, in which case teleworking adoption would be considered as the treatment. In this study, we further distinguish teleworkers based on their teleworking frequency, considering that people who frequently telework may have different travel patterns than their infrequent counterparts. Therefore, the selection model now has three alternatives, namely, non-teleworkers, non-usual teleworkers (less than 3 times/week), and usual teleworkers (3 or more times/week). Given the ordinal nature of the three categories, we can adopt the ordinal probit model to formulate the selection equation and refer to the whole model system as an ordinal probit switching regression (OPSR, Jimenez & Kugler, 1987).

[5] The highest-loading items associated with the travel-stressed attitude are: "My commute is stressful" (pattern loading 0.667); "I would pay extra to reduce the time I spend in daily traveling" (0.523); and "I'm too busy to do many things I'd like to do" (0.325).

Table 2
Weighted descriptive statistics.

	Full sample		Travel-stressed		Non-travel-stressed	
Unweighted N	1,584		836		748	
Weighted N	1,543.56		762.63		780.93	
	Count	***Share***	***Count***	***Share***	***Count***	***Share***
Gender						
Male	805.89	52.2%	403.27	52.9%	402.62	51.6%
Female	737.67	47.8%	359.36	47.1%	378.32	48.4%
Age						
18–24	128.98	8.4%	61.65	8.1%	67.33	8.6%
25–44	720.16	46.7%	437.20	57.3%	282.97	36.2%
45+	694.41	45.0%	263.78	34.6%	430.63	55.1%
Ethnicity						
Non-Hispanic	1240.42	80.4%	613.59	80.5%	626.83	80.3%
Hispanic	303.14	19.6%	149.03	19.5%	154.11	19.7%
Race						
White only	814.13	52.7%	400.73	52.5%	413.40	52.9%
Black only	311.06	20.2%	132.54	17.4%	178.52	22.9%
Other	418.37	27.1%	229.36	30.1%	189.01	24.2%
Education						
High school or lower	316.65	20.5%	143.91	18.9%	172.74	22.1%
Some college	360.50	23.4%	144.58	19.0%	215.91	27.6%
BA or higher	866.41	56.1%	474.13	62.2%	392.28	50.2%
Household income						
Less than $50,000	325.93	21.1%	145.22	19.0%	180.70	23.1%
$50,000 to $99,999	434.18	28.1%	206.44	27.1%	227.74	29.2%
$100,000 or more	783.45	50.8%	410.96	53.9%	372.49	47.7%
Worker type						
Parttime	282.75	18.3%	119.01	15.6%	163.74	21.0%
Fulltime	1260.81	81.7%	643.62	84.4%	617.19	79.0%
Teleworking frequency						
Never	872.89	56.6%	394.59	51.7%	478.30	61.2%
Less than 3 times/week	231.08	15.0%	137.33	18.0%	93.75	12.0%
3 times/week or more	439.59	28.5%	230.71	30.3%	208.89	26.7%

Note: We developed the sample weights on the original survey sample (N = 2,071), before removing cases for the present study; thus here, the unweighted and weighted total sample sizes do not exactly match.

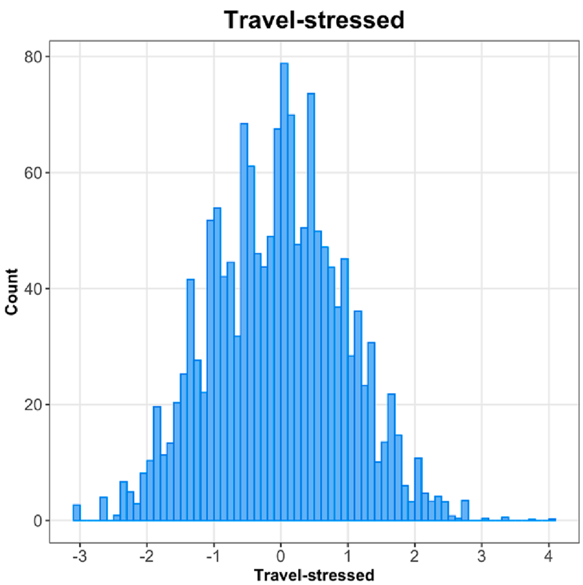

Fig. 3. Weighted distribution of the travel-stressed attitudinal factor score.

The three outcome models will be linear regressions that estimate the (log-transformed) weekly VMD for the three (tele)worker groups, respectively. OPSR models allow correlated residuals between the selection and outcome models, indicating that unobserved traits influencing the teleworking status may also influence the weekly VMD (i.e., capturing the self-selection bias). With the switching regression model, we are able to identify and interpret the self-selection bias and quantify the treatment effect of teleworking status on weekly VMD (Kim & Mokhtarian, 2023b[6]).

As shown in Eq. (1) (individual subscripts are suppressed throughout, for simplicity), we use $\mathscr{Z}$ to represent a latent teleworking frequency propensity:

$$\mathscr{Z} = \boldsymbol{W\gamma} + \varepsilon, \tag{1}$$

where $\boldsymbol{W}$ represents the vector of attributes of an individual, $\boldsymbol{\gamma}$ denotes the corresponding vector of parameters, and ε is the error term, which is assumed to follow a normal distribution $N(0,1)$.

As $\mathscr{Z}$ increases and passes some unknown but estimable thresholds, we move up the selected alternatives from non-teleworking (zero teleworking frequency) to non-usual (less than 3 times/week) and then usual teleworking (3 or more times/week):

$$Z = j \;\; if \;\kappa_{j-1} < \mathscr{Z} \le \kappa_j, \tag{2}$$

where Z is the observed ordinal selection variable, $j = 1, \cdots, J$ indexes the ordinal levels of Z, and the κ_j are the estimable thresholds (except that $\kappa_0 = -\infty$, $\kappa_J = \infty$). $J = 3$ in our application, so we estimate the thresholds κ_1 and κ_2. The probability that level j is selected is expressed as

$$\begin{aligned} P(Z = j) &= P(\kappa_{j-1} < \mathscr{Z} \le \kappa_j) \\ &= P(\kappa_{j-1} - \boldsymbol{W\gamma} < \varepsilon \le \kappa_j - \boldsymbol{W\gamma}) \\ &= \Phi(\kappa_j - \boldsymbol{W\gamma}) - \Phi(\kappa_{j-1} - \boldsymbol{W\gamma}) \end{aligned} \tag{3}$$

We formulate the outcome model using the linear regression specification as shown in Eq. (4). Here, we have J alternative-specific outcome models, where the dependent variable y_j is $\ln(VMD+1)$, $\boldsymbol{X}_j$ is the vector of observed explanatory variables associated with the j^{th} outcome model, $\boldsymbol{\beta}_j$ is the vector of associated parameters, and η_j represents the unobserved influences on y_j, which is assumed to follow a normal distribution ($\eta_j \sim N\left(0, \sigma_j^2\right)$):

$$y_j = \boldsymbol{X}_j\boldsymbol{\beta}_j + \eta_j. \tag{4}$$

The error terms of the selection and outcome models follow the multivariate normal distribution (Eq. (5), where ρ_j represents the correlation between the errors of the selection model (ε) and the j^{th} outcome model (η_j):

$$\begin{pmatrix} \varepsilon \\ \eta_1 \\ \vdots \\ \eta_j \\ \vdots \\ \eta_J \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_1\sigma_1 & \cdots & \rho_j\sigma_j & \cdots & \rho_J\sigma_J \\ \rho_1\sigma_1 & \sigma_1^2 & & & & \\ \vdots & & \ddots & & & \\ \rho_j\sigma_j & & & \sigma_j^2 & & \\ \vdots & & & & \ddots & \\ \rho_J\sigma_J & & & & & \sigma_J^2 \end{pmatrix}\right]. \tag{5}$$

Similar to the classical switching regression model with binary outcomes, there are two main methods for estimating the OPSR model: two-step estimation (e.g., Jimenez & Kugler, 1987) and maximum likelihood estimation (MLE, e.g., Chiburis & Lokshin, 2007). In this study, we applied both methods and found that they produced similar results. However, considering its higher efficiency, we ultimately selected the MLE approach for our analysis. Eq. (6) shows the log-likelihood function for the MLE approach, where $\sum_{\{j\}}$ means the summation over all the cases belonging to teleworking frequency level j. The derivation of the log-likelihood function is provided in Appendix A. Note that Chiburis & Lokshin (2007) developed a Stata package for both two-step and maximum likelihood estimation, which does not allow the outcome equations to differ by selection alternative. In contrast, we enhanced the MLE of the OPSR model by utilizing an improved log-likelihood function that accommodates alternative-specific outcome equations (as described in Eq. (6)), and we implemented this in R using our own code:

$$LL = \sum_{j=1}^{J}\sum_{\{j\}}\left\{ ln\left[\frac{1}{\sigma_j}\phi\left(\frac{y_j - \boldsymbol{X}_j\boldsymbol{\beta}_j}{\sigma_j}\right)\right] + ln\left[\Phi\left(\frac{\sigma_j(\kappa_j - \boldsymbol{W\gamma}) - \rho_j\left(y_j - \boldsymbol{X}_j\boldsymbol{\beta}_j\right)}{\sigma_j\sqrt{1-\rho_j^2}}\right) - \Phi\left(\frac{\sigma_j(\kappa_{j-1} - \boldsymbol{W\gamma}) - \rho_j\left(y_j - \boldsymbol{X}_j\boldsymbol{\beta}_j\right)}{\sigma_j\sqrt{1-\rho_j^2}}\right)\right]\right\}. \tag{6}$$

[6] While Kim and Mokhtarian (2023b) also uses the endogenous switching regression approach and has $\ln(VMD+1)$ (employing a different sample from a different region) as the outcome variable of interest, it diverges from the present study in three material ways: (1) the selection process of interest is *residential* self-selection, whereas the present study focuses on selection into *teleworking* states; (2) it uses only a *binary* selection model (urban versus less-urban residential location), whereas the present study uses an *ordinal* model, which is a nontrivial variation involving original coding of the maximum likelihood estimation routine and original derivations of the back-transformed effects; and (3) the present study provides novel visualizations of the aggregate and disaggregate factual and counterfactual VMD estimates.

5. Model results

In this section, we first present the full sample model, which identifies factors that influence teleworking adoption/frequency and VMD in general. Next, we separate the population into travel-stressed and non-travel-stressed groups and estimate separate ordered probit switching regression models to explore the differences between the two groups.

5.1. Full sample model

Table 3 presents the model result for the full sample (N = 1,584). As indicated by the selection model, which is an ordered probit model of the (during-COVID) teleworking adoption/frequency choice, people with higher education attainments and higher household income tend to have higher propensities for teleworking adoption and/or higher telework frequency[7] – as had long been known, but which also held true during the pandemic. Having a flexible work schedule and being a full-time worker are also positively associated with the propensity for more teleworking. Moreover, teleworking feasibility has a positive effect on the propensity for more teleworking, as would be expected.

Regarding attitudes, we find that people who support active modes (i.e., walking and biking) and who do not prioritize car ownership tend to have higher teleworking frequencies, which could be both/either a motive for, and/or an outcome of, teleworking. People who have a preference for teamwork and perceive that teleworking enables effective teamwork have a higher propensity for teleworking more. Moreover, those who find that teleworking increases their work enthusiasm also tend to telework more often. Lastly, both people who experience work-family conflicts (specifically, perceive that work interferes with family, WIF) and who value the location flexibility afforded by teleworking tend to have a higher propensity for teleworking more, which is consistent with our findings on teleworking-related motives (i.e., family- and flexibility-motivated teleworkers, Wang et al., 2024). More interestingly, once we divide the sample into travel-stressed and non-travel-stressed segments, the two latter attitudes are only significant to the teleworking frequency propensity of non-travel-stressed people, and are insignificant for their travel-stressed counterparts (Section 5.2). The fit of the selection model, which is measured by McFadden's pseudo-R^2, is 0.486 (equally-likely [EL] base) or 0.462 (market share [MS] base). This fit is considered relatively good for disaggregate discrete choice models.

The three outcome equations separately model the log-transformed vehicle-miles driven (VMD) for non-teleworkers (NTWers), non-usual teleworkers (NUTWers), and usual teleworkers (UTWers). The three models share some common explanatory variables, which have consistent impacts on VMD. Specifically, for all three worker groups, people with a pro-car-owning attitude and people living in less urban areas tend to have higher VMD. Moreover, living in rural areas has a substantially stronger impact on VMD for the two teleworker groups than for the non-teleworkers (0.810, 0.844 vs. 0.604; t-statistics for the null hypothesis of equal coefficients = 5.60, 10.08, respectively). For NTWers and UTWers, females and WAA residents tend to drive less than their counterparts, the latter of which may relate to the spatial scale of the two regions and residents' accessibility to the public transit system. For NTWers and NUTWers, full-time employment and a preference for larger residences (potentially indicating a preference for living in less compact areas) are associated with higher driving mileage.

