The Insidious Problem of Selection Biases:

How to Recognize Them, and What to Do about Them

Patricia L. Mokhtarian
Georgia Institute of Technology
patmokh@gatech.edu
MIT Mobility Forum
April 12, 2024

Georgia Tech

A loose outline

- Introduction, basic concepts
- The selection bias problem
- The endogenous switching regression solution
- An application: the impact of teleworking on vehicle-miles driven
- Summary and conclusions

Georgia Tech

Introduction: the scenario

- There is a ***discrete event*** of interest
  - "Treatment", "experiment", "intervention"
  - **Extreme event** (earthquake, bridge collapse, etc.)
  - "New" **technologies** (e.g. medical, agricultural, telecommunications)
  - **Policies** (e.g. promoting transportation safety, or sustainable transportation)
  - Individuals are "untreated" or "treated" (there can be multiple categorical/ordered treatments)
  - Individuals may actively choose, or passively experience, the treatment (or not)
- We want to estimate the effect of that event on an (individually-experienced, then aggregated) ***outcome*** of interest, i.e., the ***"treatment effect"***
  - **Example outcomes:** Blood pressure, crop yield, vehicle-miles driven
  - Assume we observe the outcome whether a person is treated or untreated

Some transportation examples

Treatment

- Wearing a helmet while (motor)cycling
- Buying a fuel-efficient car
- Living in a transit-friendly, mixed-use neighborhood
- Adopting teleworking (TWing)

Outcome

- Injury severity
- CO_2 emissions
- Vehicle-miles driven (VMD)

What "treatment effects" (TEs) might we be interested in?

- The ***treatment effect on the treated (TT)***
  - Difference between outcome if treated versus not, ***for those who have been treated***
- The ***treatment effect on the untreated (TUT)***
  - Difference between outcome if treated versus not, ***for those who have NOT been treated*** (i.e. *how would the outcome have been different if they **had** been treated?*)
- The ***average treatment effect (ATE)***
  - The appropriately-weighted ***average of TT and TUT***
  - The average impact of treatment for a person ***randomly-selected from the population***, which contains both treated and untreated individuals
- It's not always obvious which TE is most relevant, nor even which the authors intend!

Georgia Tech

Hm – how hard can it be to estimate a TE?

- ***The problem:*** with ***cross-sectional data***, we only observe people as treated or untreated (their ***factual*** state at the time of measurement) – we don't observe them in the opposite condition (their ***counterfactual*** state)
- Well, fine, but can't we just model an outcome Y as, say,
  - $Y = \boldsymbol{X\beta} + \delta T + \eta$, where (throughout, individual-level subscripts are omitted for simplicity)
  - $\boldsymbol{X}$ = explanatory variables; T = treatment dummy; $\{\boldsymbol{\beta}, \delta\}$ = coefficients; $\eta \sim N\ (0, \sigma^2)$?
  - $\hat{\delta}$ would be the estimated TE, wouldn't it?
    - But *which* TE? In essence, this approach would force TT = TUT = ATE
- OK, then how about separate equations for the treated and untreated?
  - $Y_T = \boldsymbol{X\beta_T} + \eta_T$ and $Y_U = \boldsymbol{X\beta_U} + \eta_U$
  - ATE = avg over the entire ***(representative)*** sample of $(\hat{Y}_T - \hat{Y}_U) = \boldsymbol{X}(\hat{\boldsymbol{\beta}}_{\boldsymbol{T}} - \hat{\boldsymbol{\beta}}_{\boldsymbol{U}})$
  - TT = avg over the treated of $(\hat{Y}_T - \hat{Y}_U)$ (where $\hat{Y}_U$ applies $\hat{\boldsymbol{\beta}}_{\boldsymbol{U}}$ to the **treated** people's Xs)
  - TUT = avg over the untreated of $(\hat{Y}_T - \hat{Y}_U)$ (where $\hat{Y}_T$ applies $\hat{\boldsymbol{\beta}}_{\boldsymbol{T}}$ to the **untreated** people's Xs)

This would be fine, *IF*…

- The treatment were ***randomly assigned*** across the population
  - Then the treated and untreated cases in a (large enough, and representative) sample would ***each*** be random samples from the population of outcomes
- But what if, instead, the treated group differs from the untreated group in ways that are relevant to the outcome? I.e. we have a ***selection bias***
  - Those receiving the new medical treatment may be sicker
  - Those trying the new fertilizer may have also adopted other innovations
  - Those wearing a helmet (before it became mandatory) may be safer (motor)cyclists (or, they may compensate with *more* risky behaviors)
  - Those buying a fuel-efficient car may want to travel more
- This is likely to happen when the individual ***voluntarily adopts*** the treatment!

Georgia Tech

What's the problem with that (in English)?

- Before treatment, if the *to-be-treated* people ***already differ*** from the *to-remain-**un**treated* people, then how can we *separate the effect of the **treatment** from the effect of those **pre-existing differences***?
  - We can account for the effect of differences in ***observed variables*** (the Xs) – those are controlled for in the model
  - But the naïve model shown previously ***does not account for the effect of differences in unobserved variables*** (the ηs)
- We should not expect the treatment to have the same effect if imposed on, or adopted by, a currently untreated person
- Even the estimated ***treatment effect on the treated (TT)*** will be biased!
- That's because our estimated TT confounds the ***true effect of the treatment*** with the ***effects due to pre-existing*** *(unobserved)* ***inclinations/traits*** that may cause a portion of the observed effect to occur *even without the treatment*
- In other words, our ***estimates of β_T and β_U will be biased***, by virtue of being based on non-random samples

What's the problem (in math)?

- We have two components to the model system (the **"endogenous switching regression model", or ESRM** , AKA Roy's model, mover-stayer model, Tobit type 5 model, etc.)
- The **selection model** (binary probit) governs selection into treatment:
  - $Z^* = \boldsymbol{W\gamma} + \varepsilon$; **treated** if $\boldsymbol{Z^* > 0}$, **untreated** if $\boldsymbol{Z^* \leq 0}$
  - $\boldsymbol{W}$ = explanatory variables, $\boldsymbol{\gamma}$ = coefficients, ε = error term
- Two **outcome models** (linear regressions):
  - If **treated:** $Y_T = \boldsymbol{X\beta_T} + \eta_T$
  - If **untreated:** $Y_U = \boldsymbol{X\beta_U} + \eta_U$
  - $\boldsymbol{X}$ = explanatory variables, $\boldsymbol{\beta_T}, \boldsymbol{\beta_U}$ = coefficients, η_T, η_U= error terms
- **Trivariate normal assumption** for the error term distribution:

$$\begin{bmatrix} \varepsilon \\ \eta_T \\ \eta_U \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho_T\sigma_T & \rho_U\sigma_U \\ \rho_T\sigma_T & \sigma_T^2 & 0 \\ \rho_U\sigma_U & 0 & \sigma_U^2 \end{bmatrix} \right)$$

