RSG
the science of insight

# Gen3 Model Development Project

## Travel Forecasting Subcommittee Meeting

September 18, 2020

IN PARTNERSHIP WITH

BASELINE MOBILITY

Metropolitan Washington Council of Governments

# Discussion Topics

- Synthetic Population Creation
- On-board Survey Data Review

# Synthetic Population Generation

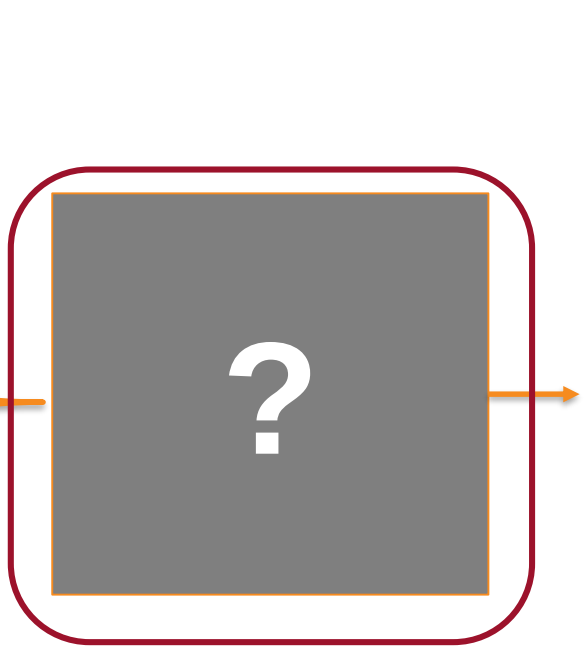
# Why is Synthetic Population Needed?

- ActivitySim simulates travel choices for individual households and persons
- This requires the generation of a ‘synthetic population’ that represents the base or future year population to be modeled
- Data available:
  - Population sample (PUMS, HH Survey)
  - Distribution of person and HH attributes (Census Summary File and/or Agency forecasts)

# Problem Statement

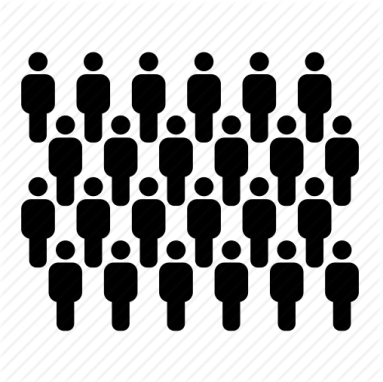
Seed Sample
(PUMS)

Marginal Controls
(Census Distributions)