In addition to the common explanatory variables, the outcome models identify unique factors that influence VMD for the three worker groups. Regarding NTWers, the estimated coefficients of age and its quadratic term indicate that VMD increases to a peak around middle age (late 50s) and decreases afterward. White NTWers drive more than their Black counterparts. NTWers with more household vehicles tend to drive more. Regarding *teleworkers*, education and household income are positively associated with VMD for NUTWers and UTWers, respectively. NUTWers who have a preference for active modes tend to drive less. UTWers with more children in their household tend to have higher VMD, potentially due to an increased demand for household-serving travel, such as chauffeuring children.

For the outcome model, we calculated four R^2 measures: a model-specific measure for each of the three worker groups (NTWers 0.178, NUTWers 0.181, UTWers 0.124), and one for the synthesized model (i.e., combining the prediction results of the three separate outcome models, 0.236). In general, the NTWer and NUTWer models have similar measures, both of which are better than the UTWer model, which suggests that the UTWers have more diverse travel patterns and thus less predictable VMD. Specifically, we conjecture that individuals who telework regularly typically enjoy greater flexibility in various aspects of their lives, including work schedules, residential location choices, and lifestyles. Consequently, they may exhibit a more diverse range of behaviors related to non-work trip purposes, departure times, travel modes, and other factors. Predicting their weekly VMD can be inherently more challenging due to this increased variability. To address this complexity, it may be necessary to collect a broader set of variables to enhance the accuracy of VMD predictions for regular teleworkers. Conversely, both non-teleworkers and those who do not telework regularly are typically bound by their daily commutes to work. As such, their travel patterns tend to be more constrained by factors like fixed work schedules, residential location, and workplace location. This reduced degree of variability in their travel behavior results in more conventional and predictable weekly VMD patterns.

Lastly, the correlation between the error term of the selection model and the UTWer outcome model is the largest (0.301) and highly significant, followed by that of the NUTWer model (0.128), indicating the existence of self-selection bias for both UTWers and

[7] To simplify the language used throughout the rest of this paper, we will refer to teleworking adoption and frequency changes using phrases such as "telework more" or "telework less", which encompass any ordinal level change in the three teleworking frequency categories. The decision to "adopt" or "stop" teleworking is also covered by such phrases– the teleworking frequency will change either from zero to a positive value ("telework more") or from a positive value to zero ("telework less").

Table 3
Full sample model results.

Selection model (N = 1,584) $\hat{\gamma}$	*Coef.*	*s.e.*	
Education (ref: high school or less)			
Some college	0.316	0.136	*
Bachelor's degree or higher	0.466	0.131	***
Annual household income (ref: less than $50,000)			
$50,000 to $99,999	0.0623	0.114	
$100,000 or more	0.247	0.112	*
Flexible work schedule[a]	0.307	0.094	**
Full-time worker	0.330	0.092	***
Teleworking feasibility	0.132	0.00504	***
Attitudes			
Pro-active-mode	0.0788	0.0385	*
Pro-car-owning	−0.0836	0.0404	*
Work interferes with family (WIF)	0.113	0.0391	**
Pro-teamwork	0.0851	0.0366	*
TW effective teamwork	0.322	0.0436	***
TW enthusiasm	0.0853	0.0372	*
TW location flexibility	0.0849	0.0392	*
Threshold			
1 | 2	1.233	0.164	***
2 | 3	2.461	0.175	***
Selection model fit (McFadden's pseudo-R^2)			
Equally likely (EL) base	*0.486*		
Market share (MS) base	*0.462*		

Outcome model	**NTWer (N = 535)**			**NUTWer (N = 322)**			**UTWer (N = 727)**		
$\hat{\beta}$	*Coef.*	*s.e.*		*Coef.*	*s.e.*		*Coef.*	*s.e.*	
Intercept	3.644	0.263	***	2.490	0.344	***	2.378	0.238	***
Female	−0.206	0.107	·	–	–		−0.364	0.111	**
Age	0.00942	0.00367	*	–	–		–	–	
Age squared	−0.000412	0.000242	·	–	–		–	–	
Race (ref: White)									
Black	−0.400	0.181	*	–	–		–	–	
Other races	−0.0573	0.180		–	–		–	–	
Education (ref: high school or less)									
Some college	–	–		0.150	0.287		–	–	
Bachelor's degree or higher	–	–		0.624	0.273	*	–	–	
Annual household income (ref: less than $50,000)									
$50,000 to $99,999	–	–		–	–		0.465	0.207	*
$100,000 or more	–	–		–	–		0.310	0.196	
Number of children	–	–		–	–		0.177	0.0579	**
Number of vehicles	0.124	0.0489	*	–	–		–	–	
Residential location (ref: urban)									
Suburban	0.0683	0.150		0.445	0.161	**	0.283	0.135	*
Small town	0.475	0.202	*	0.189	0.296		0.291	0.230	
Rural	0.604	0.215	**	0.810	0.306	**	0.884	0.269	***
Full-time worker	0.452	0.126	***	0.692	0.161	***	–	–	
Attitudes									
Pro-large-house	0.179	0.0553	**	0.180	0.0821	*	–	–	
Pro-active-mode				−0.176	0.0748	*	–	–	
Pro-car-owning	0.142	0.0632	*	0.159	0.0832	·	0.248	0.0530	***
Region indicator (WAA)	−0.247	0.109	*	–	–		−0.271	0.111	*
Outcome model fit									
R^2 (model-specific)	*0.178*			*0.181*			*0.124*		
R^2 (combined)				*0.236*					

Error terms	*Coef.*	*s.e.*		*Coef.*	*s.e.*		*Coef.*	*s.e.*	
$\hat{\sigma}$	1.178	0.0361	***	1.233	0.0485	***	1.431	0.0389	***
$\hat{\rho}$	0.054	0.0897		0.128	0.0613	*	0.301	0.0787	***

*** Coefficient is statistically significant at the 0.001 level.
** Coefficient is statistically significant at the 0.01 level.
* Coefficient is statistically significant at the 0.05 level.
· Coefficient is statistically significant at the 0.1 level.
[a] Flexible work schedule is characterized in the survey as "e.g., organizing my own work hours", and is differentiated from compressed work week and "variable start time/rotating shift" responses.

NUTWers. Specifically, a significant positive correlation means that unobserved attributes that increase people's teleworking frequency propensity also tend to increase their VMD (consistent with our presupposition that teleworkers tend to travel more than observationally equivalent others even before they start teleworking), or, the unobserved attributes that decrease people's teleworking frequency propensity also decrease their VMD. However, such an association between selection and outcome is not exhibited to a significant degree for NTWers.

5.2. Travel-stressed and non-travel-stressed models

In the full-sample model, we identified self-selection biases for the two teleworker groups. However, we did not find evidence to support the existence of a self-selection bias for non-teleworkers. This may reflect the truth – i.e. the unobservables associated with teleworking adoption/frequency choice do not influence non-teleworkers' VMD – or, the evidence of self-selection bias for non-teleworkers may be masked by other factors such as mixing teleworking-related motives.

As discussed earlier, we identified different teleworking motives in a companion paper (Wang et al., 2024). Teleworkers with different motives expect improvement in different life aspects once adopting teleworking: travel-motivated teleworkers may expect to reduce long commute trips, family-motivated teleworkers may expect to have a more flexible schedule for family duties, and workplace-discouraged teleworkers may expect to avoid unpleasant social encounters at the workplace. VMD, a critical indicator for transportation system and travel behavior studies, can be a good measure of the teleworking outcome for the travel-motivated teleworkers, but it may not reflect the dominant benefit gain for teleworkers with other motives. Mismatch between teleworking motives and the outcome measure being studied by the analyst may confound the results. For example, if travel-motivated teleworkers tend to reduce their VMD but family-motivated teleworkers often increase it, the effects of teleworking on VMD will be blurred accordingly.

For some purposes, we may in fact want to know the VMD change due to teleworking at an aggregate level, i.e., considering all teleworkers regardless of their teleworking motives (Section 5.1). However, to better understand the behavior changes and provide targeted policy suggestions, we also want to match the outcome measure to the major benefit of teleworking for the individuals in question. In the present study, we focus on the VMD outcome, and accordingly, we divide the full sample into travel-stressed and non-travel-stressed subsamples so that we can compare teleworking impacts on VMD for the case when that outcome variable *matches* the (presumed) major benefit of teleworking to the case when it does not.

Table 4 presents the switching regression model results for travel-stressed and non-travel-stressed groups, separately. Regarding the explanatory variables for the selection model (i.e., modeling the decision on teleworking frequency), sociodemographic variables such as education attainment and having a flexible work schedule are significant in both travel-stressed and non-travel-stressed models, exhibiting results that are consistent with the full sample model. For attitudinal variables, although the two subsample models still have results consistent with those of the full sample model, the significant attitudinal variables are almost disjoint between the travel-stressed and non-travel-stressed models. Specifically, the significant attitudinal variables in the travel-stressed model are mode- or work-related (e.g., pro-active-modes, performance measurable, TW effective teamwork), whereas many attitudinal variables in the non-travel-stressed model are associated with lifestyle preferences that may relate to their decision on teleworking (e.g., pro-large-house, work-interferes-with-family, TW location flexibility). The split of attitudinal variables between the two segments further demonstrates the existence of heterogeneous motivations, as we found in Wang et al., (2024).

For the outcome model (i.e., modeling the weekly VMD), the two subsample models overall have consistent results with the full sample model. To name a few examples: females have lower VMD than males across travel-stressed UTWers and non-travel-stressed NTWers and UTWers; people living in less urban areas generally have higher VMD for most of the worker groups; pro-car ownership is also positively associated with VMD. It is worth noting that individuals who prefer larger residences generally have higher VMD, which may be due to their preference for living in less dense areas. However, among travel-stressed UTWers, those who prefer larger residences exhibit a lower weekly VMD. We speculate that such UTWers who are pro-large-house may be individuals who prefer a home-centered lifestyle, with larger homes allowing for more amenities (in both indoor and outdoor spaces) in support of that preference.

Regarding the error terms, we find several positive correlations between the selection model and the outcome model errors. Specifically, unobserved attributes that increase the teleworking frequency propensity (via ε) tend to induce higher VMD (via η) for both travel-stressed and non-travel-stressed UTWers, where the two equations' unobserved terms have significantly more commonality in the latter case (correlations of 0.189 vs. 0.443, t-statistic testing equality of parameters = −41.60). The unobserved variables, most of which are likely non-travel-related (e.g., personality, lifestyle, family needs), are typically not well-captured by a travel-focused survey like ours. Therefore, such unobserved variables could well be related to the (less stereotypical) travel demand of non-travel-stressed UTWers as well as to their motives for teleworking (i.e., associated with adopting teleworking due to family needs, personal preferences, etc.). Beyond the two UTWer groups, we also find significant positive correlations for non-travel-stressed NUTWers and travel-stressed NTWers. Thus, similarly to the results for UTWers, unobserved attributes that increase (or decrease) the teleworking frequency propensity for those other groups will also tend to induce higher (or lower) VMD. Without distinguishing the travel-stressed and non-travel-stressed subsamples, the selection bias for travel-stressed non-teleworkers was masked in the full sample model.

Regarding model fit, the selection models for both the travel-stressed and non-travel-stressed groups have similar or superior performance (EL-based pseudo-R^2: 0.494/0.488, MS-based: 0.465/0.458) to that of the full-sample model (EL: 0.486, MS: 0.462). Moreover, in terms of outcome models, both the travel-stressed and non-travel-stressed models outperform the full sample model for NTWers and NUTWers. The travel-stressed UTWer model R^2 (0.117) is slightly lower than that of the corresponding full sample model (0.124), and markedly lower than the non-travel-stressed UTWer model R^2 (0.165). Overall, UTWers' (ln)VMD is harder to predict than that of NTWers and NUTWers. NUTWer model R^2s are slightly higher than, but basically similar to, those of NTWer models.

Table 4
Travel-stressed vs. non-travel-stressed model results.