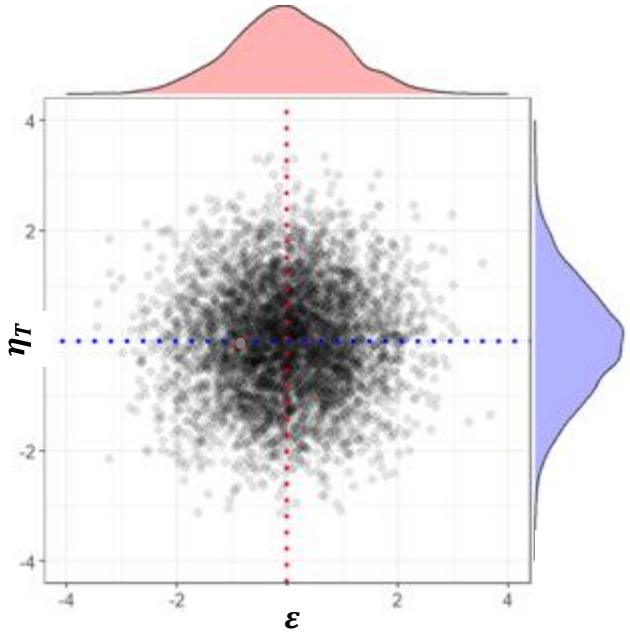
$\rho_j = Corr\,(\varepsilon, \eta_j)$

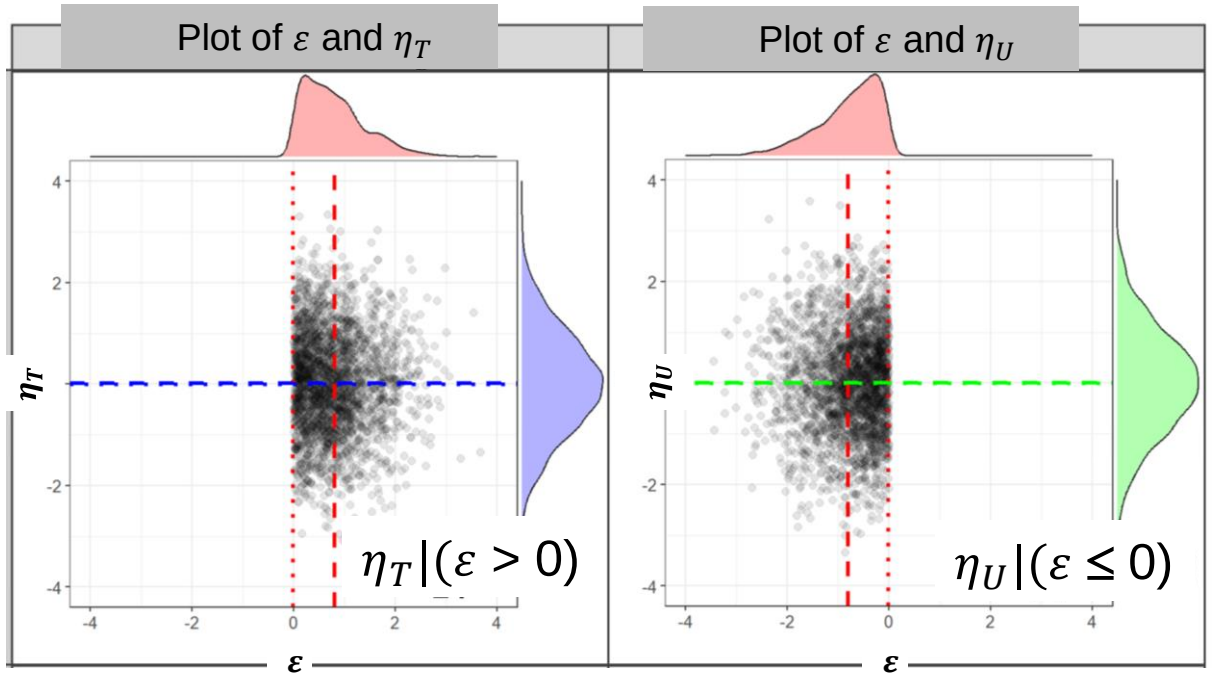
Georgia Tech

The distribution of ε is truncated for each subsample (T) and (U) (but if ε and η are ***uncorrelated***, it doesn't matter)

$Corr\ (\varepsilon,\eta_T) = \rho_T = 0$

$\rho_T = 0,\ \rho_U = 0$





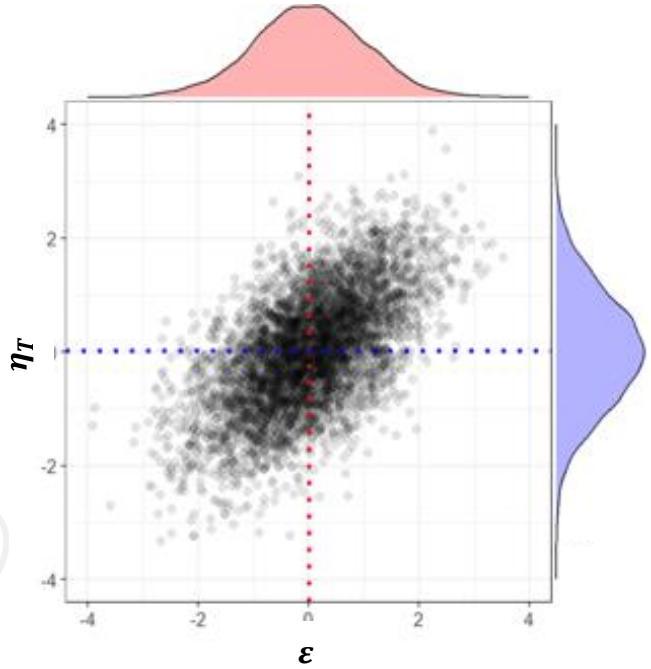
Unconditional expected value (before truncation; by color)

Conditional expected value (after truncation; by color)

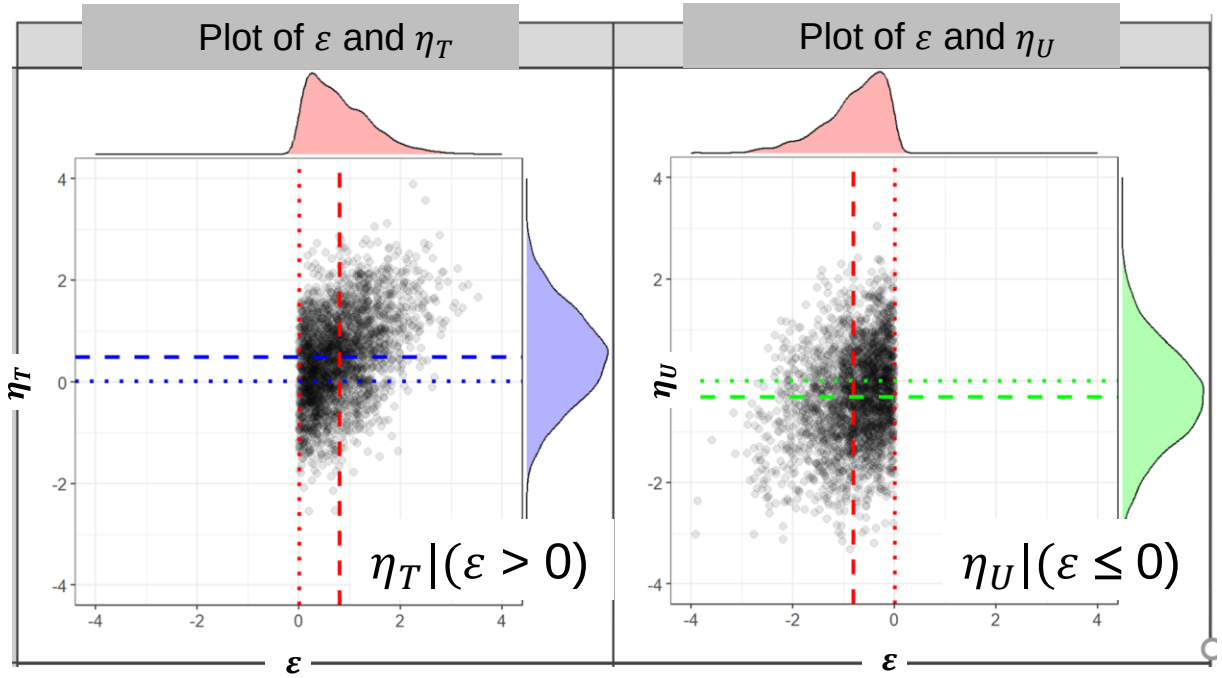
Source: Kim & Mokhtarian, 2023b

If ε and η are ***correlated***, truncation for ε propagates to the outcome model error terms, η_T and η_U

Corr $(\varepsilon, \eta_T) = \rho_T > 0$



$\rho_T > 0, \rho_U > 0$



Unconditional expected value (before truncation; by color)

Conditional expected value (after truncation; by color)

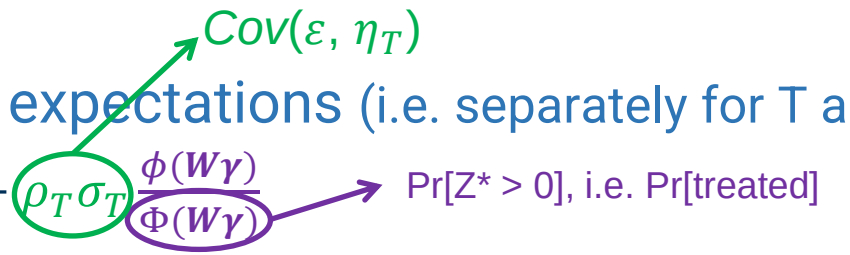
Source: Kim & Mokhtarian, 2023b

What's the problem (in math-y English)?

