# Big Data for Finite Population Inference

## Calibrating pseudo-weights based on estimated control totals using the General Regression Estimator


Michael R. Elliott[†]
Ali Rafei[†]
Carol A.C. Flannagan[‡]

[†]Michigan Program in Survey Methodology

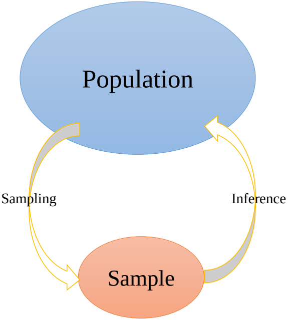
[‡]University of Michigan Transportation Research Institute


AAPOR 2019

# Big Data vs survey data

- The 21st century is witnessing a re-emergence of non-probability sampling methods for policy-making, health and social research.

- Probability sampling, which has been the "gold population inference, is declining in popularity.
  1. The upward trends of non-response rate
  2. The rising cost and complexity

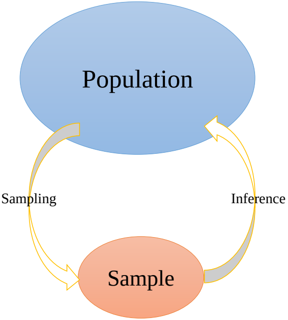



- New generations of automated processes have evolved, leading to ever-accumulation of massive volume of unstructured information, so-called "Big Data".
  1. Easier to access, less expensive to collect, and highly detailed
  2. A potentially rich treasure for producing official statistics
  3. But introducing multiple new challenges

# Big Data vs survey data

- The 21st century is witnessing a re-emergence of non-probability sampling methods for policy-making, health and social research.

- Probability sampling, which has been the "gold population inference, is declining in popularity.
  1. The upward trends of non-response rate
  2. The rising cost and complexity




- New generations of automated processes have evolved, leading to ever-accumulation of massive volume of unstructured information, so-called "Big Data".
  1. Easier to access, less expensive to collect, and highly detailed
  2. A potentially rich treasure for producing official statistics
  3. But introducing multiple new challenges

# Review of existing approaches

- Considering *ignorable* assumption in Big Data ($B$), the joint density of the outcome ($Y$) and selection indicator ($\delta_B$) conditional on $X$ is given by:

$$\begin{aligned} f(Y, \delta_B|X) &= f(Y|X)f(\delta_B|Y, X) \\ &= f(Y|X)f(\delta_B|X) \end{aligned}$$

- In presence of a reference survey ($R$) with a set of overlapping covariates ($X$), three approaches can be taken:

1. **Quasi-randomization:**
   Estimating pseudo-inclusion probabilities by modeling $f(\delta_B|X)$
2. **Super-population:**
   Predicting the outcome for non-sampled units by modeling $f(Y|X)$
3. **Doubly robust weighting:**
   Combining the two to further protect against model misspecification

- Let combine $B$ with $R$ and define $Z_i = I(i \in B)$, given $\delta_{Bi} + \delta_{Ri} = 1$.

# Review of existing approaches

- Considering *ignorable* assumption in Big Data ($B$), the joint density of the outcome ($Y$) and selection indicator ($\delta_B$) conditional on $X$ is given by:

$$\begin{aligned} f(Y, \delta_B | X) &= f(Y|X) f(\delta_B | Y, X) \\ &= f(Y|X) f(\delta_B | X) \end{aligned}$$

- In presence of a reference survey ($R$) with a set of overlapping covariates ($X$), three approaches can be taken:

  1. **Quasi-randomization:**
     Estimating pseudo-inclusion probabilities by modeling $f(\delta_B|X)$
  2. **Super-population:**
     Predicting the outcome for non-sampled units by modeling $f(Y|X)$
  3. **Doubly robust weighting:**
     Combining the two to further protect against model misspecification

- Let combine $B$ with $R$ and define $Z_i = I(i \in B)$, given $\delta_{Bi} + \delta_{Ri} = 1$.

# Quasi-randomization (QR)

- Traditionally, propensity scores are used to estimate pseudo-weights (Czajka et al., 1992; Lee., 2006; Schonlau et al., 2009).