Selection model	**Travel-stressed (N = 836)**			**Non-travel-stressed (N = 748)**		
$\widehat{\gamma}$	*Coef.*	*s.e.*		*Coef.*	*s.e.*	
Education (ref: high school or less)						
Some college	0.272	0.190		0.235	0.187	
Bachelor's degree or higher	0.578	0.176	**	0.433	0.176	*
Flexible work schedule[a]	0.387	0.136	**	0.327	0.130	*
Full-time worker	–	–		0.475	0.131	***
Teleworking feasibility	0.144	0.00713	***	0.130	0.00703	***
Attitudes						
Pro-active-modes	0.137	0.0507	**	–	–	
Pro-large-house	–	–		0.137	0.0543	*
Work interferes with family	–	–		0.141	0.0636	*
Pro-teamwork	–	–		0.111	0.0508	*
Performance measurable	−0.114	0.0495	*	–	–	
TW effective teamwork	0.350	0.0617	***	0.318	0.0585	***
TW location flexibility	–	–		0.121	0.0548	*
Region indicator (WAA)	−0.183	0.105	.	–	–	
Threshold						
1 | 2	0.894	0.174	***	1.529	0.196	***
2 | 3	2.295	0.196	***	2.591	0.214	***
Selection model fit (McFadden's pseudo-R^2)						
Equally likely (EL) base	*0.494*			*0.488*		
Market share (MS) base	*0.465*			*0.458*		

Outcome model: NTWer	**Travel-stressed (N = 230)**			**Non-travel-stressed (N = 305)**		
$\widehat{\beta}$	*Coef.*	*s.e.*		*Coef.*	*s.e.*	
Intercept	4.788	0.300	***	3.034	0.320	***
Female	–	–		−0.352	0.137	**
Age	0.00984	0.00598		–	–	
Age squared	−0.000924	0.000363	*	–	–	
Race (ref: White)						
Black	–	–		−0.645	0.213	**
Other races	–	–		−0.346	0.233	
Education (ref: high school or less)						
Some college	–	–		0.395	0.179	*
Bachelor's degree or higher	–	–		0.311	0.182	.
Household income (ref: less than $50,000)						
$50,000 to $99,999	0.371	0.216	.	–	–	
$100,000 or more	0.686	0.209	**	–	–	
Number of vehicles				0.156	0.0612	*
Full-time worker	–	–		0.473	0.143	***
Residential location (ref: urban)						
Suburban	0.199	0.210		–	–	
Small town	0.524	0.306	.	–	–	
Rural	0.876	0.347	*	–	–	
Attitudes						
Pro-large-house	–	–		0.172	0.0697	*
Pro-active-modes	–	–		−0.131	0.0732	.
Pro-car-owning	–	–		0.244	0.0878	**
Commute positive	−0.181	0.0743	*	–	–	
Work interferes with family (WIF)	0.167	0.0797	*	–	–	
Region indicator (WAA)	−0.546	0.170	**	–	–	

Outcome model: NUTWer	**Travel-stressed (N = 189)**			**Non-travel-stressed (N = 133)**		
$\widehat{\beta}$	*Coef.*	*s.e.*		*Coef.*	*s.e.*	
Intercept	3.038	0.475	***	1.710	0.452	***
Education (ref: high school or less)						
Some college	−0.758[b]	0.422	.	1.116	0.386	**
Bachelor's degree or higher	−0.212	0.396		1.263	0.379	***
Household income (ref: less than $50,000)						
$50,000 to $99,999	–	–		0.646	0.323	.
$100,000 or more	–	–		0.840	0.315	**
Full-time worker	0.823	0.210	***	0.584	0.242	*
Residential location (ref: urban)						
Suburban	0.696	0.206	***	0.237	0.230	
Small town	0.472	0.380		0.119	0.440	

(*continued on next page*)

Table 4 (*continued*)

Rural	0.743	0.474		0.686	0.379	·
Attitudes						
Pro-active-mode	−0.264	0.0973	**	–	–	
Pro-car-owning	0.248	0.100	*	–	–	

Outcome model: UTWer	**Travel-stressed (N = 417)**			**Non-travel-stressed (N = 310)**		
$\widehat{\beta}$	*Coef.*	*s.e.*		*Coef.*	*s.e.*	
Intercept	2.535	0.211	***	2.541	0.245	***
Female	−0.441	0.142	**	−0.355	0.154	*
Number of children	0.272	0.0765	***	–	–	
Residential location (ref: urban)						
Suburban	0.350	0.179	·	0.436	0.213	*
Small town	0.398	0.318		0.739	0.344	*
Rural	0.610	0.367	·	1.465	0.407	***
Attitudes						
Pro-large-house	−0.200	0.0836	*	0.194	0.0861	*
Pro-car-owning	0.335	0.0714	***	0.212	0.0806	**
Urbanite	–	–		0.195	0.0801	*
Region indicator (WAA)	–	–		−0 .431	0.158	**

Outcome model fit		
R^2 (non-TWer)	*0.237*	*0.222*
R^2 (non-usual TWer)	*0.238*	*0.229*
R^2 (usual TWer)	*0.117*	*0.165*
R^2 (combined)	*0.270*	*0.264*

Error terms	**Travel-stressed**			**Non-travel-stressed**		
	Coef.	*s.e.*		*Coef.*	*s.e.*	
$\widehat{\sigma}_{NTW}$	1.220	0.0714	***	1.125	0.0465	***
$\widehat{\sigma}_{NUTW}$	1.231	0.063	***	1.186	0.0736	***
$\widehat{\sigma}_{UTW}$	1.446	0.0505	***	1.375	0.0592	***
$\widehat{\rho}_{NTW}$	0.478	0.192	*	0.0649	0.116	
$\widehat{\rho}_{NUTW}$	0.0624	0.0818		0.341	0.0866	***
$\widehat{\rho}_{UTW}$	0.189	0.103	·	0.443	0.108	***

*** Coefficient is statistically significant at the 0.001 level.

** Coefficient is statistically significant at the 0.01 level.

* Coefficient is statistically significant at the 0.05 level.

· Coefficient is statistically significant at the 0.1 level.

[a] Flexible work schedule is characterized in the survey as "e.g., organizing my own work hours", and is differentiated from compressed work week and "variable start time/rotating shift" responses.

[b] The marginally significant (t-statistic = -1.797, p-value = 0.072) negative influence on VMD of having some college (relative to a high school education or less) is unexpected, and unique to this model. Investigation revealed that there were only 12 travel-stressed NUTWers having a high school education or less, and among those, two cases reported 350 weekly miles driven and a second reported 300, which yielded an average weekly VMD for the category of 149.9 mi/week (compared to 125.9 for those with some college, and 139.0 for those with a bachelor's degree or higher: 139.0). Thus, we believe that this result is an artifact of the small sample for the reference category, and therefore should not be given too much emphasis. We retain education, however, as a useful control variable overall.

6. Discussion: Treatment effects

In this section, we apply the three switching regression models presented in Section 5 to analyze the treatment effect of teleworking adoption/frequency choices on VMD. We first present the technical details of the treatment effect calculations in Section 6.1. In Sections 6.2 and 6.3, we discuss the treatment effect for the full working sample and the travel-stressed vs. non-travel-stressed samples, respectively.

6.1. Technical details on treatment effects

At the individual level, the treatment effect reflects the difference between the expectation of the variable of interest (i.e., VMD in this study) if the individual is treated and if she is not treated (i.e., teleworks, at either the usual or the non-usual level, or not).

Therefore, to calculate the treatment effect, we first define the conditional expectation in Eq. (7), where the "condition" is the (un)treated status at level j ($Z = j$):

$$\begin{aligned}\mathbb{E}\left[y_j|Z=j\right] &= \boldsymbol{X}_j\boldsymbol{\beta}_j + \mathbb{E}\left[\eta_j|\kappa_{j-1} - \boldsymbol{W\gamma} \ < \ \varepsilon \leq \kappa_j - \boldsymbol{W\gamma}\right] \\ &= \boldsymbol{X}_j\boldsymbol{\beta}_j - \rho_j\sigma_j\frac{\phi\left(\kappa_j - \boldsymbol{W\gamma}\right) - \phi\left(\kappa_{j-1} - \boldsymbol{W\gamma}\right)}{\Phi\left(\kappa_j - \boldsymbol{W\gamma}\right) - \Phi\left(\kappa_{j-1} - \boldsymbol{W\gamma}\right)},\end{aligned} \tag{7}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the density function and cumulative distribution function of the standard normal distribution, and $-\frac{\phi(\kappa_j - \boldsymbol{W\gamma}) - \phi(\kappa_{j-1} - \boldsymbol{W\gamma})}{\Phi(\kappa_j - \boldsymbol{W\gamma}) - \Phi(\kappa_{j-1} - \boldsymbol{W\gamma})}$ is the ordered probit switching regression model counterpart to the inverse Mills ratio (IMR) term of a binary switching regression model, which we will call the "quasi-IMR term" (see Appendix B for the derivation[8]). Since, in this case, the potential teleworking frequency level (i.e., the subscript of the dependent variable y) matches the observed teleworking state ($Z = j$), we refer to the associated expected outcomes as the "factual VMD".

However, to obtain unbiased treatment effects, we must also evaluate the "counterfactual VMD", which reflects the expected VMD measures when the potential teleworking frequency level *does not* match the observed teleworking frequency. Specifically, Eq. (8) is the expected weekly $\ln(VMD+1)$ of an individual at teleworking level j, *if she were to adopt* teleworking level j' (where $j' = j$ is now the factual case, and $j' \neq j$ yields the counterfactual cases):

$$\begin{aligned}\mathbb{E}\left[y_{j'}|Z=j\right] &= \boldsymbol{X}_{j'}\boldsymbol{\beta}_{j'} + \mathbb{E}\left[\eta_{j'}|\kappa_{j-1} - \boldsymbol{W\gamma} < \varepsilon \leq \kappa_j - \boldsymbol{W\gamma}\right] \\ &= \boldsymbol{X}_{j'}\boldsymbol{\beta}_{j'} - \rho_{j'}\sigma_{j'}\frac{\phi\left(\kappa_j - \boldsymbol{W\gamma}\right) - \phi\left(\kappa_{j-1} - \boldsymbol{W\gamma}\right)}{\Phi\left(\kappa_j - \boldsymbol{W\gamma}\right) - \Phi\left(\kappa_{j-1} - \boldsymbol{W\gamma}\right)}.\end{aligned} \tag{8}$$

The difference between the factual VMD and counterfactual VMD counterparts consists of the difference in the impact of the variables on VMD ($\boldsymbol{\beta}_j$ vs. $\boldsymbol{\beta}_{j'}$) and the impact of self-selection $\{(\rho_j\sigma_j)$ quasi-IMR$(\cdot)\}$ vs. $\{(\rho_{j'}\sigma_{j'})$ quasi-IMR$(\cdot)\}$.

Please note, Eqs. (7–8) provide the conditional expectation of the dependent variable, which is the *natural*-log-*transformed VMD*. To obtain the treatment effect in the original VMD scale for ease of understanding, we need to back-transform $\ln(VMD+1)$. Here, we adopt the Yen & Rosinski (2008) derivation process and adjust it for the OPSR model (Eqs. (9–10)). We have not seen these equations, and an application to an ordered probit switching regression model with a log-transformed outcome, elsewhere in the literature:

$$\mathbb{E}\left[VMD_j|Z=j\right] = e^{X_j\beta_j + \frac{\sigma_j^2}{2}}\left[\frac{\Phi\left(\kappa_j - \boldsymbol{W\gamma} - \rho_j\sigma_j\right) - \Phi\left(\kappa_{j-1} - \boldsymbol{W\gamma} - \rho_j\sigma_j\right)}{\Phi\left(\kappa_j - \boldsymbol{W\gamma}\right) - \Phi\left(\kappa_{j-1} - \boldsymbol{W\gamma}\right)}\right] - 1 \tag{9}$$

for the factual case, and

$$\mathbb{E}\left[VMD_{j'}|Z=j\right] = e^{X_{j'}\beta_{j'} + \frac{\sigma_{j'}^2}{2}}\left[\frac{\Phi\left(\kappa_j - \boldsymbol{W\gamma} - \rho_{j'}\sigma_{j'}\right) - \Phi\left(\kappa_{j-1} - \boldsymbol{W\gamma} - \rho_{j'}\sigma_{j'}\right)}{\Phi\left(\kappa_j - \boldsymbol{W\gamma}\right) - \Phi\left(\kappa_{j-1} - \boldsymbol{W\gamma}\right)}\right] - 1 \tag{10}$$

for the counterfactual case (when $j' \neq j$).

As discussed, the OPSR model permits estimation of VMD for all cases under both factual and counterfactual conditions. In this study, we generate the VMD when untreated (i.e., if not teleworking), when NUTW-treated (i.e., if teleworking less than 3 times a week), and when UTW-treated (i.e., if teleworking 3 or more times a week), as shown in Fig. 4. For each *column* in Fig. 4, we compare VMD across different (tele)worker segments while holding the "potential" teleworking status constant. Recalling the illustration in Fig. 1, the VMD estimates are now comparable since all cases are at the "same time point" (i.e., under the same treatment status). For each *row*, we compare VMD across different potential and actual (tele)working statuses for the same (tele)worker group. These latter differences are known as *treatment effects (*i.e.*, treated VMD – untreated VMD)*. In the figure, we illustrate the treatment effect on the untreated (if NUTWing and if UTWing), treatment effect on the NUTW-treated, and treatment effect on the UTW-treated.