- We only observe people in one state (treated or untreated)
- Therefore, we are sampling from ***truncated distributions*** of the error terms (unobserved influences, ε) of the *model determining which state people are in* (i.e. the ***selection model***)
- Because those ε error terms are ***correlated*** with the η error terms of the outcome equations, that truncates the distributions of the error terms of the ***outcome equations*** as well
- Thus, although *across the population* we can assume the outcome errors have mean 0, *in our biased sample they have a non-zero mean*
- If we *do a regular regression we are assuming a 0 mean* of the outcome errors, so that non-zero mean (the selectivity bias) interferes with (i.e. biases) the estimates of the true parameters of the model

What's the solution (in math and English)?

- The ESRM "knows" the amount of the selectivity bias, and *corrects for it*, to produce *unbiased estimates* of the model parameters, and thus the treatment effect:
  - We *want* the ***unconditional*** (i.e. the *population-wide*) averages of Y:
    - $E(Y_T) = \boldsymbol{X\beta_T} + E[\varepsilon_T] = \boldsymbol{X\beta_T}$ (because the pop.-wide $E[\varepsilon_T] = E[\varepsilon_U] = 0$)
    - $E(Y_U) = \boldsymbol{X\beta_U} + E[\varepsilon_U] = \boldsymbol{X\beta_U}$
  - Instead, we have the (*factual*) ***conditional*** expectations (i.e. separately for T and U):
    - $E(Y_{T\ for\ a\ T}) = \boldsymbol{X\beta_T} + E[\varepsilon_T|\boldsymbol{Z^* > 0}] = \boldsymbol{X\beta_T} + \rho_T\sigma_T \frac{\phi(\boldsymbol{W\gamma})}{\Phi(\boldsymbol{W\gamma})}$
    - $E(Y_{U\ for\ a\ U}) = \boldsymbol{X\beta_U} + E[\varepsilon_U|\boldsymbol{Z^* \leq 0}] = \boldsymbol{X\beta_U} + \rho_U\sigma_U \left[\frac{-\phi(\boldsymbol{W\gamma})}{1-\Phi(\boldsymbol{W\gamma})}\right]$
  - And we can also have the *counterfactual* ***conditional*** expectations:
    - $E(Y_{T\ for\ a\ U}) = \boldsymbol{X\beta_T} + E[\varepsilon_T|\boldsymbol{Z^* \leq 0}] = \boldsymbol{X\beta_T} + \rho_T\sigma_T \left[\frac{-\phi(\boldsymbol{W\gamma})}{1-\Phi(\boldsymbol{W\gamma})}\right]$
    - $E(Y_{U\ for\ a\ T}) = \boldsymbol{X\beta_U} + E[\varepsilon_U|\boldsymbol{Z^* > 0}] = \boldsymbol{X\beta_U} + \rho_U\sigma_U \frac{\phi(\boldsymbol{W\gamma})}{\Phi(\boldsymbol{W\gamma})}$



(what E[Y] would be for an ***untreated*** person with traits **X, W**, ***if she were to be treated***)

(what E[Y] would have been for a **treated** person with traits **X, W**, ***before she was treated***)

Kim & Mokhtarian, 2023b

To restate the problem (with the additional math we've now seen)

- If we estimate an equation like
  - $Y_{T\ for\ a\ T} = \boldsymbol{X\beta_T} + [\varepsilon_T | \boldsymbol{Z^* > 0}]$,
- *Assuming* that $E[\varepsilon_T | \boldsymbol{Z^* > 0}] = 0$ when it is really $\rho_T \sigma_T \frac{\phi(\boldsymbol{W\gamma})}{\Phi(\boldsymbol{W\gamma})}$, the non-zero bias term gets absorbed into the $\boldsymbol{X\beta_T}$ term, which biases the estimates of $\boldsymbol{\beta_T}$ (and similarly for $\boldsymbol{\beta_U}$)

What's the solution (in math and English)? (cont'd)

- Once we properly account for the selection bias ***(giving us unbiased estimates of β_T and β_U)***, then we can compute the treatment effects as:
  - **Average treatment effect (ATE):**
    - $E(Y_T) - E(Y_U) = \boldsymbol{X(\beta_T - \beta_U)}$
  - **Treatment effect on the treated (TT):**
    - $E(Y_{T\ for\ a\ T}) - E(Y_{U\ for\ a\ T}) = \boldsymbol{X\beta_T} + E[\varepsilon_T|\boldsymbol{Z^* > 0}] - (\boldsymbol{X\beta_U} + E[\boldsymbol{\varepsilon_U}|\boldsymbol{Z^* > 0}])$
  - **Treatment effect on the untreated (TUT):**
    - $E(Y_{T\ for\ a\ U}) - E(Y_{U\ for\ a\ U}) = \boldsymbol{X\beta_T} + E[\boldsymbol{\varepsilon_T}|\boldsymbol{Z^* \leq 0}] - (\boldsymbol{X\beta_U} + E[\varepsilon_U|\boldsymbol{Z^* \leq 0}])$
  - Whereas the **observed difference** is:
    - $E(Y_{T\ for\ a\ T}) - E(Y_{U\ for\ a\ U}) = \boldsymbol{X\beta_T} + E[\varepsilon_T|\boldsymbol{Z^* > 0}] - (\boldsymbol{X\beta_U} + E[\varepsilon_U|\boldsymbol{Z^* \leq 0}])$
    - (Where, in estimation, the selection bias terms are invisibly absorbed into the β estimates, thereby biasing them)
    - Which is different from all those other quantities

counterfactual

Example application: impact of working from home (WFH) on vehicle-miles driven (VMD)

Dr. Xinyi Wang
xinyi174@mit.edu

Acronyms:

WFH = work(ing) from home

TW = telework, **TWer** = teleworker, **TWing** = teleworking

For this talk, **TW ≈ WFH**

NTWer: non-TWer, does not WFH

NUTWer: non-usual TWer, WFH < 3 days / wk

UTWer: usual TWer, WFH 3+ days / wk

VMD = (weekly) vehicle-miles driven (key travel indicator)

Georgia Tech

Some *potential* impacts of WFH on vehicle travel

- There will be local variations
- ***Multiple factors will counteract each other***
- What will the overall net impact be?

Expected (net) impact on → due to changes in ↓	vehicle-miles	vehicle-trips
Commute travel	↘	↘
Nonwork travel	?	?
Residential relocation	↗	↘
Vehicle ownership	?	?
Mode choice (e.g. transit to solo driving)	↗	↗

A tale of two types of travel diary studies of TWing

TW program evaluations

- Early (1980s – 1990s)
- Small, unrepresentative samples
- Focused on TWing
- Panel data (before-after)
- ***Found travel reductions*** *(TW **decreased** travel)*