Data Fitting Technique
a.k.a.
Population Synthesizer

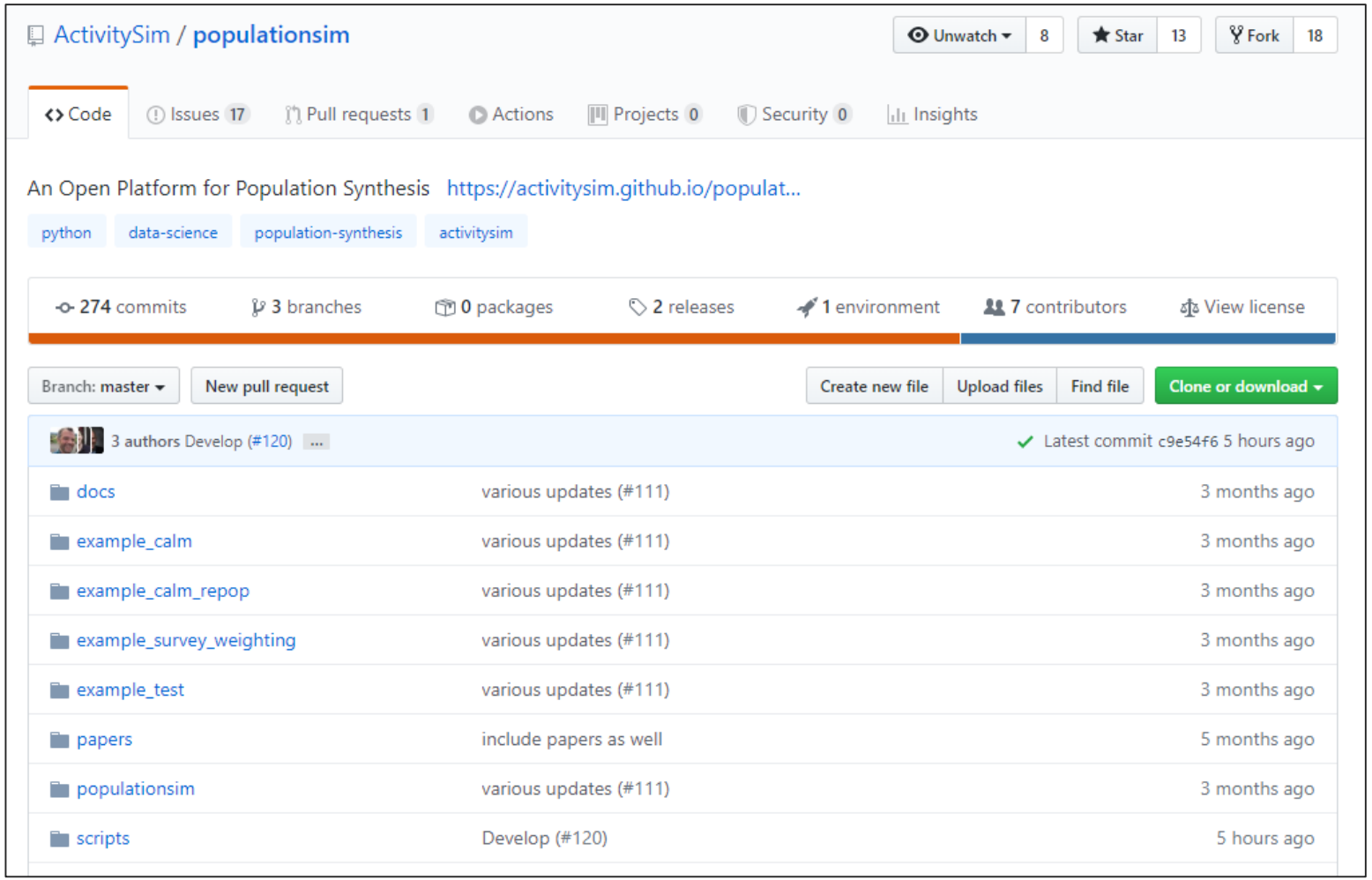
Synthetic Population



# Terminology

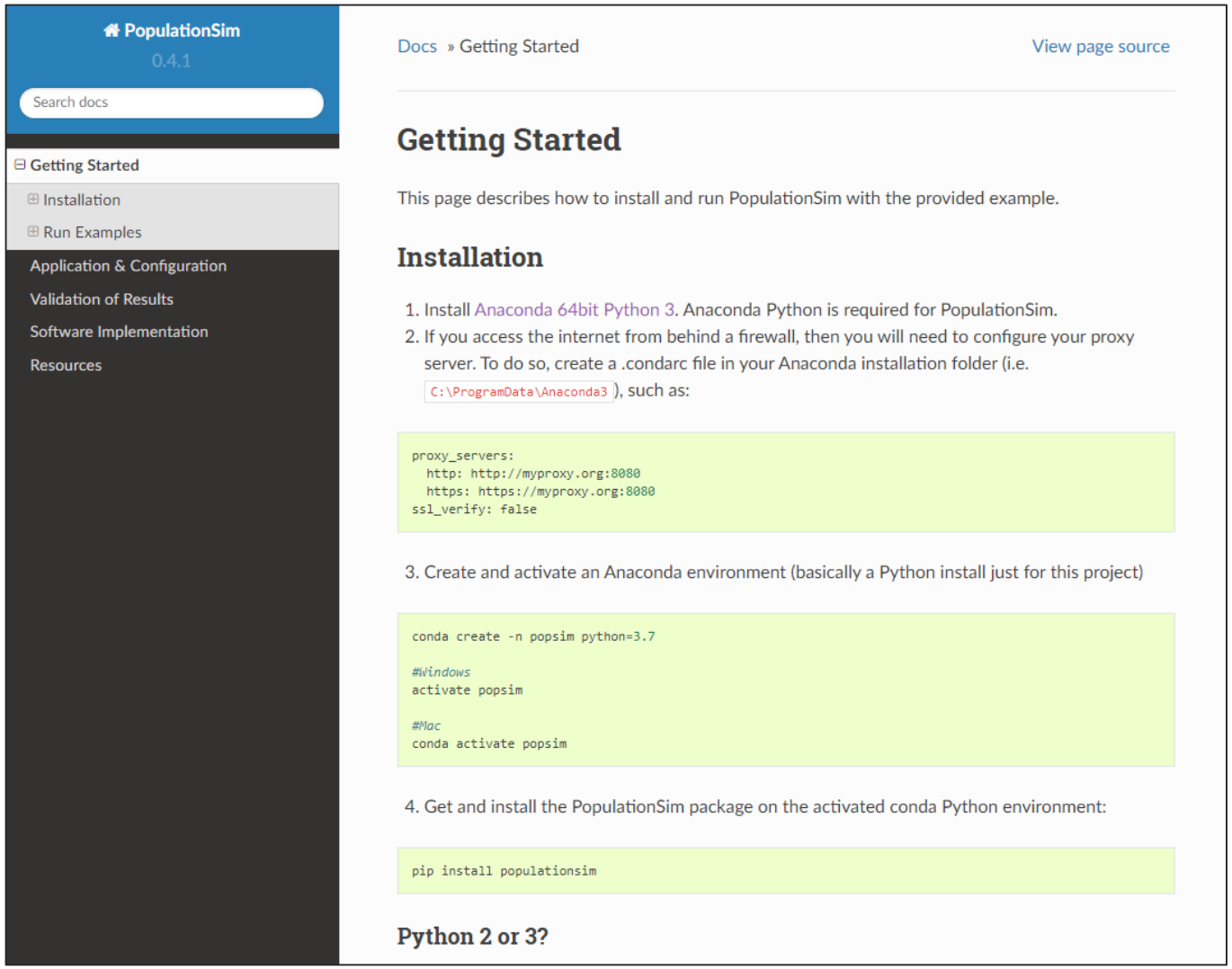
- ***Geography***
  - Geographic resolution or the unit(s) of geography (e.g., MAZ, TAZ, TRACT)
- ***Seed Sample***
  - Sample of household and person records representing a population [e.g., American Community Survey - Public Use Microdata Sample (PUMS)]
- ***Marginal Controls [or Targets]***
  - Aggregate distribution of households or persons for a control variable at a geographic resolution [e.g., Households by Income Quintiles in each TAZ]
- ***Synthetic Population***
  - Set of household and person records representing a region's entire population
- ***Balancing/Fitting***
  - Process of reweighting a disaggregate sample of agents to satisfy aggregate constraints
- ***Integerization***
  - Process of converting the floating point weights to integer weights while satisfying the aggregate constraints