6.2. Treatment effects for the full sample

The treatment effect results of the full sample model are displayed in Fig. 5. Fig. 5(a)-(c) shows the distributions of predicted VMD by (tele)worker group. Each figure presents a pair of (un)treated teleworking statuses as the margins. Fig. 5(d) shows the treatment effect by (tele)worker group at the aggregate level using the weighted average VMD, where the teleworking frequencies of 0.80 days/week (for NUTWing) and 4.42 days/week (for UTWing) are the empirical average values for the corresponding observed teleworker

[8] Chiburis & Lokshin (2007) presented similar equations for their two-step estimates. However, in their operationalization of the approach they required the outcome equations to have the same specification for all selection alternatives (although the coefficients could differ), whereas we allowed the outcome equations to differ by selection alternative (Section 4). We use Eqs. (7–8) to calculate the factual and counterfactual ln(VMD+1), respectively, taking into account the impact of alternative-specific equations. To provide readers with a better understanding of the process, we offer derivations of the key equations in Appendix B.

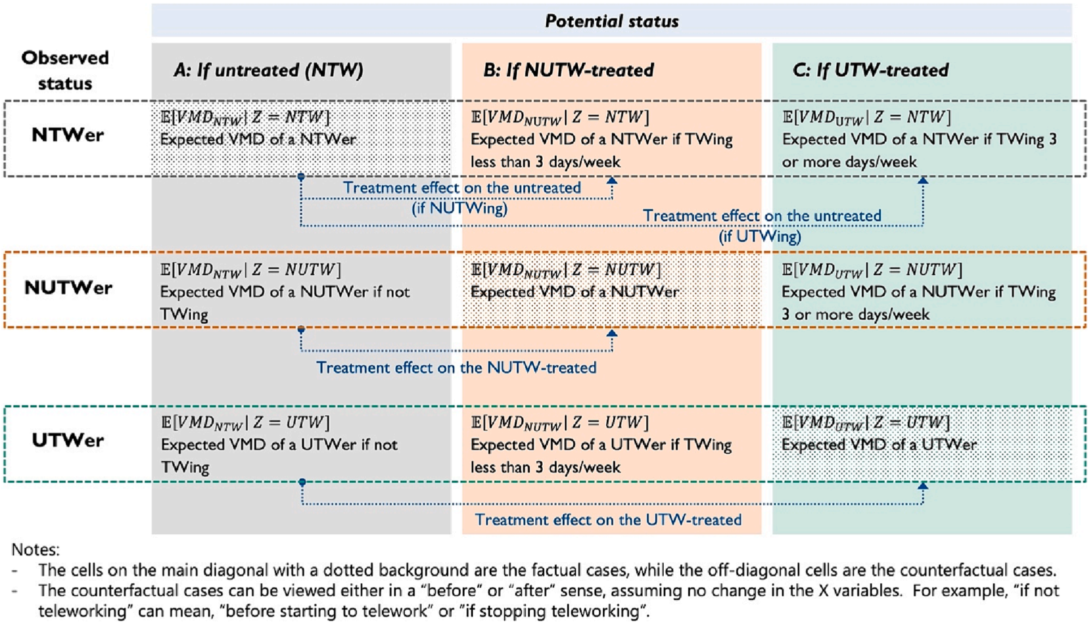

Fig. 4. Factual and counterfactual VMD definitions.

groups.

In Fig. 5(a)-(c), the dashed reference line marks the instances where VMD is equal for both of the paired (un)treated teleworking states. Taking Fig. 5(a) as an example, cases under the reference line generate *more VMD when they do not telework*, whereas cases above the reference line generate *more VMD when they non-usually telework (*i.e.*, telework less than three times a week)*. Overall, in Fig. 5(a) the cases are fairly symmetrically distributed along the reference line, but with those who *decrease VMD when (non-usually) teleworking* (compared to not teleworking) dominating those who *increase* VMD (54% to 46%).[9] For Fig. 5(b) and 5(c), most cases would have reduced VMD if they shifted from non-teleworking to usual teleworking (5(b)), or if they increased teleworking frequency from NUTWing to UTWing (5(c)).

As shown in Fig. 5(d), on average, the three groups of (tele)workers have similar untreated VMD, while UTWers have the highest NUTW-treated VMD, followed by NUTWers and NTWers. The two teleworker groups have similar VMD when UTW-treated, both of which are higher than for NTWers when UTW-treated. When increasing teleworking frequency from zero to the NUTWing level, VMD slightly decreases for NTWers and NUTWers, but slightly increases for UTWers. When the teleworking frequency further increases to the UTWing level, the weekly VMD drops for all three (tele)worker groups.

Focusing specifically on the treatment effects illustrated in Fig. 4, the ***treatment effect on the NUTW-treated*** is −7.54 miles per week (−9.42 miles per teleworking day), which is not statistically significant (see Table 5). In other words, adopting non-usual teleworking (less than 3 times/week) did not significantly reduce VMD for NUTWers. On the other hand, ***the treatment effect on the UTW-treated*** is significant: weekly VMD is reduced by 92.87 miles (21.01 miles per teleworking occasion). For the untreated group, i.e. NTWers, both ***NUTWing and UTWing treatment effects (on the untreated)*** are significant, namely –32.74 miles/week (−40.93 miles/TW occasion) and −111.38 (−25.19), respectively, indicating that NTWers would reduce VMD (on average) once they adopt any level of teleworking.

In summary, adopting UTWing can reduce VMD for most cases, while adopting NUTWing does not necessarily reduce VMD. Many individuals even have a higher VMD once they start NUTWing (Fig. 5(a) and footnote 9). In fact, for UTWers, their *average* NUTW-treated VMD is even higher than their untreated VMD (Fig. 5(d)). Noting that NUTWers are likely to end up being the larger group of teleworkers (Wang et al., 2023), it suggests that we will see increases in the number of teleworkers who generate more vehicle travel than they save. In the next section, we will update our conclusions with more nuanced insights gained from the examination of travel-stressed and non-travel-stressed models.

[9] Be aware that the scatterplots cannot reflect the sample weights. By (tele)worker group, the weighted ratios of decreasing VMD to increasing it are 66:34 for NTWers, 35:65 for NUTWers, and 54:46 for UTWers. Thus, nearly twice as many NUTWers have *higher* VMD when teleworking as have *lower* VMD, although the average VMD for NUTWers is still (slightly) lower when they are teleworking than when they are not (Fig. 5(d); also see discussion in the main text).

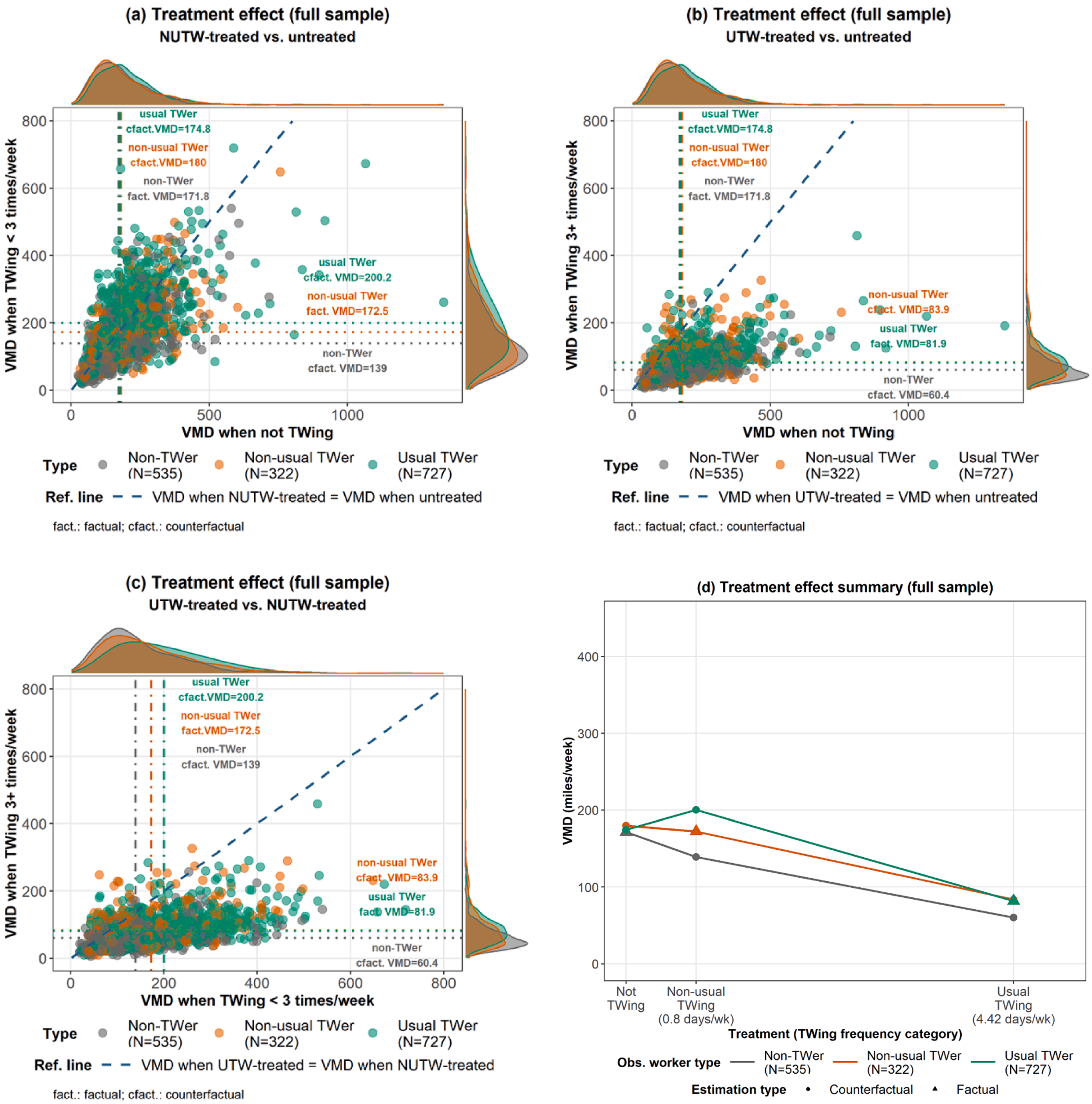

Fig. 5. Treatment effects of teleworking, full sample.

6.3. *Treatment effects for the travel-stressed and non-travel-stressed subsamples*

In this section, we present the treatment effect results for travel-stressed and non-travel-stressed models side by side, allowing for easy comparison. Starting with a visualization and a concise discussion at a disaggregate level, we provide a broad overview of the VMD distribution under paired treatment statuses (Section 6.3.1, Fig. 6). We derive our major conclusions through comparing the aggregate level results, namely the weighted average VMD by (tele)worker group under different treatment statuses (Section 6.3.2, Fig. 7).

6.3.1. *Disaggregate analysis*

Fig. 6(a/c/e) and 6(b/d/f) presents the distributions of VMD for the travel-stressed and non-travel-stressed subsamples,

Table 5
Estimated factual and counterfactual mean VMD and treatment effects.

	Observed worker group								
	Full sample			*Travel-stressed*			*Non-travel-stressed*		
	NTWer	NUTWer	UTWer	NTWer	NUTWer	UTWer	NTWer	NUTWer	UTWer
Observed mean VMD	137.51	142.62	63.19	160.92	138.98	70.38	118.19	147.96	55.26
Estimated mean VMD									
NTWing	**171.79**	180.03	174.80	**211.24**	294.13	411.71	**144.90**	165.98	140.66
NUTWing	139.05	**172.49**	200.22	202.62	**212.28**	216.53	87.37	**148.96**	186.11
UTWing	60.41	83.89	**81.93**	73.41	90.33	**91.53**	51.93	85.74	**74.48**
Treatment effect, mi/week (p-value)									
NTWing→NUTWing	**−32.74**[3a]	−7.54[1]	**25.42**	−8.61[3a]	**−81.84**[1]	**−195.18**	**−57.53**[3a]	−17.02[1]	**45.45**
	(0.000)	(0.266)	**(0.000)**	(0.586)	**(0.000)**	**(0.000)**	**(0.000)**	(0.114)	**(0.000)**
NTWing→UTWing	**−111.38**[3b]	**−96.13**	**−92.87**[2]	**−137.82**[3b]	**−203.79**	**−320.18**[2]	**−92.97**[3b]	**−80.24**	**−66.18**[2]
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
NUTWing→UTWing	**−78.64**	**−88.60**	**−118.29**	**−129.21**	**−121.95**	**−125.01**	**−35.45**	**−63.22**	**−111.63**
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**

Note: The bolded estimated VMDs are factual VMDs, whereas the rest are counterfactual VMDs. The treatment effects are based on the (tele)working status changes displayed in the row header. For each (tele)worker group (displayed in the column header), we conducted a weighted *t*-test on the difference of means to ascertain whether the VMDs are significantly different while changing from one (tele)working status to another (displayed in the row header). Bolded p-values are statistically significant at the 0.05 level.