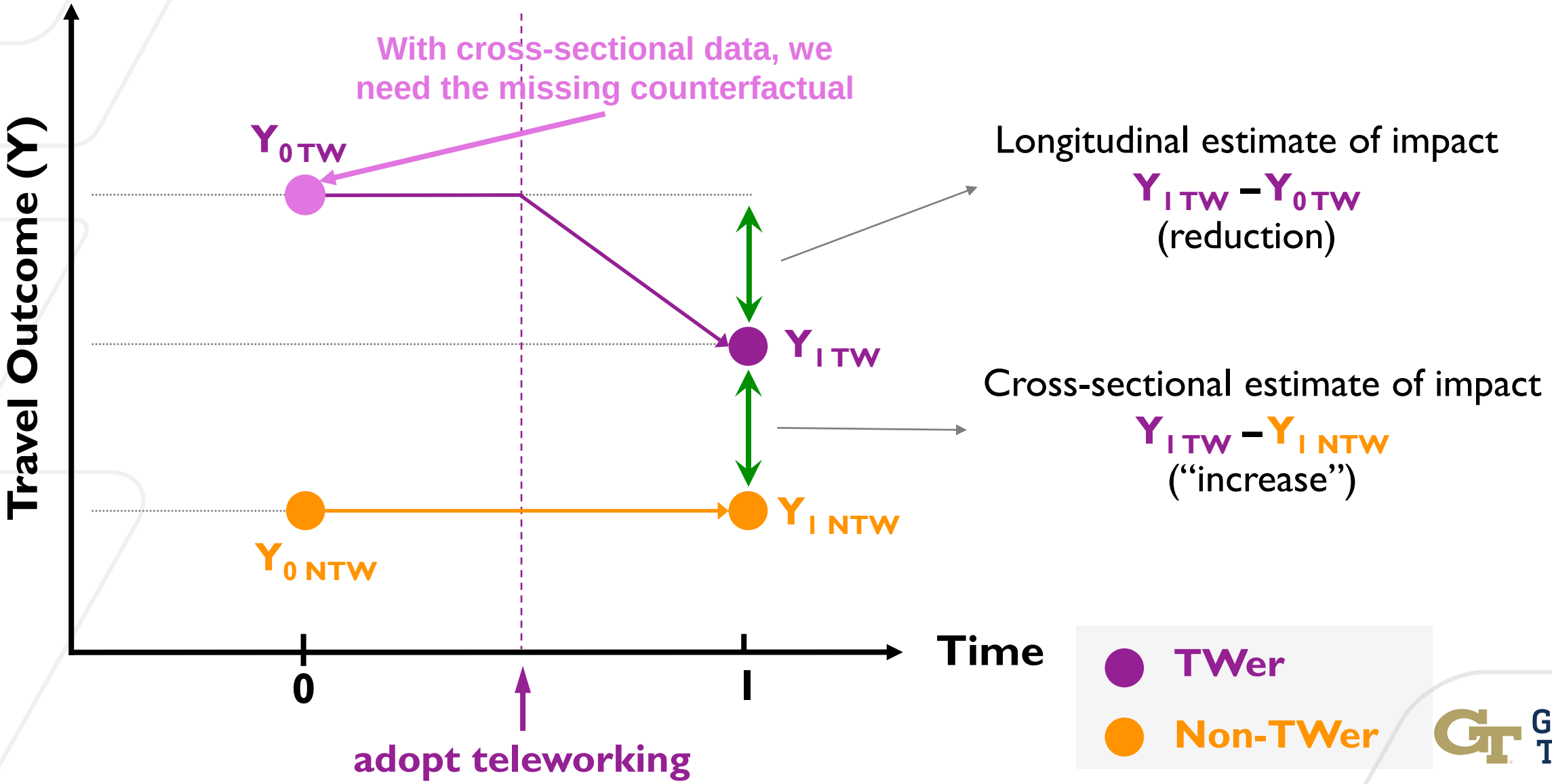
General travel surveys

- Later (2000s – 2010s)
- Large, representative samples
- No emphasis on TWing
- Cross-sectional data
- ***Finding complementarity*** *(TW **increases** travel)*

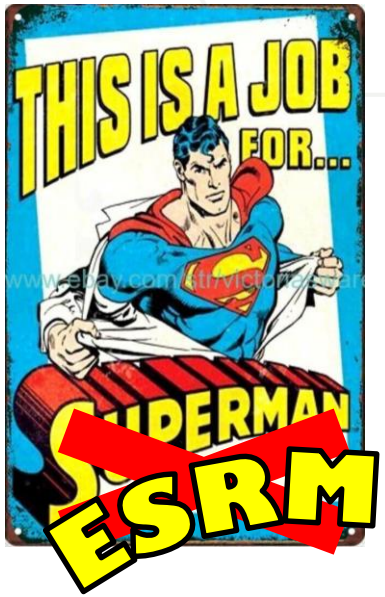
Obeid et al. 2022; Mokhtarian et al. 1995; He & Hu 2015; Kim et al. 2015; Zhu 2013; Zhu & Mason 2014

Georgia Tech

Selection bias: longitudinal vs. cross-sectional inference

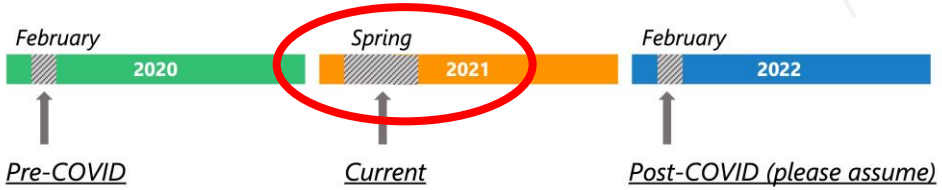


This is a job for ESRM!

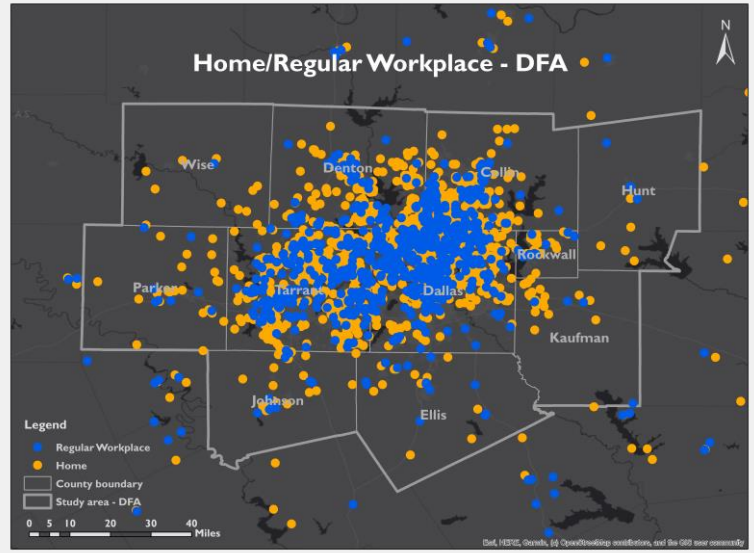


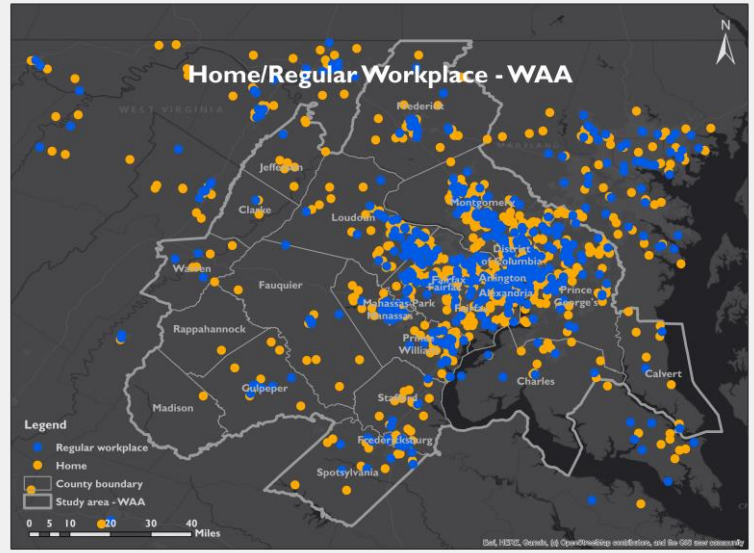
- Some of Xinyi's contributions:
  - Two TWing treatments (**NUTW** & **UTW**), plus untreated (**NTW**)
  - Estimated **ordinal** as well as **multinomial** *selection models*
  - With **log-transformed** VMD (then **back-transformed** treatment effects)
  - **Innovative visualizations** of aggregate and disaggregate treatment effect components
  - **R scripts** to improve the flexibility and statistical efficiency of previously-existing estimation software
  - **Careful exposition**, including clearing up some critical ambiguities in Dubin & McFadden (1984)

Data overview (N = 1,584)



- **Funded by Cintra (Ferrovial)**
  - Impact of COVID-influenced TW on toll revenues
- **Survey focus**
  - Telework and work patterns **before**, **during**, and **after** COVID-19
- **Study areas**
  - Dallas-Fort Worth-Arlington (DFA)
  - Washington-Arlington-Alexandria (WAA)
- **Respondent sources**
  - Cintra database (DB): current and potential customers who consented to be surveyed
  - Online panel (OP): three vendor companies
- **Data collection Feb. 24 - April 30, 2021**