**PS weighting:**

$$\hat{\omega}_i^{PS} = \frac{1 - \hat{e}(x_i)}{\hat{e}(x_i)}, \ \ \forall i \in B$$

where $\hat{e}_i$ is predicted by:

$$\hat{e}(x_i) = \hat{P}(Z_i = 1 | X_i = x_i) = \frac{exp\{x_i^T \hat{\beta}\}}{1 + exp\{x_i^T \hat{\beta}\}}, \forall i \in B \cup R$$

- When $R$ is NOT SRS, Valliant & Dever (2011) recommend using *pseudo-MLE* to estimate $\beta$, i.e. solving the estimating equations:

$$\sum_{i \in B \cup R} \omega_i x_i^T [y_i - e(x_i)] = 0$$

where

$$\omega_i = \begin{cases} \pi_i^{-1}, & \text{for } i \in R \\ 1, & \text{for } i \in B \end{cases}$$

# Quasi-randomization (QR)

- Traditionally, propensity scores are used to estimate pseudo-weights (Czajka et al., 1992; Lee., 2006; Schonlau et al., 2009).

**PS weighting:**

$$\hat{\omega}_i^{PS} = \frac{1-\hat{e}(x_i)}{\hat{e}(x_i)}, \ \ \forall i \in B$$

where $\hat{e}_i$ is predicted by:

$$\hat{e}(x_i) = \hat{P}(Z_i = 1 | X_i = x_i) = \frac{exp\{x_i^T \hat{\beta}\}}{1 + exp\{x_i^T \hat{\beta}\}}, \forall i \in B \cup R$$

- When $R$ is NOT SRS, Valliant & Dever (2011) recommend using *pseudo-MLE* to estimate $\beta$, i.e. solving the estimating equations:

$$\sum_{i \in B \cup R} \omega_i x_i^T [y_i - e(x_i)] = 0$$

where

$$\omega_i = \begin{cases} \pi_i^{-1}, & \text{for } i \in R \\ 1, & \text{for } i \in B \end{cases}$$

# Quasi-randomization (QR)

- Elliott et al. (2010) derive pseudo-weights directly in the combined samples by multiply applying the Bayes rule.

**Pseudo-weighting:**

$$\hat{\omega}_i^{PW} = \hat{\pi}(x_i)^{-1} \times \frac{1 - \hat{e}(x_i)}{\hat{e}(x_i)}$$

where $\hat{\pi}(x_i) = \hat{P}(\delta_{Ri} = 1 | X_i = x_i)$ can be modeled via *Beta* regression.

- When $R$ is SRS, then $\hat{\pi}(x_i)^{-1} \propto 1$, so $\hat{\omega}_i^{PW} = \hat{\omega}_i^{PS}$.
- When $R$ and $B$ are similar in terms of the dist. of $X$, then $\hat{e}(x_i) = 1/2$, so $\hat{\omega}_i^{PW} = \hat{\pi}(x_i)^{-1}$.
- Propensity weighting lacks adequate theory when $R$ is not SRS.
- It is expected $\hat{\omega}_i^{PW}$ performs better than $\hat{\omega}_i^{PS}$ in bias reduction when one sample overrepresents $X$ and the other underrepresent $X$.

# Quasi-randomization (QR)

- Elliott et al. (2010) derive pseudo-weights directly in the combined samples by multiply applying the Bayes rule.

**Pseudo-weighting:**

$$\hat{\omega}_i^{PW} = \hat{\pi}(x_i)^{-1} \times \frac{1-\hat{e}(x_i)}{\hat{e}(x_i)}$$

where $\hat{\pi}(x_i) = \hat{P}(\delta_{Ri} = 1 | X_i = x_i)$ can be modeled via *Beta* regression.

- When $R$ is SRS, then $\hat{\pi}(x_i)^{-1} \propto 1$, so $\hat{\omega}_i^{PW} = \hat{\omega}_i^{PS}$.
- When $R$ and $B$ are similar in terms of the dist. of $X$, then $\hat{e}(x_i) = 1/2$, so $\hat{\omega}_i^{PW} = \hat{\pi}(x_i)^{-1}$.
- Propensity weighting lacks adequate theory when $R$ is not SRS.
- It is expected $\hat{\omega}_i^{PW}$ performs better than $\hat{\omega}_i^{PS}$ in bias reduction when one sample overrepresents $X$ and the other underrepresent $X$.

# Super-population (SP)

- Dever & Valliant (2016) propose using General Regression (GREG) estimator based on estimated control totals from a benchmark sample.

**GREG for population total**

$$\hat{y}_U = \sum_{i \in B} y_i + (\hat{t}_R - \hat{t}_B)\hat{\beta}$$

where $\beta$ can be estimated from $y = X^T\beta + \epsilon_i$ and $\epsilon_i \sim N(0, \sigma^2)$

- The main advantage of GREG is that it produces a single set of calibration weights that can be applied to any outcome variable.

**Calibration weights based on GREG**

$$\hat{\omega}_i^{GR} = 1 + (\hat{t}_R - \hat{t}_B)(X^TX)^{-1}x_i^T$$

- Simulations show that GREG performs well even for non-normal outcomes. However, it is possible GREG leads to negative weights.