[1] Treatment effect on the NUTW-treated.

[2] Treatment effect on the UTW-treated.

[3a] NUTWing treatment effect on the untreated.

[3b] UTWing treatment effect on the untreated.

respectively. Each part of the figure presents a pair of (un)treated teleworking statuses as the margins. Fig. 6(a) and 6(b) focuses on the treatment of NUTWing, where the x-axis is the untreated VMD and the y-axis is the NUTW-treated VMD. For travel-stressed individuals (Fig. 6(a)), a weighted 62%[10] fall below the reference line, indicating that they would have a reduced VMD once adopting NUTWing. However, for the non-travel-stressed individuals (Fig. 6b), a large share of teleworkers, especially UTWers (70%), would have a *higher* VMD if adopting NUTWing, compared to the NTWing status (i.e., cases above the reference line).

In terms of UTWing vs. NTWing (Fig. 6(c) and 6(d)) and UTWing vs. NUTWing (Fig. 6(e) and 6(f)), most cases are below the reference line, indicating that adopting UTWing or increasing teleworking frequency from NUTWing to UTWing would lead to a VMD reduction for most cases, whether they are travel-stressed or not.

6.3.2. *Aggregate analysis*

In this section and Appendix C, we analyze the weighted average VMD for the nine cells of Fig. 4, from several different perspectives. In Section 6.3.2.1, we present and analyze the weighted average VMD for each observed worker segment, in each (counter) factual status (i.e., comparing each *row* in Fig. 4 or estimates on *the same trajectory line* in Fig. 7). From these we obtain the treatment effects for each worker group, in terms of both weekly miles driven, and miles per teleworking occasion. In Appendix C, we analyze the weighted average VMD across different (tele)worker segments while controlling for the "potential" teleworking status (i.e., comparing each *column* in Fig. 4 or estimates on *the same vertical line* in Fig. 7). In Section 6.3.2.2, we take a closer look at the effects of non-usual teleworking adoption, since that form of teleworking shows greater nuance in its travel impacts than the adoption of usual teleworking does.

6.3.2.1. Aggregate treatment effects. Fig. 7 displays the central tendencies (i.e., the weighted average VMD) for travel-stressed and non-travel-stressed individuals in each possible observed status/potential status combination. We compare VMD across different (tele) working statuses for the same (tele)worker group to analyze the treatment effect. Each trajectory line in Fig. 7, represented by a distinct color, corresponds to a specific (tele)worker group. We compare the estimates on each trajectory to analyze the treatment effect (i.e., the same as comparing across each row in Fig. 4).

We begin by presenting several key treatment effect estimates and comparing them to the conclusions drawn from the full sample model (Table 5). Further details on this analysis can be found in the remainder of this section. As a general observation, we find that the travel-stressed treatment effects are always greater in magnitude than the corresponding non-travel-stressed effects, if both are statistically significant. Moreover, all treatment effects on the treated are negative, which indicates that, on average, teleworking has led to a reduction in VMD for those observed teleworkers.

In terms of the ***treatment effect on the NUTW-treated***, we observe respective reductions of 81.84 and 17.02 miles per week for travel-stressed and non-travel-stressed NUTWers. Notably, the latter result is not statistically significant. In light of these findings, we refine our conclusion from the full sample model to state that adopting non-usual teleworking (less than 3 times/week) leads to a

[10] Be aware that the scatterplots cannot reflect the sample weights. The unweighted share, as portrayed in Fig. 6(a), is 77%.

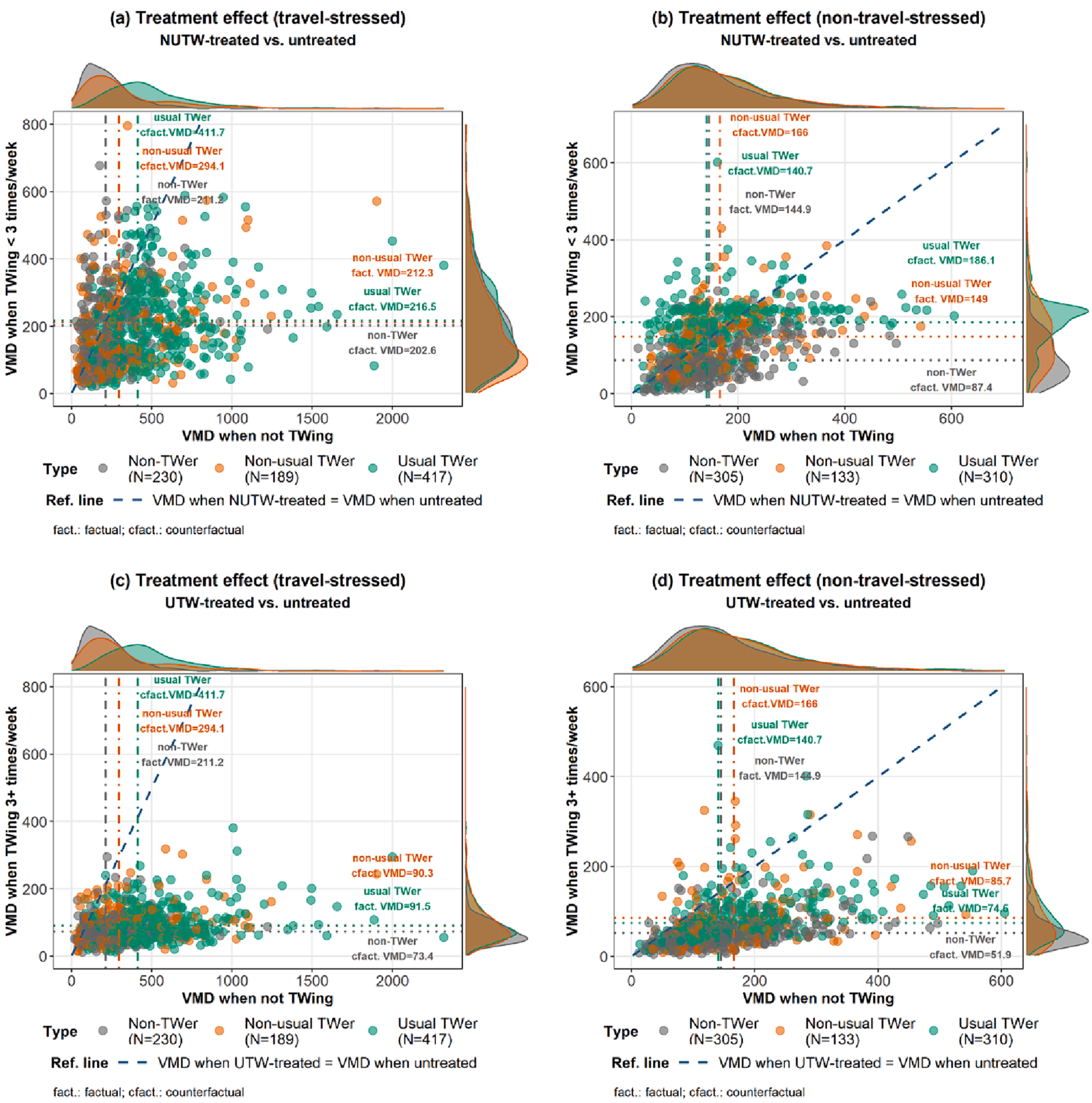

Fig. 6. Disaggregate treatment effects of teleworking adoption/frequency, travel-stressed (a/c/e) vs. non-travel-stressed (b/d/f).

significant reduction in VMD for travel-stressed NUTWers, while the same treatment does not significantly reduce VMD for non-travel-stressed individuals.

Regarding the ***treatment effect on the UTW-treated***, we find that travel-stressed UTWers experience a significant VMD reduction of 320.18 miles per week, while non-travel-stressed participants experience a reduction of 66.18 miles per week, which is also statistically significant. These results are consistent with the conclusion drawn from the full sample model, which finds that adopting usual teleworking (3 + times/week) significantly reduces VMD for UTWers.

When analyzing the ***NUTWing treatment effect on the untreated***, we find that it is insignificant for travel-stressed NTWers (−8.61 mi/week) and significant for non-travel-stressed NTWers (−57.53 mi/week). By contrast, the ***UTWing treatment effect on the untreated*** is significant for both travel-stressed (−137.82 mi/week) and non-travel-stressed (−92.97 mi/week) NTWers, and greater (in magnitude) than the corresponding NUTWing treatment effects. These results suggest that NTWers would experience a reduction in VMD if they were to adopt UTWing, while (on average) only non-travel-stressed NTWers would see a meaningful VMD reduction with NUTWing treatment. This conclusion refines our findings from the full sample model.

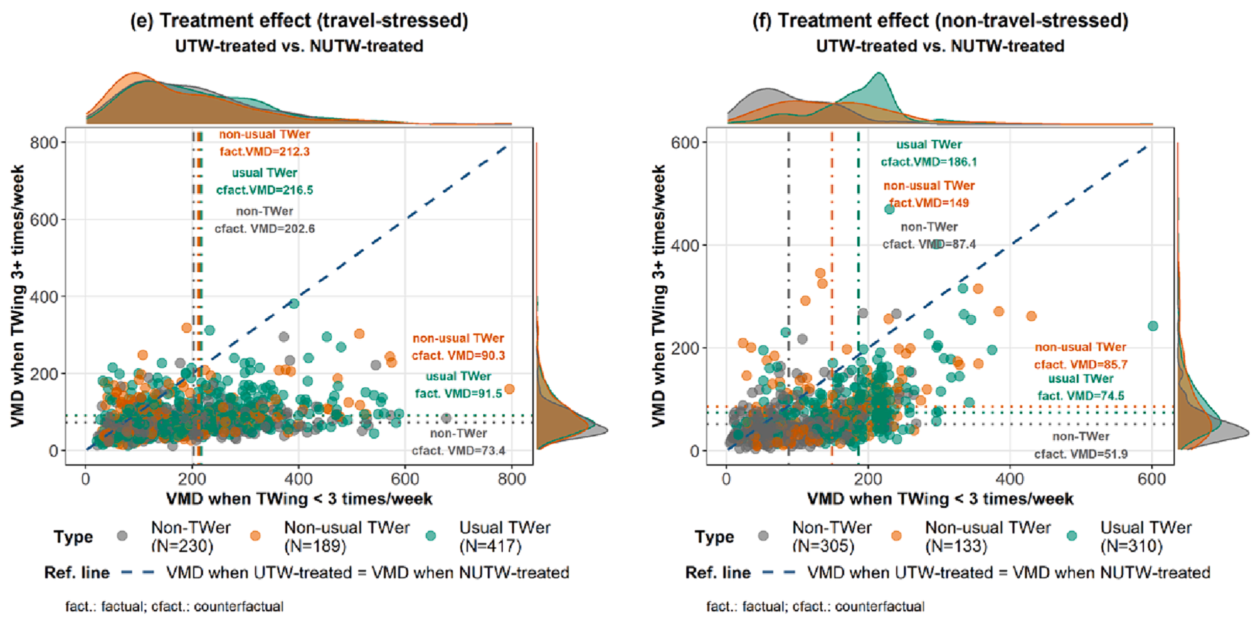

Fig. 6. (*continued*).