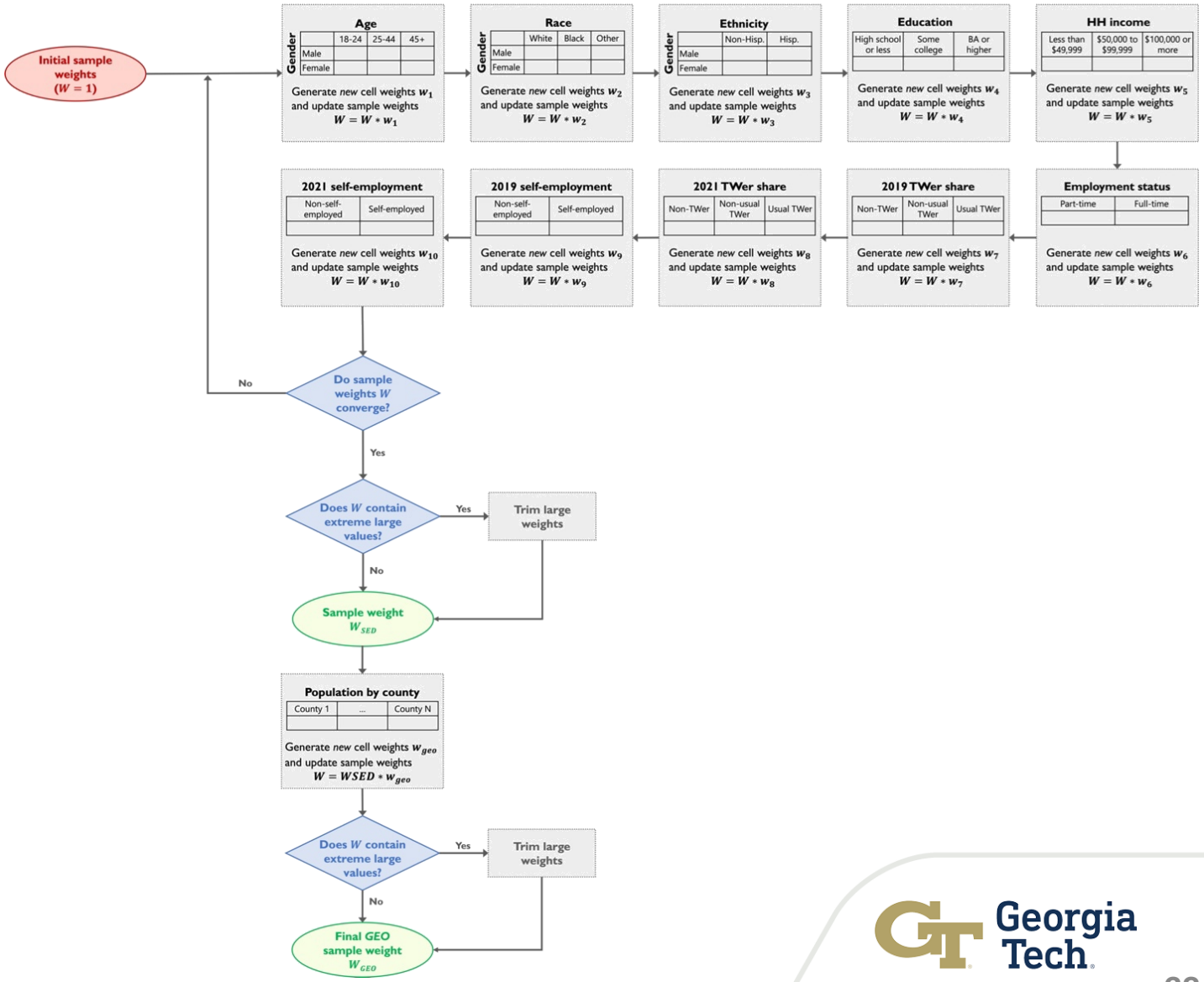




Sample weights

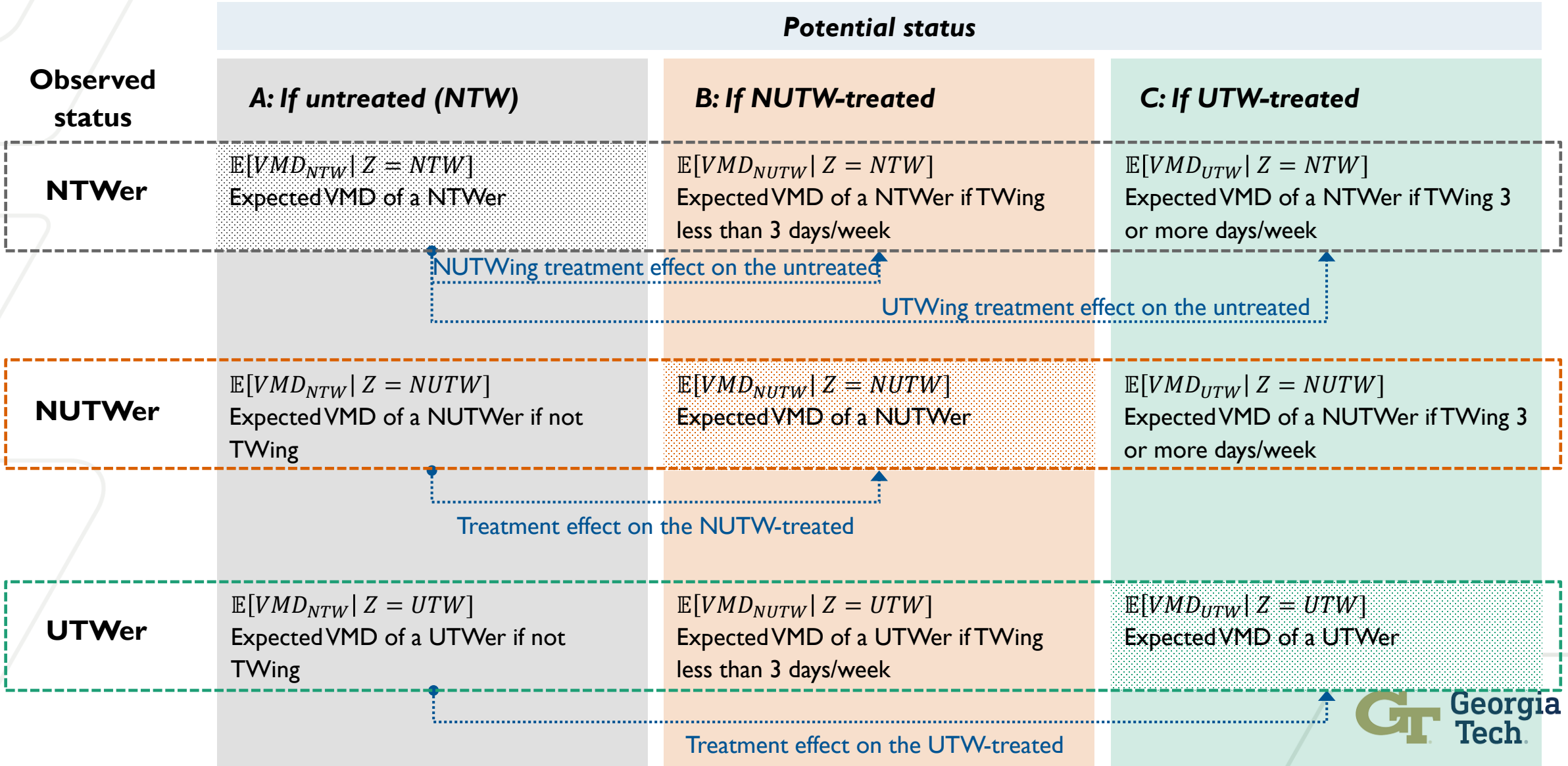
Sample was weighted (by region) to reflect pop. distributions on:

- Gender
- Age
- Race
- Ethnicity
- Education
- Household income
- Employment status
- **2019 (pre-COVID) and 2021 (during-COVID) shares of**
  - **Non-TWers**
  - **Non-usual TWers (< 3 days/wk)**
  - **Usual TWers (3+ days/wk)**
- 2019 (pre-COVID) and 2021 (during COVID) self-employment shares
- Employed population by county