# Super-population (SP)

- Dever & Valliant (2016) propose using General Regression (GREG) estimator based on estimated control totals from a benchmark sample.

**GREG for population total**

$$\hat{y}_U = \sum_{i \in B} y_i + (\hat{t}_R - \hat{t}_B)\hat{\beta}$$

where $\beta$ can be estimated from $y = X^T\beta + \epsilon_i$ and $\epsilon_i \sim N(0, \sigma^2)$

- The main advantage of GREG is that it produces a single set of calibration weights that can be applied to any outcome variable.

**Calibration weights based on GREG**

$$\hat{\omega}_i^{GR} = 1 + (\hat{t}_R - \hat{t}_B)(X^T X)^{-1} x_i^T$$

- Simulations show that GREG performs well even for non-normal outcomes. However, it is possible GREG leads to negative weights.

# Doubly robust (DR) weighting

- Both QR and SP approaches assume models are correctly specified.
- Robins et al (1994) propose a class of adjustment methods such that estimates are consistent if either QR or SP model holds.
- We combine PW with GREG and show that calibrating pseudo-weighted estimates based on GREG with estimated totals is doubly robust (Wu & Sitter, 2001).
- Yet, this method yield a single set of weights, which we call "doubly robust weights".

**Doubly Robust weighting:**

$$\begin{aligned}\hat{\omega}_i^{DR} &= \hat{\omega}_i^{PW} \times [1 + (\hat{t}_R - \hat{t}_{\tilde{w}B})(X^T \tilde{W} X)^{-1} x_i^T] \\ &= \hat{\pi}(x_i)^{-1} \frac{1 - \hat{e}(x_i)}{\hat{e}(x_i)} [1 + (\hat{t}_R - \hat{t}_{\hat{w}B})(X^T \hat{W} X)^{-1} x_i^T]\end{aligned}$$

where $\hat{t}_{\hat{\omega}B}$ is the pseudo-weighted estimate of total in $B$, and $\hat{W} = diag\{\hat{\omega}_i^{PW}\}$.

# Doubly robust (DR) weighting

- Both QR and SP approaches assume models are correctly specified.
- Robins et al (1994) propose a class of adjustment methods such that estimates are consistent if either QR or SP model holds.
- We combine PW with GREG and show that calibrating pseudo-weighted estimates based on GREG with estimated totals is doubly robust (Wu & Sitter, 2001).
- Yet, this method yield a single set of weights, which we call "doubly robust weights".

**Doubly Robust weighting:**

$$\hat{\omega}_i^{DR} = \hat{\omega}_i^{PW} \times [1 + (\hat{t}_R - \hat{t}_{\tilde{w}B})(X^T \tilde{W} X)^{-1} x_i^T]$$
$$= \hat{\pi}(x_i)^{-1} \frac{1 - \hat{e}(x_i)}{\hat{e}(x_i)} [1 + (\hat{t}_R - \hat{t}_{\hat{w}B})(X^T \hat{W} X)^{-1} x_i^T]$$

where $\hat{t}_{\hat{\omega}B}$ is the pseudo-weighted estimate of total in $B$, and $\hat{W} = diag\{\hat{\omega}_i^{PW}\}$.

# Simulation study

- Two correlated covariates were generated as below:

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim MVN\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 & 0.4 \\ 0.4 & 1 \end{pmatrix}\right)$$

- We assumed $Y$ is a binary outcome with Bernoulli distribution variable as below:

$$Y_c|X \sim N(2 - 3x_1 + 2x_2 + 6x_1x_2, \sigma^2 = 1), \quad Y_b|X \sim b\left(\frac{e^{-2-3x_1+2x_2+6x_1x_2}}{1 + e^{-2-3x_1+2x_2+6x_1x_2}}\right)$$

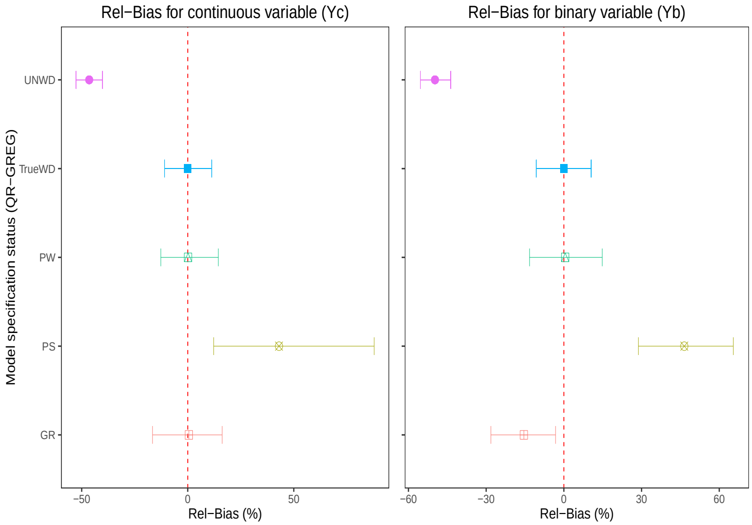
- Each units in the population were assigned two sets of unequal probabilities of selection, which were correlated with $W$ through a *logistic* link as below:

$$P(\delta_{Ri} = 1|X) = \frac{e^{-5.9+0.3x_1-0.5x_2+0.1x_1x_2}}{1 + e^{-5.9+0.3x_1-0.5x_2+0.1x_1x_2}}, P(\delta_{Bi} = 1|x) = \frac{e^{-9.5-x_1+x_2-x_1x_2}}{1 + e^{-9.5-x_1+x_2-x_1x_2}}$$

- The simulation was iterated $K = 1000$ times, and rel-Bias and 95%CI coverage rates were computed.
- Different scenarios of model misspecification were examined.

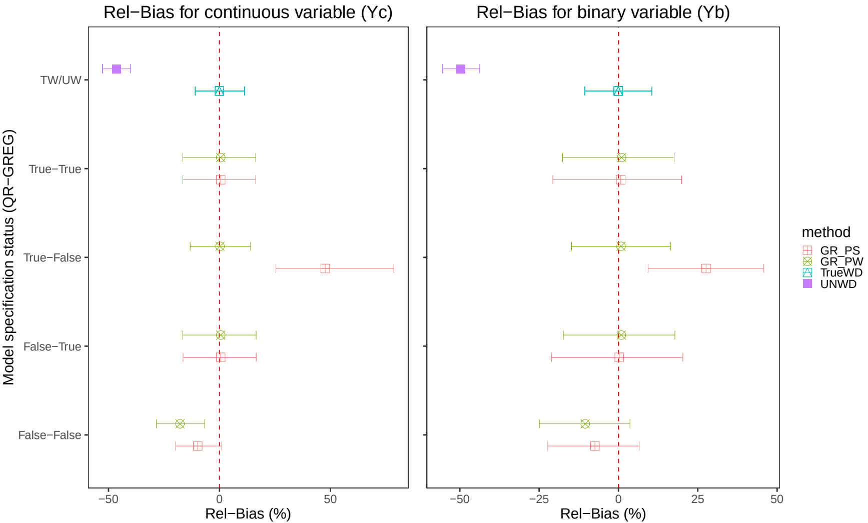
# Simulation results

- The simulation results for $n_R = 200$ and $n_B = 1000$



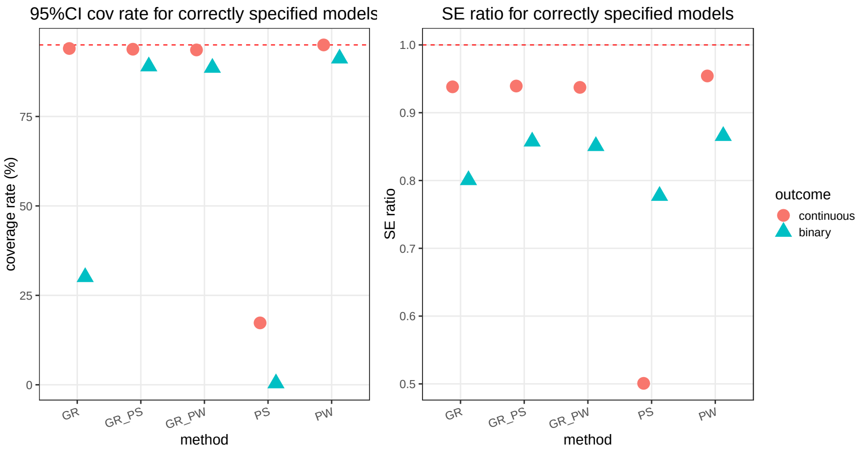
# Simulation results

- The simulation results for $n_R = 200$ and $n_B = 1000$



# Simulation results

- The simulation results for $n_R = 200$ and $n_B = 1000$



# Naturalistic Driving Studies (NDS)

- One real-world application of sensor-based Big Data.
- Driving behaviors are continuously monitored via instrumented vehicles.
- A rich resource for exploring crash causality, traffic safety, and travel dynamics.
- Launched in 2010, the 2nd Strategic Highway Research Program (SHRP2) is the world's largest NDS to date.
- ∼5 million trips and ∼50 million driven miles were recorded for a total of 3,700 participant-year.
- Participants were selected from six sites across the US.
- A combination of quota and convenience sampling was used to recruit samples; Youngest/eldest groups were oversampled.