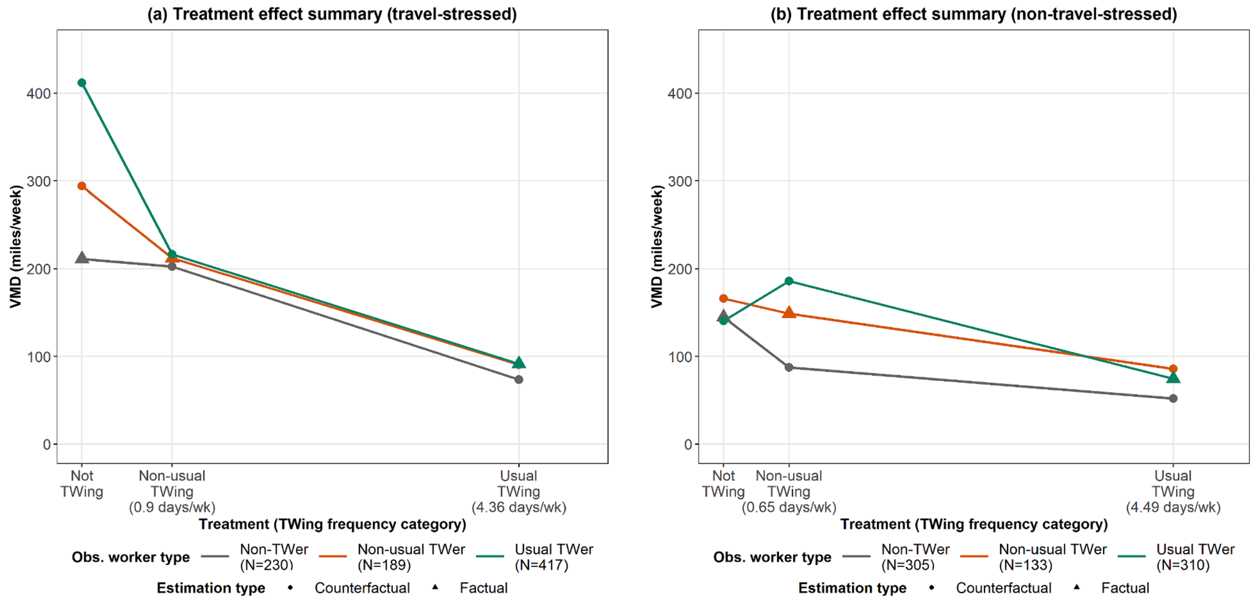

Fig. 7. Aggregate treatment effects of teleworking adoption/frequency, travel-stressed (a) vs. non-travel-stressed (b).

To facilitate more in-depth analysis, Table 6 provides additional information such as empirical average teleworking frequencies, average two-way commute distances, and the unit treatment effect. We present the results for all three models for completeness. In general, the empirical NUTWing frequencies are less than 1 day/week, and the UTWing frequencies are more than 4 days/week. Compared to the non-travel-stressed individuals, travel-stressed people tend to have higher commute distances and larger unit treatment effects. Interestingly, the average teleworking frequency is also notably larger for travel-stressed NUTWers (0.90 days/week) than for their non-travel-stressed counterparts (0.65 days/week), which makes their larger unit treatment effects on the treated (−91 versus −26 miles per teleworking occasion) all the more striking.

As discussed above, adopting UTWing or increasing teleworking frequency from NUTWing to UTWing would result in a significant reduction in average weekly VMD for all three travel-stressed worker groups. From Table 6 we now observe that among all three travel-stressed groups, notably, the unit treatment effect of increasing teleworking frequency from NUTWing to UTWing is comparable to the corresponding two-way commute distance. Regarding the two travel-stressed teleworker groups, the unit treatment effects of adopting

Table 6
Unit treatment effects.

	Observed worker group								
	Full sample			*Travel-stressed*			*Non-travel-stressed*		
	NTWer	NUTWer	UTWer	NTWer	NUTWer	UTWer	NTWer	NUTWer	UTWer
Teleworking frequency, days/week	0.00	0.80	4.42	0.00	0.90	4.36	0.00	0.65	4.49
Two-way commute distance, mi	31.21	30.95	27.98	40.14	36.91	33.27	23.85	22.23	22.12
Unit treatment effect, mi/TWing occasion									
NTWing→NUTWing	*−40.93*	**−9.42**	31.77	*−9.58*	**−91.00**	−217.00	*−87.91*	**−26.00**	69.46
NTWing→UTWing	*−25.19*	−21.75	**−21.01**	*−31.61*	−46.74	**−73.44**	*−20.71*	−17.88	**−14.74**
NUTWing→UTWing	−21.72	−24.47	-32.67	−37.34	−35.24	−36.12	−9.25	−16.49	−29.12

Notes: The unit treatment effect is calculated by dividing the total treatment effect (i.e., VMD per week, Table 5) by the corresponding average teleworking frequency. Bolded numbers are treatment effects on the treated, and italicized numbers are treatment effects on the untreated (see Fig. 4).

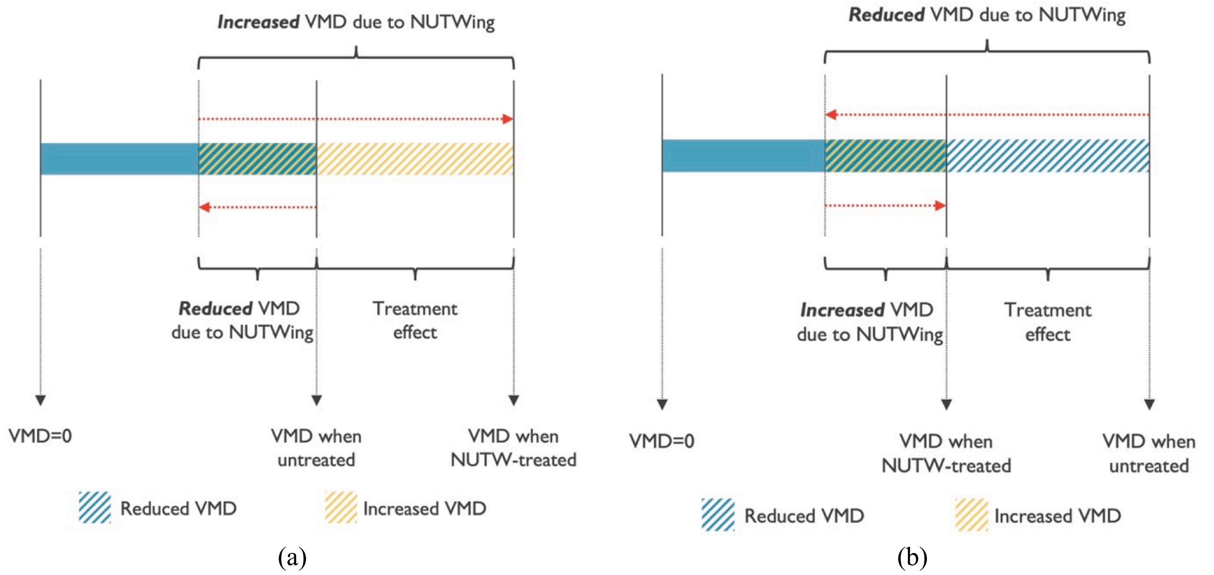

Fig. 8. Illustration of different NUTWing treatment effects (positive (a) or negative (b)) as the net outcome of increases and reductions in VMD.

NUTWing from NTWing are far higher than those of increasing teleworking frequency from NUTWing to UTWing. This indicates a non-linearity of the treatment effect, suggesting a diminishing marginal benefit of additional teleworking. Our hypothesis is that the initial adoption of teleworking may be associated with other behavior changes, such as residential relocation and changes in non-work travel patterns. These changes may amplify the unit treatment effect of NUTWing. For travel-stressed NTWers, by contrast, the adoption of NUTWing makes an insignificant impact (Table 5). This may explain why they have not adopted teleworking, as their VMD would only significantly fall with a very high level of teleworking frequency, which may not be feasible for them.

6.3.2.2. A closer look at (potential) NUTWing adoption. As shown in the NTWing→NUTWing row of Table 5, the adoption of NUTWing makes significant but reversed impacts on non-travel-stressed NTWers and UTWers, while showing no significant effects on NUTWers. Specifically, ***non-travel-stressed UTWers*** would have an *increased* VMD after adopting NUTWing (compared to NTWing), indicating that their non-VMD-reduction teleworking motivations (e.g., family-motivated, flexibility-motivated) induce more non-commute VMD than the VMD saved from commuting (as shown schematically in Fig. 8(a)). ***Non-travel-stressed NUTWers***, by contrast, may experience both VMD reduction and VMD generation due to teleworking, which largely offset each other, resulting in the insignificant net effect. Regarding ***non-travel-stressed NTWers***: on the one hand, they may not have strong enough teleworking motivations that are non-VMD-reduction related, to lead them to adopt NUTWing like their teleworking counterparts. Therefore, even if they were to adopt NUTWing, they may not have much induced (non-commute) VMD. On the other hand, the non-travel-stressed NTWers are also not travel-stressed enough to adopt teleworking for VMD reduction reasons. Still, they could have reduced (commute) VMD if they were to adopt NUTWing. The reduced (commute) VMD resulting from NUTWing would outweigh the NUTWing-induced (non-commute) VMD, as shown schematically in Fig. 8(b). Overall, the non-travel-stressed UTWers are more likely than the other groups to generate new travel demand due to NUTWing, despite not having the highest untreated VMD, as shown in Fig. 7(b).

To summarize, adopting teleworking at a low frequency level (i.e., less than 3 times a week) may not significantly reduce VMD for all worker groups, regardless of whether they are travel-stressed or non-travel-stressed. In contrast, adopting teleworking at a high frequency level or increasing teleworking frequency from low to high levels (i.e., 3 + days/week) leads to a reduction in average VMD for all worker groups, travel stressed and non-travel-stressed alike. It is also worth reiterating that the average *treatment effects on the treated* – i.e. the effects on NUTWers of adopting NUTWing, and the effects on UTWers of adopting UTWing – generally constitute significant reductions in VMD for both travel-stressed and non-travel-stressed TWers (except that the reduction is insignificant for non-travel-stressed NUTWers; Table 5).

7. Conclusion

This study quantifies and compares the "treatment effect" of teleworking adoption/frequency choices on (log-transformed) weekly vehicle-miles driven (VMD) for three different types of (tele)workers (non-teleworkers, non-usual teleworkers, and usual teleworkers), segmented by teleworking-related motivation. Specifically, we applied an ordered probit switching regression (OPSR) model to identify factors that influence the teleworking adoption/frequency choice and weekly vehicle-miles driven while accounting for self-selection biases. We developed three models with survey data collected in Spring 2021: a full-sample model and two subsample models, for travel-stressed workers and non-travel-stressed workers, respectively. We separated travel-stressed and non-travel-stressed workers because reducing VMD may not be the most-desired benefit for all types of teleworkers when they adopt teleworking, which may thus confound the net impact of teleworking on VMD.

The OPSR model contains a selection model that predicts the teleworking decision, and a set of outcome models that predict log-transformed VMD for each (tele)worker group. The self-selection bias is captured by the correlation between the error terms of the selection and outcome models. The full sample model identifies self-selection biases for both teleworker groups. In particular, for both UTWers and NUTWers, the unobserved attributes that increase teleworking propensity also tend to increase their VMD.[11] However, such an association between selection and outcome is not exhibited to a significant degree for NTWers. When we separate the full sample into travel-stressed and non-travel-stressed, we find that unobserved attributes that increase the teleworking propensity will induce higher VMD for both travel-stressed and non-travel-stressed UTWers. However, the correlation is stronger for the non-travel-stressed group, suggesting that unobserved attributes have a greater impact on increasing VMD for non-travel-stressed UTWers than for their travel-stressed counterparts. Beyond the two UTWer groups, we also find significant positive correlations for non-travel-stressed NUTWers and travel-stressed NTWers. Similarly, unobserved attributes that increase (or decrease) their teleworking propensity will induce higher (or lower) VMD. Without distinguishing the travel-stressed and non-travel-stressed subsamples, the selection bias for travel-stressed NTWers was masked in the full sample model.

After calibrating the OPSR models, we provide technical details on calculating factual and counterfactual model estimates, treatment effects, and VMD estimates that have been back-log-transformed. In terms of the treatment effect, the three models have generally consistent results, while travel-stressed and non-travel-stressed models provide more nuanced insights. In terms of the ***treatment effect on the NUTW-treated***, we conclude that adopting non-usual teleworking (less than 3 times/week) leads to a significant reduction in VMD for travel-stressed NUTWers (−81.84 mi/week), while the same treatment does not significantly reduce VMD for non-travel-stressed individuals. Regarding the ***treatment effect on the UTW-treated***, we find that adopting usual teleworking (3 + times/week) significantly reduces VMD for UTWers, regardless of the travel stress level (−320.18 mi/week if travel-stressed, −66.18 mi/week if non-travel-stressed). When analyzing the ***UTWing and NUTWing treatment effects on the untreated***, the results suggest that NTWers would experience a reduction in VMD if they were to adopt UTWing (−137.82 mi/week if travel-stressed, −92.97 mi/week if non-travel-stressed), while only non-travel-stressed NTWers would see a significant VMD reduction with NUTWing treatment (−57.53 mi/week).

We calculate unit treatment effects per teleworking occasion and compare them with the corresponding two-way commute distance for reality checks. For travel-stressed individuals, the unit treatment effect of increasing teleworking frequency from NUTWing to UTWing is comparable to the corresponding two-way commute distance. However, for the two teleworker groups, the unit treatment effects of adopting NUTWing are far higher than they are for increasing teleworking frequency from NUTWing to UTWing. This indicates a non-linearity of the treatment effect, suggesting a diminishing marginal benefit of additional teleworking. Moreover, travel-stressed people tend to have higher commute distances and unit treatment effects than the non-travel-stressed individuals. For the latter group, the adoption of NUTWing significantly *reduces* VMD for NTWers, significantly *increases* VMD for UTWers, and shows *no significant effects* on NUTWers.