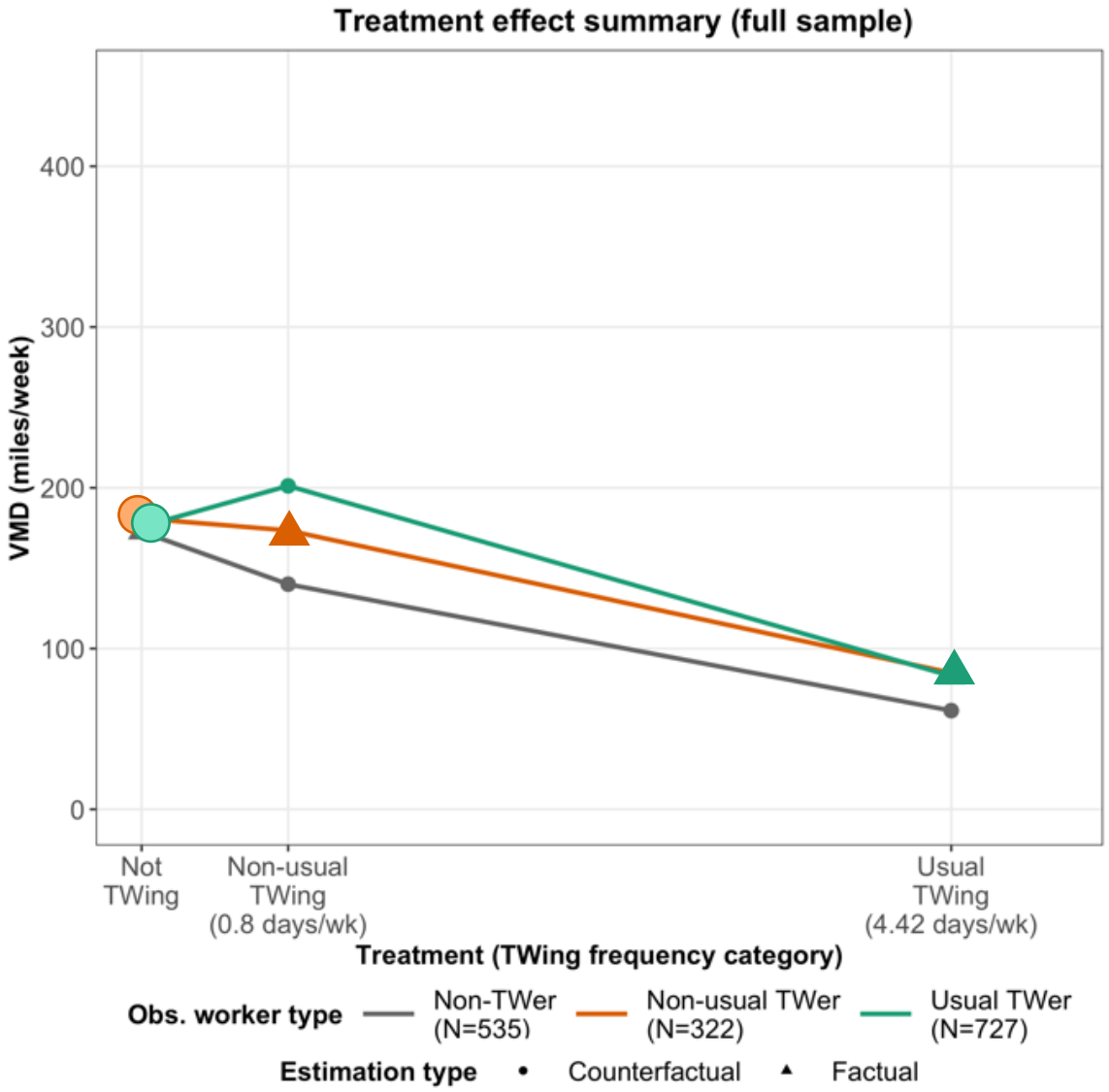


Components of treatment effects (TEs)

$\mathbb{E}[VMD_{j'}|Z = j]$: Expected weekly VMD if {in, moved to} state j' given observed to belong to group j