# Research objective and benchmark

- **Objective:**
  To adjust for potential selection bias in SHRP2 dataset using National Household Travel Survey (NHTS) 2017 national probability as benchmark.
- 15 covariates were identified in common between NHTS and SHRP2 datasets.
  - **Participants' demographics**: gender, age, race, ethnicity, education, HH income, HH size, house ownership, birth country, job status
  - **Vehicle characteristics:** vehicle age, vehicle type, vehicle manufacturer, mileage
- To make the two datasets comparable, the following steps were taken:
  - Only drivers of the trips were kept in NHTS.
  - Trips with modes other than car/SUV/van/truck were removed.
  - trips for which public transportation was used were omitted.
  - all the trips with average speed $< 20$Km/h and $> 120$Km/h were dropped.
- The sample sizes were $n_{NHTS} = 447,493$ and $n_{SHRP2} = 3,458,826$.

# Research objective and benchmark

- **Objective:**
  To adjust for potential selection bias in SHRP2 dataset using National Household Travel Survey (NHTS) 2017 national probability as benchmark.
- 15 covariates were identified in common between NHTS and SHRP2 datasets.
  - **Participants' demographics**: gender, age, race, ethnicity, education, HH income, HH size, house ownership, birth country, job status
  - **Vehicle characteristics:** vehicle age, vehicle type, vehicle manufacturer, mileage
- To make the two datasets comparable, the following steps were taken:
  - Only drivers of the trips were kept in NHTS.
  - Trips with modes other than car/SUV/van/truck were removed.
  - trips for which public transportation was used were omitted.
  - all the trips with average speed $<$ 20Km/h and $>$ 120Km/h were dropped.
- The sample sizes were $n_{NHTS} = 447,493$ and $n_{SHRP2} = 3,458,826$.

# Research objective and benchmark

- **Objective:**
  To adjust for potential selection bias in SHRP2 dataset using National Household Travel Survey (NHTS) 2017 national probability as benchmark.
- 15 covariates were identified in common between NHTS and SHRP2 datasets.
  - **Participants' demographics**: gender, age, race, ethnicity, education, HH income, HH size, house ownership, birth country, job status
  - **Vehicle characteristics:** vehicle age, vehicle type, vehicle manufacturer, mileage
- To make the two datasets comparable, the following steps were taken:
  - Only drivers of the trips were kept in NHTS.
  - Trips with modes other than car/SUV/van/truck were removed.
  - trips for which public transportation was used were omitted.
  - all the trips with average speed $< 20$Km/h and $> 120$Km/h were dropped.
- The sample sizes were $n_{NHTS} = 447,493$ and $n_{SHRP2} = 3,458,826$.

# Research objective and benchmark

- **Objective:**
  To adjust for potential selection bias in SHRP2 dataset using National Household Travel Survey (NHTS) 2017 national probability as benchmark.
- 15 covariates were identified in common between NHTS and SHRP2 datasets.
  - **Participants' demographics**: gender, age, race, ethnicity, education, HH income, HH size, house ownership, birth country, job status
  - **Vehicle characteristics:** vehicle age, vehicle type, vehicle manufacturer, mileage
- To make the two datasets comparable, the following steps were taken:
  - Only drivers of the trips were kept in NHTS.
  - Trips with modes other than car/SUV/van/truck were removed.
  - trips for which public transportation was used were omitted.
  - all the trips with average speed $< 20$Km/h and $> 120$Km/h were dropped.
- The sample sizes were $n_{NHTS} = 447,493$ and $n_{SHRP2} = 3,458,826$.

# Research objective and benchmark

- **Objective:**
  To adjust for potential selection bias in SHRP2 dataset using National Household Travel Survey (NHTS) 2017 national probability as benchmark.

- 15 covariates were identified in common between NHTS and SHRP2 datasets.
  - **Participants' demographics**: gender, age, race, ethnicity, education, HH income, HH size, house ownership, birth country, job status
  - **Vehicle characteristics:** vehicle age, vehicle type, vehicle manufacturer, mileage

- To make the two datasets comparable, the following steps were taken:
  - Only drivers of the trips were kept in NHTS.
  - Trips with modes other than car/SUV/van/truck were removed.
  - trips for which public transportation was used were omitted.
  - all the trips with average speed $< 20$Km/h and $> 120$Km/h were dropped.