Overall, adopting teleworking at a low frequency level (i.e., less than 3 times a week) may not significantly reduce VMD for all worker groups; sometimes it may even induce a higher VMD. For non-travel-stressed UTWers, the average NUTW-treated VMD is even higher than their untreated VMD. Noting that NUTWers are likely to end up being the larger group of teleworkers (Wang et al., 2022b), it suggests that we will see increases in the number of teleworkers who generate more vehicle travel than they save. However, adopting

[11] This is partially consistent with our presupposition that teleworkers tend to travel more than observationally equivalent others even before they start teleworking – i.e. it supports the presupposition with respect to *unobserved* influences on driving. But as Fig. 5(d) and Table C1 show, when accounting for the average impacts of both observed and unobserved influences together, the (counter)factual mean VMD in the untreated (NTWing) condition does not differ significantly by worker group (although it should be noted that those means do not guarantee equivalence across group with respect to the observed explanatory variables). On the other hand, the corresponding (counter)factual VMD means for the untreated (NTWing) condition in Fig. 7(a) show that the presupposition that teleworkers tend to travel more than observationally equivalent others before adopting teleworking *is* supported specifically for *travel-stressed* teleworkers.

teleworking at a high frequency level or increasing teleworking frequency from low to high levels (i.e., 3 + days/week) always leads to a reduction in VMD, on average, for all worker groups, whether travel stressed or non-travel-stressed.

The study has several limitations, which could be improved in future studies. First, the variables utilized in the analysis are derived from self-reported survey data, including the measure of travel demand, which is represented by weekly VMD. It is important to acknowledge that self-reported measures inherently suffer from memory biases and differing degrees of skill at estimating such quantities. To address these issues, we recommend incorporating travel diary data combined with odometer readings, preferably through passive data collection methods like GPS tracking, to mitigate inaccuracies in self-reported VMD.

Moreover, considering the efficiency of the estimated coefficients (Winship and Radbill, 1994), we made a deliberate choice not to incorporate sample weights into our *modeling* process (although all *descriptive* results, including treatment effect estimates, are based on the weighted sample). Instead, we incorporated an array of socioeconomic and demographic variables as control factors. Future researchers might find it beneficial to incorporate sample weights in their modeling procedures, especially when the sample is small in size and is unrepresentative of the target population. To derive the correct coefficient estimator variance when sample weights are included in the modeling procedure, one may either formulate a mathematical representation that accounts for sample weights or adopt bootstrapping techniques for estimation.

Lastly, we established the threshold of 3 times/week to classify teleworkers into two categories: non-usual teleworkers (those who work from home less than 3 times/week) and usual teleworkers (those who work from home 3 or more times/week). This segmentation is based on empirical experience and the U.S. Census (or American Community Survey, ACS) question wording. The threshold also aligns with recent policy changes in many companies, where employees are required to work on site for at least 3 days per week (which would unambiguously classify their teleworkers as non-usual, according to our definition). The empirical results of our study demonstrate distinct teleworking outcomes between the two groups in terms of weekly VMD, supporting the division between usual and non-usual teleworkers. However, future studies could explore alternative and additional split points for segmenting teleworkers, to identify the optimal classification that represents teleworkers with more homogeneous travel patterns and/or better reflects the evolving empirical share of teleworkers as teleworking policies continue to evolve.

To summarize, this study provides a unique perspective on understanding the impact of teleworking on VMD. The methodology used here, including the visualizations of factual and counterfactual effects, can also be used to shed light on related topics, such as examining the impact of residential location on how much people travel after accounting for self-selection bias (also see Kim and Mokhtarian, 2023a,b). Indeed, it can be applied to numerous other contexts in which selectivity bias is an issue, making the study relevant to a wide range of research domains.

CRediT authorship contribution statement

Xinyi Wang: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **Patricia L. Mokhtarian:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The study is funded by the private transport company Cintra (https://www.ferrovial.com/en-us/cintra/) as an independent research investigation, as well as by the Center for Teaching Old Models New Tricks (TOMNET, which receives funding from the U.S. Department of Transportation under Grant No. 69A3551747116). Any opinions, findings, and conclusions or recommendations expressed in this study are those of the authors and do not necessarily reflect the views of the sponsor organization. The authors are grateful for the constructive comments of the anonymous reviewers, which have improved the paper.

Appendix A. Derivation of the log-likelihood function

Assuming independence of observations, the joint density or likelihood, $\mathscr{L}$, of the sample may be written as:

$$\mathscr{L} = \prod_{\{j=1\}} P(Z=1)\, f(y_1|Z=1) \cdots \prod_{\{j=J\}} P(Z=J)\, f\left(y_J|Z=J\right).$$

Given that the two components in the likelihood function are

$$P(Z=j) = P(\kappa_{j-1} - \boldsymbol{W\gamma} \ < \ \varepsilon \leq \kappa_j - \boldsymbol{W\gamma}) = \int_{\kappa_{j-1}-\boldsymbol{W\gamma}}^{\kappa_j-\boldsymbol{W\gamma}} \phi(\varepsilon)d\varepsilon$$

and

$$f\left(y_j|Z=j\right) = f\left(y_j\middle|\kappa_{j-1} - \boldsymbol{W\gamma} \ < \ \varepsilon \leq \kappa_j - \boldsymbol{W\gamma}\right) = \frac{\int_{\kappa_{j-1}-\boldsymbol{W\gamma}}^{\kappa_j-\boldsymbol{W\gamma}} f\left(y_j|\varepsilon\right)\phi(\varepsilon)d\varepsilon}{\int_{\kappa_{j-1}-\boldsymbol{W\gamma}}^{\kappa_j-\boldsymbol{W\gamma}} \phi(\varepsilon)d\varepsilon} = \frac{\int_{\kappa_{j-1}-\boldsymbol{W\gamma}}^{\kappa_j-\boldsymbol{W\gamma}} f\left(y_j,\varepsilon\right)d\varepsilon}{\int_{\kappa_{j-1}-\boldsymbol{W\gamma}}^{\kappa_j-\boldsymbol{W\gamma}} \phi(\varepsilon)d\varepsilon},$$

the likelihood can be expressed as (eventually reversing the direction of conditionality):

$$\begin{aligned}\mathscr{L} &= \prod_{\{j=1\}}\int_{-\infty}^{\kappa_1-\boldsymbol{W\gamma}} f(y_1,\varepsilon)d\varepsilon \prod_{\{j=2\}}\int_{\kappa_1-\boldsymbol{W\gamma}}^{\kappa_2-\boldsymbol{W\gamma}} f(y_2,\varepsilon)d\varepsilon\cdots\prod_{\{j=J\}}\int_{\kappa_{J-1}-\boldsymbol{W\gamma}}^{\infty} f(y_J,\varepsilon)d\varepsilon \\ &= \prod_{j'=1}^{J}\prod_{\{j=j'\}}\int_{\kappa_{j-1}-\boldsymbol{W\gamma}}^{\kappa_j-\boldsymbol{W\gamma}} f\left(y_j,\varepsilon\right)d\varepsilon \\ &= \prod_{j'=1}^{J}\prod_{\{j=j'\}}\int_{\kappa_{j-1}-\boldsymbol{W\gamma}}^{\kappa_j-\boldsymbol{W\gamma}} f\left(\varepsilon|y_j\right)f\left(y_j\right)d\varepsilon \\ &= \prod_{j'=1}^{J}\prod_{\{j=j'\}} f\left(y_j\right)\int_{\kappa_{j-1}-\boldsymbol{W\gamma}}^{\kappa_j-\boldsymbol{W\gamma}} f\left(\varepsilon\middle|y_j\right)d\varepsilon.\end{aligned}$$

The conditional distribution of ε given y_j is univariate normal:

$$\varepsilon|y_j \sim N\left(\frac{\rho_j}{\sigma_j}\left(y_j - \boldsymbol{X}_j\boldsymbol{\beta}_j\right), \left(1-\rho_j^2\right)\right),$$

since

$$\varepsilon \sim \ N(0,1),$$

$$y_j = \boldsymbol{X}_j\boldsymbol{\beta}_j + \eta_j \sim N\left(\boldsymbol{X}_j\boldsymbol{\beta}_j, \sigma_j^2\right),$$

$$\text{and}\begin{pmatrix}\varepsilon \\ y_j\end{pmatrix} \sim N\left(\begin{bmatrix}0 \\ \boldsymbol{X}_j\boldsymbol{\beta}_j\end{bmatrix}, \begin{bmatrix}1 & \rho_j\sigma_j \\ \rho_j\sigma_j & \sigma_j^2\end{bmatrix}\right).$$

Accordingly, the likelihood and log-likelihood functions, respectively, are

$$\mathscr{L} = \prod_{j'=1}^{J}\prod_{\{j=j'\}}\left\{\frac{1}{\sigma_j}\phi\left(\frac{y_j - \boldsymbol{X}_j\boldsymbol{\beta}_j}{\sigma_j}\right)\left[\Phi\left(\frac{\sigma_j\left(\kappa_j - \boldsymbol{W\gamma}\right) - \rho_j\left(y_j - \boldsymbol{X}_j\boldsymbol{\beta}_j\right)}{\sigma_j\sqrt{1-\rho_j^2}}\right) - \Phi\left(\frac{\sigma_j\left(\kappa_{j-1} - \boldsymbol{W\gamma}\right) - \rho_j\left(y_j - \boldsymbol{X}_j\boldsymbol{\beta}_j\right)}{\sigma_j\sqrt{1-\rho_j^2}}\right)\right]\right\}, \text{ and}$$

$$LL = \sum_{j'=1}^{J}\sum_{\{j\}}\left\{ln\left[\frac{1}{\sigma_j}\phi\left(\frac{y_j - \boldsymbol{X}_j\boldsymbol{\beta}_j}{\sigma_j}\right)\right] + ln\left[\Phi\left(\frac{\sigma_j\left(\kappa_j - \boldsymbol{W\gamma}\right) - \rho_j\left(y_j - \boldsymbol{X}_j\boldsymbol{\beta}_j\right)}{\sigma_j\sqrt{1-\rho_j^2}}\right) - \Phi\left(\frac{\sigma_j\left(\kappa_{j-1} - \boldsymbol{W\gamma}\right) - \rho_j\left(y_j - \boldsymbol{X}_j\boldsymbol{\beta}_j\right)}{\sigma_j\sqrt{1-\rho_j^2}}\right)\right]\right\}.$$

Appendix B. Derivation of the conditional expectation

$$
\begin{aligned}
\mathbb{E}\left[\eta_j \middle| \kappa_{j-1} - \boldsymbol{W\gamma} < \varepsilon < \kappa_j - \boldsymbol{W\gamma}\right] &= \frac{\int_{\kappa_{j-1}-\boldsymbol{W\gamma}}^{\kappa_j-\boldsymbol{W\gamma}} \mathbb{E}\left[\eta_j|\varepsilon = \varepsilon_0\right]\phi(\varepsilon_0)d\varepsilon_0}{\Pr\left(\kappa_{j-1} - \boldsymbol{W\gamma} < \varepsilon < \kappa_j - \boldsymbol{W\gamma}\right)} \\
&= \frac{1}{\Phi\left(\kappa_j - \boldsymbol{W\gamma}\right) - \Phi\left(\kappa_{j-1} - \boldsymbol{W\gamma}\right)} \int_{\kappa_{j-1}-\boldsymbol{W\gamma}}^{\kappa_j-\boldsymbol{W\gamma}} \mathbb{E}\left[\eta_j|\varepsilon = \varepsilon_0\right]\phi(\varepsilon_0)d\varepsilon_0 \\
&= \frac{1}{\Phi\left(\kappa_j - \boldsymbol{W\gamma}\right) - \Phi\left(\kappa_{j-1} - \boldsymbol{W\gamma}\right)} \int_{\kappa_{j-1}-\boldsymbol{W\gamma}}^{\kappa_j-\boldsymbol{W\gamma}} \rho_j\sigma_j\varepsilon_0 \frac{1}{\sqrt{2\pi}}e^{-\frac{\varepsilon_0^2}{2}}d\varepsilon_0 \\
&= \frac{\rho_j\sigma_j}{\Phi\left(\kappa_j - \boldsymbol{W\gamma}\right) - \Phi\left(\kappa_{j-1} - \boldsymbol{W\gamma}\right)} \left(-\frac{1}{\sqrt{2\pi}}e^{-\frac{\varepsilon_0^2}{2}}d\varepsilon_0\right)\Bigg|_{\kappa_{j-1} - \boldsymbol{W\gamma}}^{\kappa_j - \boldsymbol{W\gamma}} \\
&= -\rho_j\sigma_j\frac{\phi\left(\kappa_j - \boldsymbol{W\gamma}\right) - \phi\left(\kappa_{j-1} - \boldsymbol{W\gamma}\right)}{\Phi\left(\kappa_j - \boldsymbol{W\gamma}\right) - \Phi\left(\kappa_{j-1} - \boldsymbol{W\gamma}\right)},
\end{aligned}
$$

where

$$
\begin{pmatrix} \varepsilon \\ \eta_1 \\ \vdots \\ \eta_j \\ \vdots \\ \eta_J \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_1\sigma_1 & \cdots & \rho_j\sigma_j & \cdots & \rho_J\sigma_J \\ \rho_1\sigma_1 & \sigma_1^2 & & & & \\ \vdots & & \ddots & & & \\ \rho_j\sigma_j & & & \sigma_j^2 & & \\ \vdots & & & & \ddots & \\ \rho_J\sigma_J & & & & & \sigma_J^2 \end{pmatrix}\right],
$$