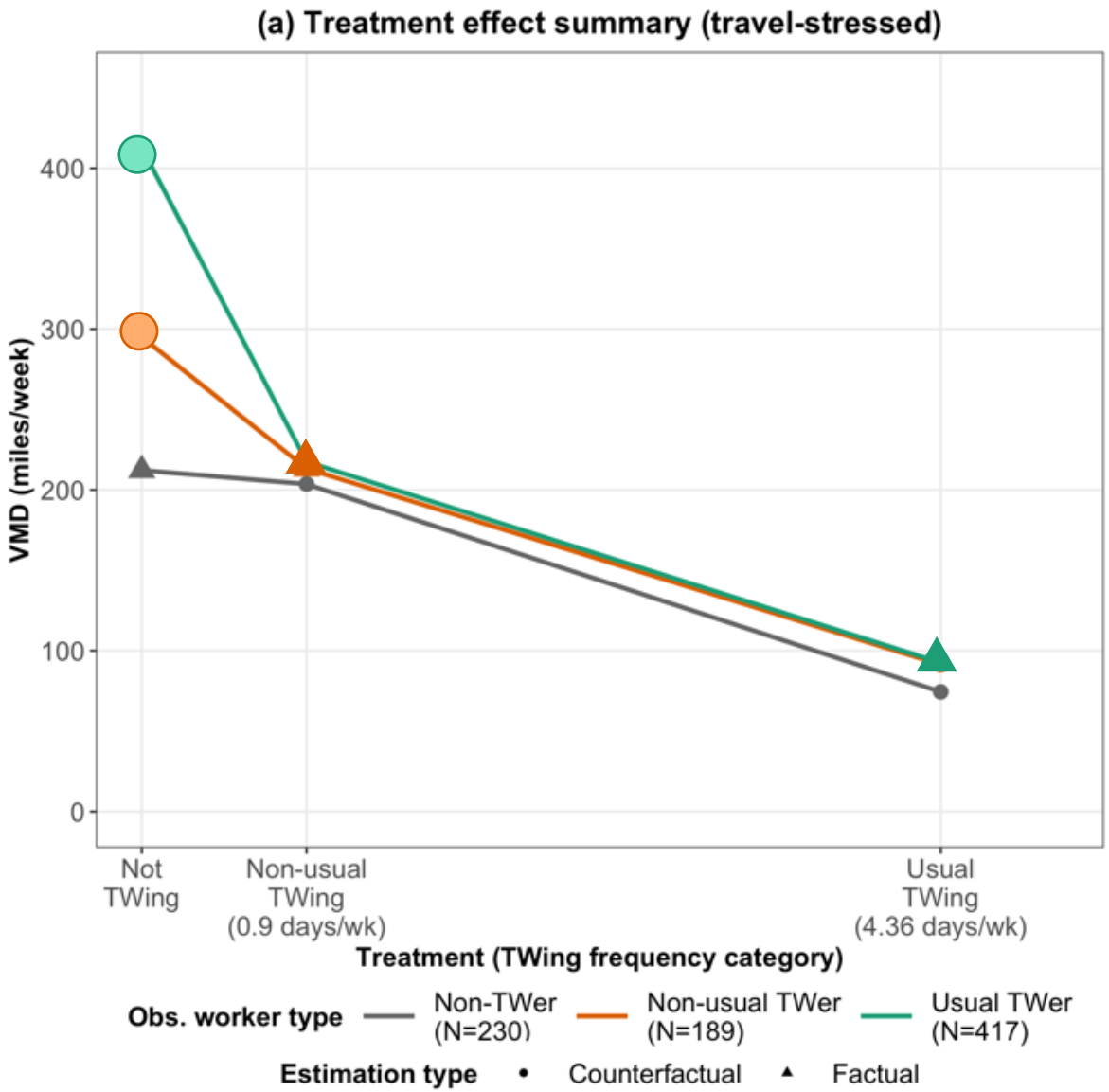


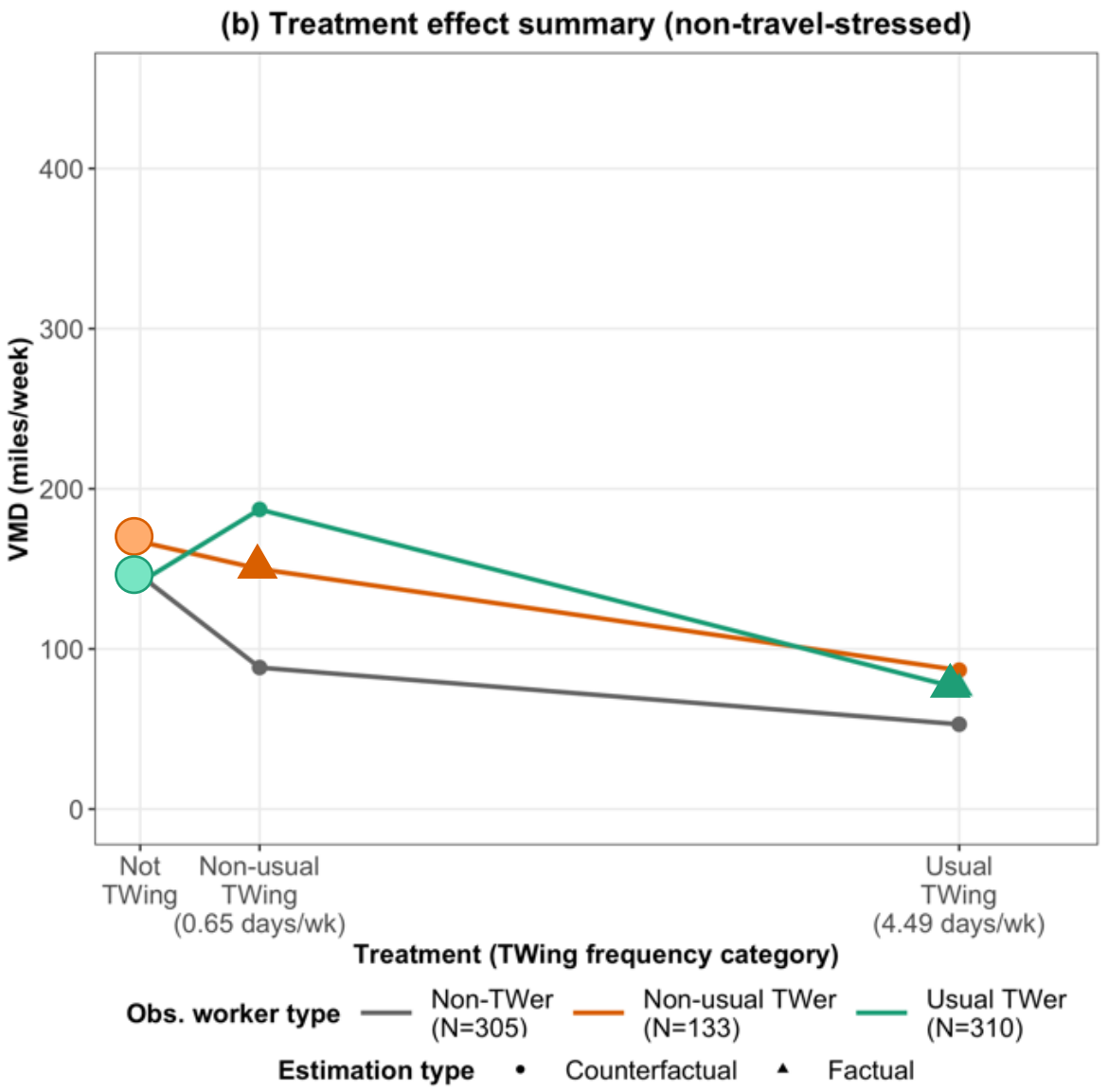
Full sample model (2 treatments: NUTWing & UTWing)



- Focusing on the TEs (compared to not TWing) for the two observed TWer groups:
  - ▲, ▲ = factual
  - ●, ● = (NTW) counterfactual
- **Ave. VMD of non-usual TWers (15% of the sample) barely declines** (not statistically significant)
  - -7.5 mi/wk, or, -9.4 mi/TWing occasion
- **Ave. VMD of usual TWers (28% of the sample) declines substantially** (statistically significant)
  - -92.9 mi/wk, or, -21.0 mi/TWing occ.

Comparison of travel-stressed and non-travel-stressed

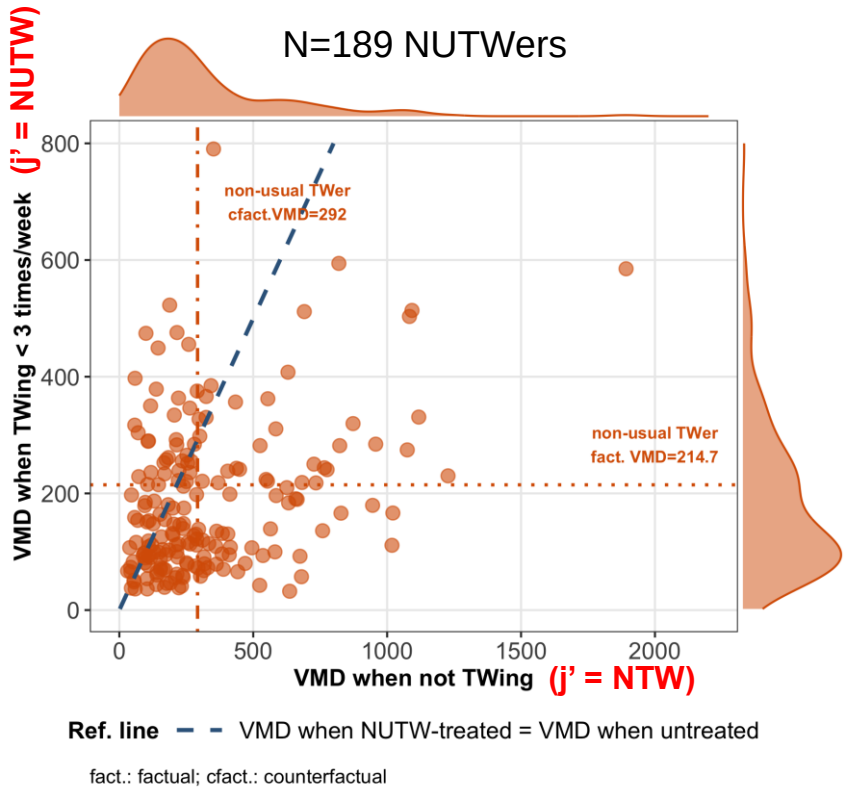




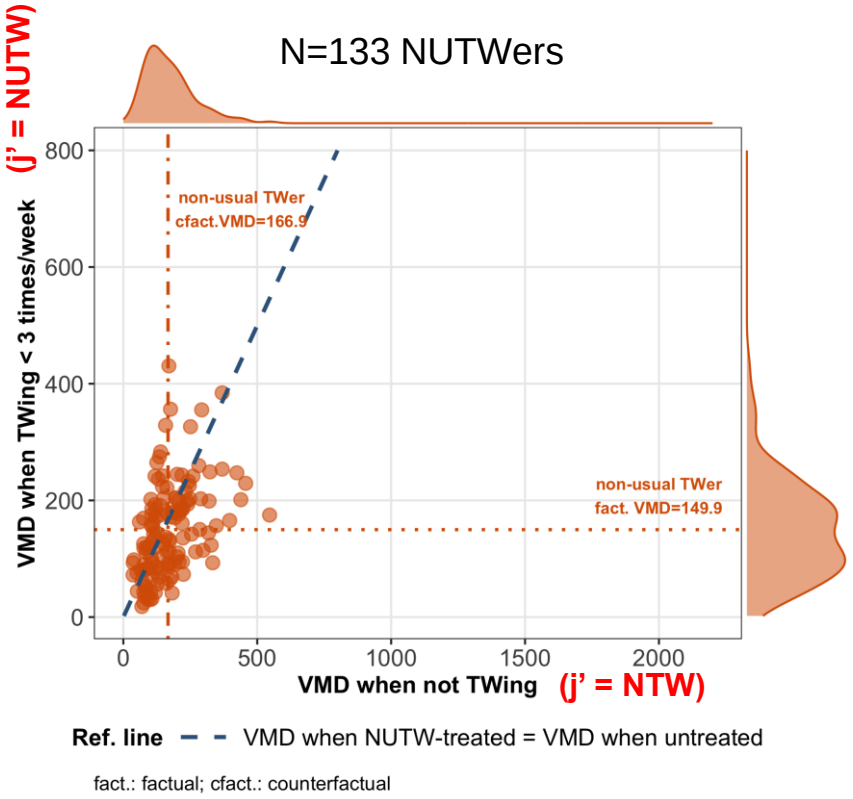
$\mathbb{E}[VMD_{j'}|Z = NUTW]$

Disaggregate TE components for NUTW treatment

Travel-stressed



Non-travel-stressed



- Individuals **below** the reference line generate ***more*** VMD when they ***do not telework***
- Individuals **above** the reference line generate ***more*** VMD when they ***adopt (non-usual) teleworking***
- 29.6% of ***travel-stressed*** NUTWers increase VMD after adopting non-usual TWing
- 38.3% of ***non-travel-stressed*** NUTWers increase VMD after adopting non-usual TWing