- The sample sizes were $n_{NHTS} = 447,493$ and $n_{SHRP2} = 3,458,826$.

# Research objective and benchmark

- **Objective:**
  To adjust for potential selection bias in SHRP2 dataset using National Household Travel Survey (NHTS) 2017 national probability as benchmark.

- 15 covariates were identified in common between NHTS and SHRP2 datasets.
  - **Participants' demographics**: gender, age, race, ethnicity, education, HH income, HH size, house ownership, birth country, job status
  - **Vehicle characteristics:** vehicle age, vehicle type, vehicle manufacturer, mileage

- To make the two datasets comparable, the following steps were taken:
  - Only drivers of the trips were kept in NHTS.
  - Trips with modes other than car/SUV/van/truck were removed.
  - trips for which public transportation was used were omitted.
  - all the trips with average speed $< 20$Km/h and $> 120$Km/h were dropped.

- The sample sizes were $n_{NHTS} = 447,493$ and $n_{SHRP2} = 3,458,826$.

# Research objective and benchmark

- **Objective:**
  To adjust for potential selection bias in SHRP2 dataset using National Household Travel Survey (NHTS) 2017 national probability as benchmark.

- 15 covariates were identified in common between NHTS and SHRP2 datasets.
  - **Participants' demographics**: gender, age, race, ethnicity, education, HH income, HH size, house ownership, birth country, job status
  - **Vehicle characteristics:** vehicle age, vehicle type, vehicle manufacturer, mileage

- To make the two datasets comparable, the following steps were taken:
  - Only drivers of the trips were kept in NHTS.
  - Trips with modes other than car/SUV/van/truck were removed.
  - trips for which public transportation was used were omitted.
  - all the trips with average speed $< 20$Km/h and $> 120$Km/h were dropped.

- The sample sizes were $n_{NHTS} = 447,493$ and $n_{SHRP2} = 3,458,826$.

# Research objective and benchmark

- **Objective:**
  To adjust for potential selection bias in SHRP2 dataset using National Household Travel Survey (NHTS) 2017 national probability as benchmark.

- 15 covariates were identified in common between NHTS and SHRP2 datasets.
  - **Participants' demographics**: gender, age, race, ethnicity, education, HH income, HH size, house ownership, birth country, job status
  - **Vehicle characteristics:** vehicle age, vehicle type, vehicle manufacturer, mileage

- To make the two datasets comparable, the following steps were taken:
  - Only drivers of the trips were kept in NHTS.
  - Trips with modes other than car/SUV/van/truck were removed.
  - trips for which public transportation was used were omitted.
  - all the trips with average speed $< 20$Km/h and $> 120$Km/h were dropped.

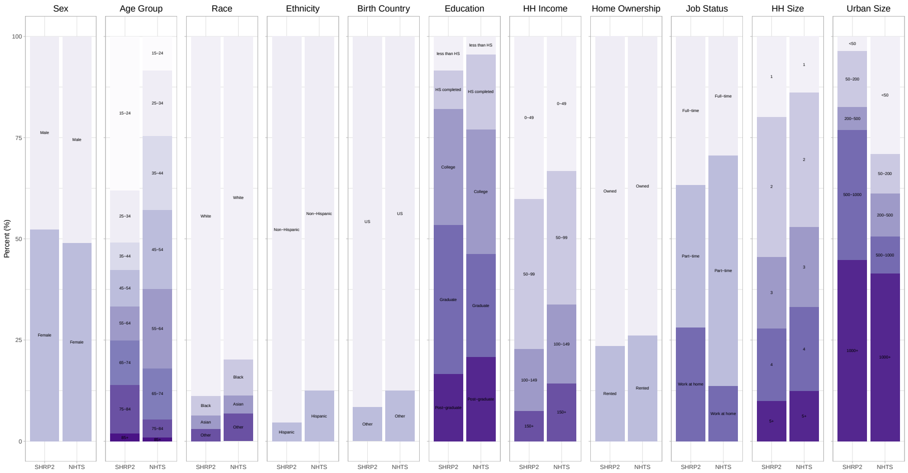
- The sample sizes were $n_{NHTS} = 447,493$ and $n_{SHRP2} = 3,458,826$.