$\mathbb{E}\left[\eta_j|\varepsilon = \varepsilon_0\right] = \rho_j\sigma_j\varepsilon_0$, and

$$\phi(\mu) = \frac{1}{\sqrt{2\pi}}e^{-\frac{\mu^2}{2}}.$$

Therefore,

$$\mathbb{E}\left[y_j|Z = j\right] = \mathbb{E}\left[\boldsymbol{X}_j\boldsymbol{\beta}_j + \eta_j|Z = j\right] = \boldsymbol{X}_j\boldsymbol{\beta}_j + \mathbb{E}\left[\eta_j\middle|\kappa_{j-1} - \boldsymbol{W\gamma} < \varepsilon \leq \kappa_j - \boldsymbol{W\gamma}\right]$$

$$= \boldsymbol{X}_j\boldsymbol{\beta}_j - \rho_j\sigma_j\frac{\phi\left(\kappa_j - \boldsymbol{W\gamma}\right) - \phi\left(\kappa_{j-1} - \boldsymbol{W\gamma}\right)}{\Phi\left(\kappa_j - \boldsymbol{W\gamma}\right) - \Phi\left(\kappa_{j-1} - \boldsymbol{W\gamma}\right)}.$$

Appendix C. VMD for different worker groups having the same “potential” teleworking status

In this appendix, we compare VMD across different (tele)worker segments while controlling the “potential” teleworking status (i.e., comparing each column in Fig. 4 or estimates on the same vertical lines in Fig. 7). In addition to the visualization in Fig. 7, we present the weighted two-sample *t*-test results in Table C1, testing the equality of VMD means between pairs of *observed* groups in the same *potential* teleworking status. For completeness, we present the results for all three models: full sample, travel-stressed, and non-travel-stressed.

Focusing on the NTWing status first, the untreated VMDs are significantly different across the three travel-stressed worker groups: UTWer > NUTWer > NTWer. Travel-stressed UTWers have the highest untreated (i.e., if not adopting teleworking) travel demand, and the order is consistent with their current teleworking frequency: the higher the travel demand (in a non-teleworking state), the higher the teleworking frequency. When it comes to the non-travel-stressed group, the NUTWers have a (significantly) higher untreated travel demand than UTWers and NTWers do (165.98 vs. 140.66 and 144.90 miles/week, on average). However, upon combining the travel-stressed and non-travel-stressed subsamples (Fig. 5(d) and the full sample NTWing column of Table C1), we find that the mean NTWing VMDs of the three groups of workers are similar, and there is no significant difference among them.

If all three groups of travel-stressed (tele)workers engage in NUTWing, there is no significant difference in their mean NUTW-treated VMDs, despite their significantly different untreated VMDs. In contrast, the mean NUTW-treated VMDs of non-travel-stressed individuals vary significantly across the three groups of (tele)workers, with the UTWers exhibiting the highest VMD, followed by NUTWers and NTWers. The same pattern is observed in the full sample model (Fig. 5(d) and the full sample NUTWing column of Table C1).

Lastly, the mean UTW-treated VMDs show similar patterns across all three models, with UTWers and NUTWers displaying similar UTW-treated VMDs, which are significantly (even if only modestly) higher than that of NTWers.

To summarize, our findings suggest that observed NTWers consistently exhibit similar or lower mean VMDs than their teleworker counterparts, when the treatment is held constant. In the case of *travel-stressed teleworkers*, UTWers exhibit a substantially higher mean untreated VMD than NUTWers, which is indicative of their high travel demand prior to adopting teleworking. However, after applying and controlling the treatment, the two groups of teleworkers exhibit comparable mean VMDs, regardless of the level of treatment. When it comes to *non-travel-stressed teleworkers*, NUTWers have slightly (but significantly) higher mean VMDs than UTWers ($p = 0.024$) in the absence of treatment. However, after applying the NUTWing treatment to both groups, UTWers tend to exhibit higher VMDs than NUTWers, and the difference between the two groups becomes insignificant as the treatment gets stronger (i.e., with an increased teleworking frequency).

Table C1
Weighted two-sample *t*-test p-values testing the equality of VMD means between observed (tele)worker groups with the same (potential) treatment status.

	Treatment status								
	Full sample			*Travel-stressed*			*Non-travel-stressed*		
Observed worker group	NTWing	NUTWing	UTWing	NTWing	NUTWing	UTWing	NTWing	NUTWing	UTWing
NTWer vs. NUTWer	0.212	**0.000**	**0.000**	**0.000**	0.480	**0.000**	**0.024**	**0.000**	**0.000**
NTWer vs. UTWer	0.592	**0.000**	**0.000**	**0.000**	0.182	**0.000**	0.480	**0.000**	**0.000**
NUTWer vs. UTWer	0.434	**0.000**	0.564	**0.000**	0.688	0.791	**0.002**	**0.000**	0.066

Note: The numbers in this table represent the p-value of the corresponding weighted two-sample *t*-test indicated in the row header. Bolded numbers are statistically significant at the 0.05 level. For example, the 0.592 indicates that NTWers and UTWers have similar (not significantly different) VMDs if they do not telework (based on the full sample model).

References

Amemiya, T., 1985. Advanced Econometrics. Harvard University Press.

Burrows, M., Burd, C., Mckenzie B., 2023. Home-Based Workers and the COVID-19 Pandemic (American Community Survey Reports), U.S. Census Bureau. Available at: https://www.census.gov/library/publications/2023/acs/acs-52.html.

Cameron, A.C., Trivedi, P.K., 2005. Microeconometrics: Methods and Applications. Cambridge University Press.

Chakrabarti, S., 2018. Does telecommuting promote sustainable travel and physical activity? J. Transp. Health 9, 19–33.

Chiburis, R., Lokshin, M., 2007. Maximum likelihood and two-step estimation of an ordered-probit selection model. Stata J. 7 (2), 167–182.

Choo, S., Mokhtarian, P.L., Salomon, I., 2005. Does telecommuting reduce vehicle-miles traveled? An aggregate time series analysis for the U.S. Transportation 32 (1), 37–64.

Currie, G., Jain, T., Aston, L., 2021. Evidence of a post-COVID change in travel behaviour–Self-reported expectations of commuting in Melbourne. Transp. Res. A 153, 218–234.

de Abreu e Silva, J., Melo, P.C., 2018. Does home-based telework reduce household total travel? A path analysis using single and two worker British households. J. Transp. Geogr., 73, 148-162.

de Abreu e Silva, J., 2022. Residential preferences, telework perceptions, and the intention to telework: insights from the Lisbon Metropolitan Area during the COVID-19 pandemic. Reg. Sci. Pol. Pract. 14, 142-161.

de Vos, D., Meijers, E., van Ham, M., 2018. Working from home and the willingness to accept a longer commute. Ann. Reg. Sci. 61 (2), 375–398.

Elldér, E., 2020. Telework and daily travel: New evidence from Sweden. J. Transp. Geogr. 86, 102777.

Erhardt, G.D., Hoque, J.M., Goyal, V., Berrebi, S., Brakewood, C., Watkins, K.E., 2022. Why has public transit ridership declined in the United States? Transp. Res. A 161, 68–87.

He, S.Y., Hu, L., 2015. Telecommuting, income, and out-of-home activities. Travel Behav. Soc. 2 (3), 131–147.

Heckman, J.J., 1979. Sample selection bias as a specification error. Econometrica 47 (1), 153–161.

Heckman, J., Tobias, J.L., Vytlacil, E., 2001. Four parameters of interest in the evaluation of social programs. South. Econ. J. 68 (2), 211–223.

Jimenez, E., Kugler, B., 1987. The earnings impact of training duration in a developing country: an ordered probit selection model of Colombia's Servicio Nacional de Aprendizaje (SENA). J. Hum. Resour. 22 (2), 228–247.

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

Kim, S.H., Mokhtarian, P.L., 2023a. Comparisons of observed and unobserved parameter heterogeneity in modeling vehicle-miles driven. Transp. Res. A 172, 103614.

Kim, S.H., Mokhtarian, P.L., 2023b. A note on the sample selection (switching regression) model and treatment effects for a log-transformed outcome variable, in the context of residential self-selection. Transportation. https://doi.org/10.1007/s11116-023-10384-2.

Mishra, G.S., 2017. Estimating the Travel Behavior Effects of Technological Innovations from Cross-sectional Observed Data: Applications to Carsharing and Telecommuting. PhD dissertation, Transportation Technology and Policy program, University of California, Davis. Available at https://www.proquest.com/docview/1947638403.
Mokhtarian, P.L., Handy, S.L., Salomon, I., 1995. Methodological issues in the estimation of the travel, energy, and air quality impacts of telecommuting. Transp. Res. A 29 (4), 283–302.
Nakosteen, R.A., Zimmer, M., 1980. Migration and income: The question of self-selection. South. Econ. J. 46 (3), 840–851.
Obeid, H., Anderson, M.L., Bouzaghrane, M.A., Walker, J., 2024. Does telecommuting reduce trip-making? Evidence from a US panel during the COVID-19 pandemic. Transportation Research Part A: Policy and Practice 180, 103972.
Ory, D.T., Mokhtarian, P.L., 2006. Which came first, the telecommuting or the residential relocation? An empirical analysis of causality. Urban Geogr. 27 (7), 590–609.
Ravalet, E., Rérat, P., 2019. Teleworking: decreasing mobility or increasing tolerance of commuting distances? Built Environ. 45 (4), 582–602.
Roodman, D., 2011. Fitting fully observed recursive mixed-process models with cmp. Stata J. 11 (2), 159–206.
Stiles, J., Smart, M.J., 2021. Working at home and elsewhere: daily work location, telework, and travel among United States knowledge workers. Transportation 48 (5), 2461–2491.
Su, R., McBride, E.C., Goulias, K.G., 2021. Unveiling daily activity pattern differences between telecommuters and commuters using human mobility motifs and sequence analysis. Transp. Res. A 147, 106–132.
Thompson, R.J., Payne, S.C., Alexander, A.L., Gaskins, V.A., Henning, J.B., 2021. A taxonomy of employee motives for telework. Occupat. Health Sci. 1–32.
Wang, X., 2023. The Teleworking Treatment Effect on Travel Behavior Changes: An In-Depth Exploration of Endogenous Switching Regression Models (Doctoral dissertation). Georgia Institute of Technology, Atlanta, Georgia. Available at https://hdl.handle.net/1853/72741.
Wang, X., Kim, S.H., Mokhtarian, P.L., 2023. Teleworking behavior pre-, during, and expected post-COVID: Identification and empirical description of trajectory types. Travel Behav. Soc. 33, 100628.
Wang, X., Kim, S.H., Mokhtarian, P.L., 2024. Identifying teleworking-related motives and comparing telework frequency expectations in the post-pandemic world: A latent class choice modeling approach. Transp. Res. Part A. https://doi.org/10.1016/j.tra.2024.104070.
Winship, C., Radbill, L., 1994. Sampling weights and regression analysis. Sociol. Methods Res. 23, 230–257.
Yen, S.T., Rosinski, J., 2008. On the marginal effects of variables in the log-transformed sample selection models. Econ. Lett. 100 (1), 4–8.
Yum, S., 2021. Differences between telecommuters and commuters: the case of the Twin Cities metropolitan area. Transp. Plan. Technol. 44 (3), 303–318.
Zhu, P., 2012. Are telecommuting and personal travel complements or substitutes? Ann. Reg. Sci. 48 (2), 619–639.
Zhu, P., 2013. Telecommuting, household commute and location choice. Urban Stud. 50 (12), 2441–2459.
Zhu, P., Mason, S.G., 2014. The impact of telecommuting on personal vehicle usage and environmental sustainability. Int. J. Environ. Sci. Technol. 11 (8), 2185–2200.