Georgia Tech

Summary: empirical

- We quantified and compared the impact of **teleworking** on **vehicle-miles driven (VMD)** for different types of teleworkers
  - By teleworking **frequency categories**: non-TWer, non-usual TWer, usual TWer
  - By teleworking-related **motive**: travel-stressed or not
- In all models, **TWing reduced VMD** on average, for its adopters
- TWing reduced VMD ***most, for travel-stressed TWers***
- Substantially more **non-travel-stressed TWers *increased* VMD** after beginning to TW than **travel-stressed TWers** did

Georgia Tech

Summary: methodological

- The endogenous switching regression model (ESRM) is designed to deal with selection biases in a cross-sectional setting, where we only observe people in ***one state (the factual state)***
- ESRMs can generate estimates of the outcome of interest as if those people were in ***other states (counterfactual states)***
- Since selection biases are controlled, ESRMs generate unbiased estimates of treatment effects in a cross-sectional setting
- In doing this, the cross-sectional results *can* be consistent with the longitudinal ones

Georgia Tech

Conclusion (1)

- What are the **generic ingredients** of a selection bias like the ones we have been discussing?
  - **Discrete intervention / experiment / treatment**
  - Need to **evaluate the impact of the treatment** (for the treated, the untreated, and/or the population at large) **on an outcome** of interest
  - **Cross-sectional data**, containing **untreated** (non-adopters, those who didn't experience the intervention) and one or more types of **treated** groups
  - Treated and untreated cases **differ from each other in unobserved ways that influence the outcome**
- **Generic question:** could (part of) the (apparent) effect of the treatment on the outcome be due to *unobserved pre-treatment differences* between the treated and the untreated??? (E.g., differences in the *willingness to be treated*???)
  - If so, then (without correcting for the bias) you *cannot* assume that future treatment/ adoption (especially if becoming *mandatory*) will provide the same results

Georgia Tech

Conclusion (2)

- Other transportation applications

Treatment	Outcome	Implication of a selection bias
Built environment (residential self-selection)	Travel behavior (various)	People who choose a neighborhood consistent with their travel predispositions will probably travel differently from those who are nudged into it for other reasons (policy, constraints, other objectives)
Mode choice	Satisfaction with trip	The fact that mode "*x*" users are the most satisfied doesn't mean that non-users would have the same satisfaction if they were to use that mode…
Voluntary safety measure (seat belt, helmet)	Accident propensity/ rate	People who opt in may be more risk averse in other ways, or may instead compensate with riskier behavior; either way, their outcomes may not match those of people forced (by law) to adopt
Voluntary behavior change program	Travel behavior (various)	Results for people who opt in will probably not match those for people who are pushed (via policy, law, cost) to change

Conclusion (3)

sunghookim@hanyang.ac.kr

- There are many other applications beyond transportation
- We *very often* ***only*** *have* ***cross-sectional*** *data*, so…
  - We hope that our collection of papers explicating, demonstrating, and visually illustrating variants of this approach will lead to greater interest in it and adoption of it

- **Kim SH & PL Mokhtarian** (2023a) Comparisons of observed and unobserved parameter heterogeneity in modeling vehicle-miles driven. *Transportation Research Part A* **172**, 103614, 2023.
- **Kim SH & PL Mokhtarian** (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://link.springer.com/article/10.1007/s11116-023-10384-2.
- **Wang X & PL Mokhtarian** (2024) Examining the treatment effect of teleworking on vehicle-miles driven: Applying an ordered probit selection model and incorporating the role of travel stress. *Transportation Research A*, in press. Available from the authors.
- **Wang X & PL Mokhtarian** (2023) Reintroducing multinomial logit switching regression models: Examining the treatment effect of two teleworking frequency categories on vehicle-miles driven. Paper # TRBAM-S-23-02767 presented at the Annual Meeting of the Transportation Research Board, January 2024. Available from the authors.

Georgia Tech

Thank you! Questions?

patmokh@gatech.edu

Happy to share the slides

Georgia Tech

References (1/2)

- **Choo S, PL Mokhtarian, & I Salomon** (2005) Does telecommuting reduce vehicle-miles traveled? An aggregate time series analysis for the U.S. *Transportation* **32(1)**, 37-64.
- **Dubin JA & DL McFadden** (1984) An econometric analysis of residential electric appliance holdings and consumption. *Econometrica* **52(2),** 345-362.
- **He S & L Hu** (2015) Telecommuting, income, and out-of-home activities. *Travel Behaviour and Society* **2(3)**, 131-147.
- **Heckman J, JL Tobias, & E Vytlacil** (2001) Four Parameters of Interest in the Evaluation of Social Programs. *Southern Economic Journal* **68(2)**, 211-223.
- **Kim SN, S Choo, & PL Mokhtarian** (2015) Home-based telecommuting and intra-household interactions in work and non-work travel: A seemingly unrelated censored regression approach. *Transportation Research A* **80**, 197-214.
- **Kim SH & PL Mokhtarian** (2023a) Comparisons of observed and unobserved parameter heterogeneity in modeling vehicle-miles driven. *Transportation Research Part A* **172**, 103614, 2023.
- **Kim SH & PL Mokhtarian** (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://link.springer.com/article/10.1007/s11116-023-10384-2.
- **Mishra GS** (2017) *Estimating the Travel Behavior Effects of Technological Innovations from Cross-sectional Observed Data: Applications to Carsharing and Telecommuting*. Doctoral dissertation. University of California, Davis, available at https://www.proquest.com/docview/1947638403.
- **Mokhtarian PL, SL Handy, & I Salomon** (1995) Methodological issues in the estimation of the travel, energy, and air quality impacts of telecommuting. *Transportation Research A* **29(4)**, 283-302.
- **Obeid H, ML Anderson, MA Bouzaghrane, & J Walker** (2022) Does telecommuting reduce tripmaking? Evidence from a U.S. panel during the COVID-19 impact and recovery periods. SSRN-id 4213516, available from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4213516.

Georgia Tech

References (2/2)

- **Su R, EC McBride, & KG Goulias** (2021) Unveiling daily activity pattern differences between telecommuters and commuters using human mobility motifs and sequence analysis. *Transportation Research A* **147**, 106-132.
- **Wang X & PL Mokhtarian** (2024) Examining the treatment effect of teleworking on vehicle-miles driven: Applying an ordered probit selection model and incorporating the role of travel stress. *Transportation Research A*, in press. Available from the authors.
- **Wang X & PL Mokhtarian** (2023) Reintroducing multinomial logit switching regression models: Examining the treatment effect of two teleworking frequency categories on vehicle-miles driven. Paper # TRBAM-S-23-02767 accepted for presentation at the Annual Meeting of the Transportation Research Board, January 2024. Available from the authors.
- **Wang X, SH Kim, & PL Mokhtarian** (2024) Identifying teleworking-related motives and comparing telework frequency expectations in the post-pandemic world: A latent class modeling approach. *Transportation Research A*, in press. Available from the authors.
- **Wang X, SH Kim, & PL Mokhtarian** (2023d) Teleworking behavior pre-, during, and expected post-COVID: Identification and empirical description of trajectory types. *Travel Behaviour and Society* **33**, 100628.
- **Zhu P** (2013) Telecommuting, household commute and location choice. *Urban Studies* **50(12)**, 2441-2459.
- **Zhu P & SG Mason** (2014) The impact of telecommuting on personal vehicle usage and environmental sustainability. *International Journal of Environmental Science and Technology* **11(8)**, 2185-2200.

Georgia Tech