# Research objective and benchmark

- **Objective:**
  To adjust for potential selection bias in SHRP2 dataset using National Household Travel Survey (NHTS) 2017 national probability as benchmark.
- 15 covariates were identified in common between NHTS and SHRP2 datasets.
  - **Participants' demographics**: gender, age, race, ethnicity, education, HH income, HH size, house ownership, birth country, job status
  - **Vehicle characteristics:** vehicle age, vehicle type, vehicle manufacturer, mileage
- To make the two datasets comparable, the following steps were taken:
  - Only drivers of the trips were kept in NHTS.
  - Trips with modes other than car/SUV/van/truck were removed.
  - trips for which public transportation was used were omitted.
  - all the trips with average speed $<$ 20Km/h and $>$ 120Km/h were dropped.
- The sample sizes were $n_{NHTS} = 447,493$ amd $n_{SHRP2} = 3,458,826$.

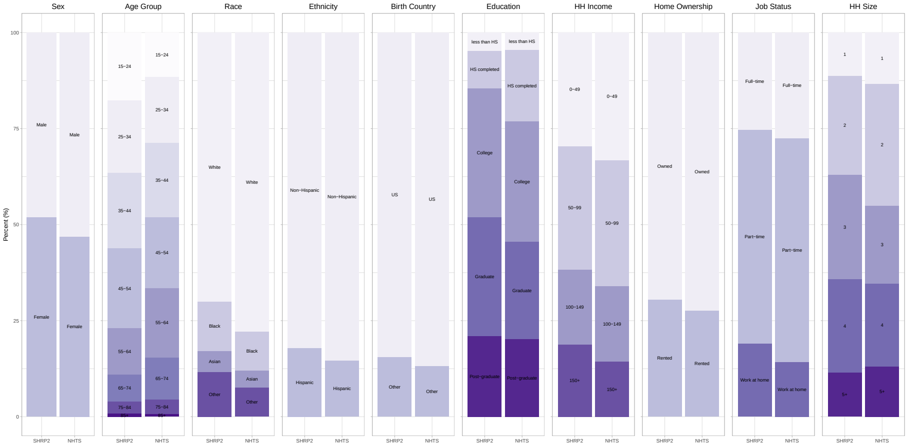
# Results on SHRP2

- Comparing the distribution of demographic covariates unweighted SHRP2 vs weighted NHTS



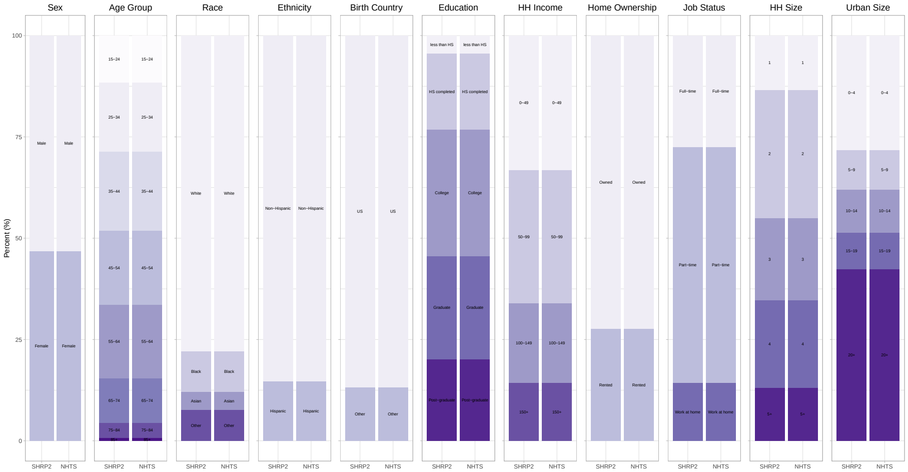
# Results on SHRP2

- Comparing the distribution of demographic covariates pseudo-weighted SHRP2 (PW) vs weighted NHTS

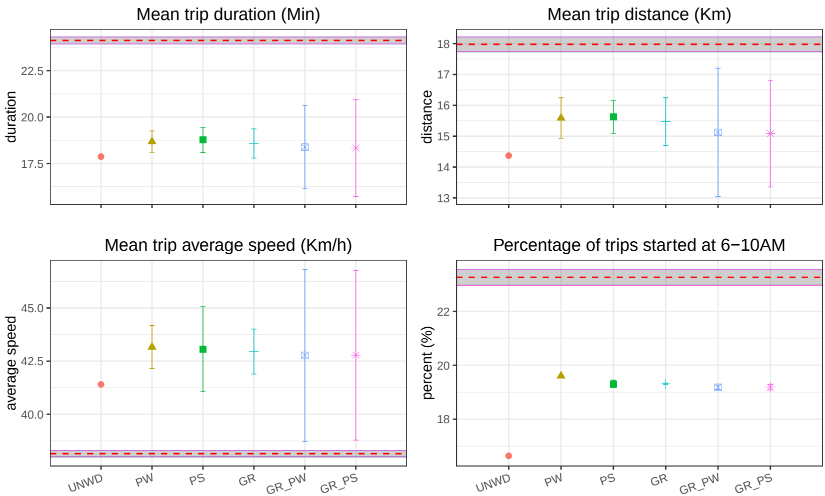


# Results on SHRP2

- Comparing the distribution of demographic covariates pseudo-weighted SHRP2 (GREG) vs weighted NHTS

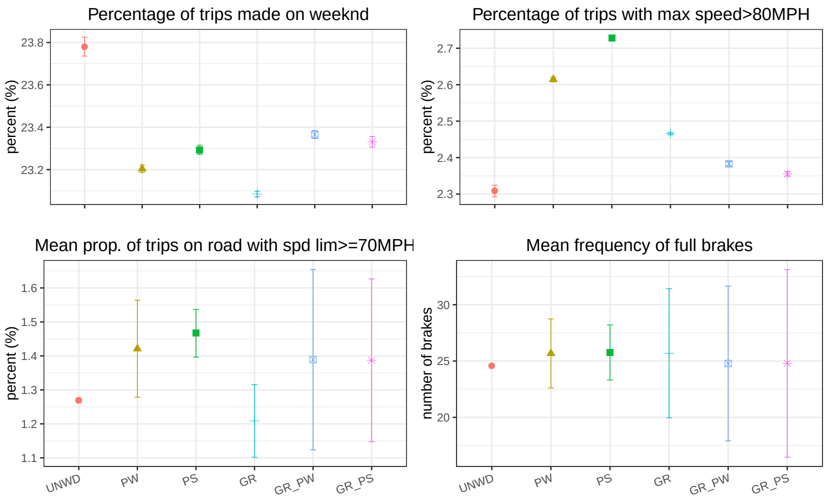


- Comparing adjusted point estimates of some trip-related outcome variables in SHRP2 with weighted estimates in NHTS



# Result on SHRP2

- Adjusted point estimates and associated 95%CIs for some SHRP2-specific outcome variables



# Conclusion

- We proposed doubly robust weighting by combining PW with GREG.
- Findings from SHRP2 data reflects substantial measurement differences between individual's report and machine's reports.
- Simulation study demonstrates doubly robustness of our adjustments.
- PW outperforms PS when selection mechanism is significantly different in the two samples.
- The Jackknife variance estimator tends to underestimate the variance especially when the outcome variable is binary.
- One might use sandwich-type methods for variance estimation, which is computationally more efficient.
- An alternative approach can be built by imputing the outcome variable for units in the reference survey (Chen et al, 2018).
- For high-dimensional data, we recommend using *LASSO* technique when fitting PS and GREG models (Chen et al. 2019).

**Thanks for your attention!**

Email: mrelliot@umich.edu

# References

Elliott, M., Valliant, R. (2017)
Inference for nonprobability samples
*Statistical Science* 32(2), 249–264.

Wu, Changbao  Sitter, Randy R. (2001)
A model-calibration approach to using complete auxiliary information from survey data
*Journal of the American Statistical Association* 96(453), 185–193.

Robins, J. M., Rotnitzky, A.,  Zhao, L. P. (1994)
Estimation of regression coefficients when some regressors are not always observed
*Journal of the American statistical Association* 89(427), 846-866.

Elliott, M., Resler, A., Flannagan, C., Rupp, J. (2010)
Appropriate analysis of CIREN data: Using NASS-CDS to reduce bias in estimation of injury risk factors in passenger vehicle crashes
*Accident analysis and prevention* 42(2), 530–539.

Deville, J. C.,  Srndal, C. E. (1992)
Calibration estimators in survey sampling
*Journal of the American statistical Association* 87(418), 376-382.

# References


Czajka, J. L., Hirabayashi, S. M., Little, R. J., Rubin, D. B. (1992)
Projecting from advance data using propensity modeling: An application to income and tax statistics
*Journal of Business Economic Statistics* 10(2), 117-131.

Valliant, R., Dever, J. A. (2011).
Estimating propensity adjustments for volunteer web surveys.
*Sociological Methods Research*. 40(1), 105-137.

Chen, Y., Li, P., Wu, C. (2018).
Doubly robust inference with non-probability survey samples.
*arXiv preprint arXiv*. 1805.06432.

Schonlau, M., Van Soest, A., Kapteyn, A., Couper, M. (2009).
Selection bias in web surveys and the use of propensity scores.
*Sociological Methods Research*. 37(3), 291-318.

Lee, S. (2006).
Propensity score adjustment as a weighting scheme for volunteer panel web surveys.
*Journal of official statistics*. 22(2), 329.