# PopulationSim Software




https://github.com/ActivitySim/populationsim

# PopulationSim Wiki




https://activitysim.github.io/populationsim/

# PopulationSim – Key Advantages

- Algorithmic advantages including
  - Entropy maximization-based list balancing
  - Linear programming-based population generation
- Specify controls at multiple levels of geography

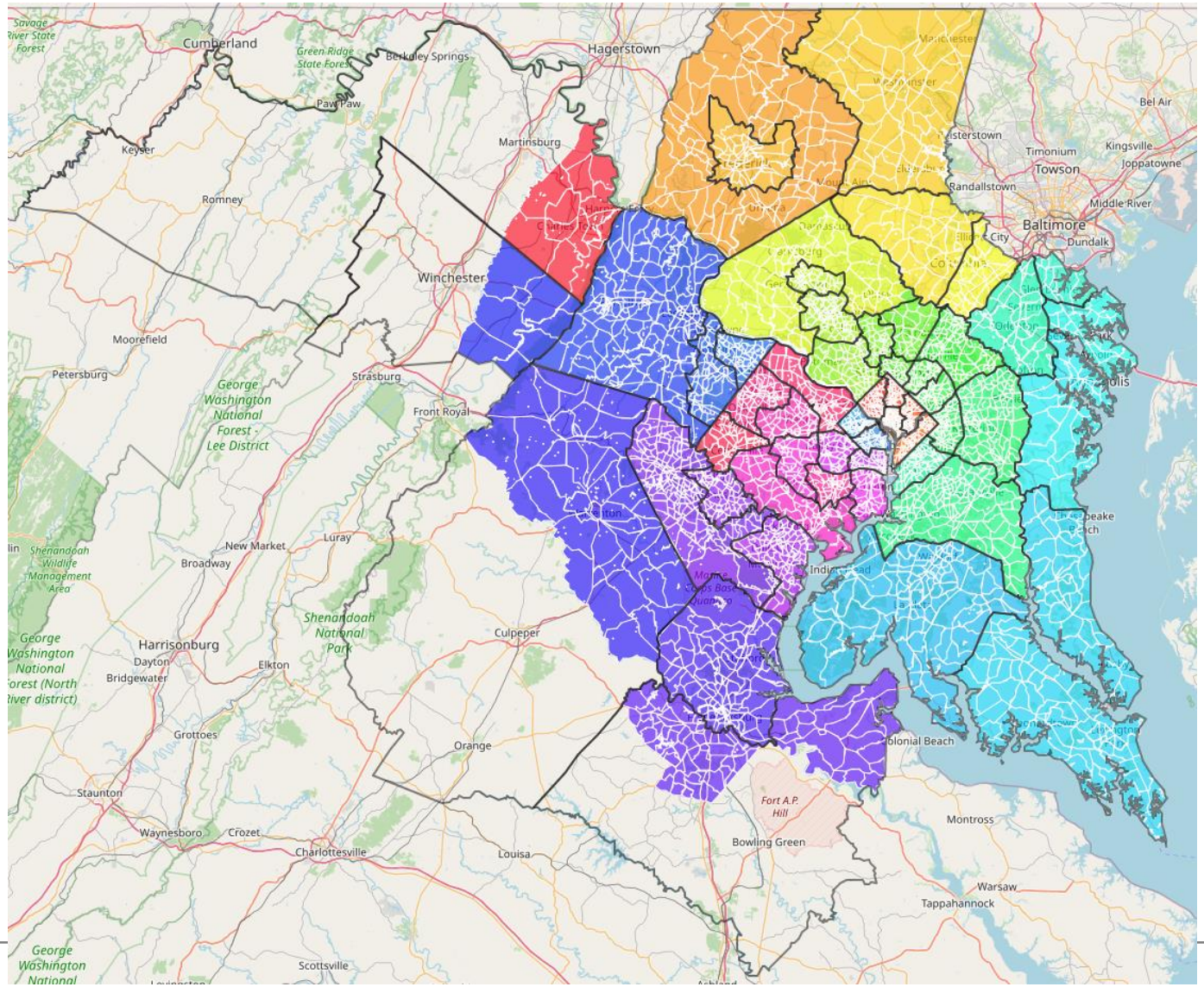
- Implemented in ActivitySim framework and planned for consortium management going forward
- Fast runtime
  - DVRPC setup runs 15x times faster than PopGen
- Re-populate feature
  - Generate synthetic population for traffic impact studies and similar small area applications

# MWCOG PopulationSim Inputs

Three main inputs:

- Disaggregate household and person samples
  - *2014-18 5-Year ACS PUMS*
- TAZ-level marginal controls
  - Tract-level Census distributions
  - MWCOG land use data (Round 9.1a Cooperative Forecasts)
- Geographic crosswalk (TAZ-to-PUMA)
  - TAZs assigned to PUMAs based on maximum overlap

# TAZ – PUMA Crosswalk



# MWCOG PopulationSim Controls

| CONTROL VARIABLE | CATEGORIES | PUMS FIELD | CONTROL SOURCE | GEOG |
|---|---|---|---|---|
| **Household Variables** | | | | |
| **Total number of households** | | WGTP | Round 9.1 Cooperative Forecast | TAZ |
| **Household Size** | 1, 2, 3, 4, 5+ | NP | 2018 ACS 5-year. Census Tract [Table S2501] | TAZ |
| **Household Income*** | 0-$30K, $30K-$60K, $60K-$100K, $100k-$150K, $150K+ | HINCP | 2018 ACS 5-year. Census Tract [Table B19001] | TAZ |
| **Number of Workers** | 0, 1, 2, 3+ | ESR | 2018 ACS 5-year. Census Tract [Table B08202] | TAZ |
| **Presence of Children** | 0, 1 | HUPAC | 2018 ACS 5-year. Census Tract [Table S1101] | TAZ |
| **Person Variables** | | | | |
| **Person Age** | 0-4, 5-19, 20-34, 35-64, 65+ | AGEP | 2018 ACS 5-year. Census Tract [Table S0101] | TAZ |
| **Person Race/Ethnicity** | White, Hispanic, Black, Asian, Other | HISP, RAC1P | 2018 ACS 5-year. Census Tract [Table DP05] | TAZ |

* Categories subject to change

# MWCOG PopulationSim Controls (GQ)

| CONTROL VARIABLE | CATEGORIES | PUMS FIELD | SOURCE | GEOG |
|---|---|---|---|---|
| **Total GQ units** | | TYPE | Round 9.1 Cooperative Forecast | TAZ |
| **GQ Type** | University, Other Non-Institutional | SCHG & TYPE | 2018 ACS 5-year. Census Tract [Table B26101] | TAZ |

# MWCOG PopulationSim Implementation

**Data Download**
- Auto download Census data and shape files
- Use Census API functionality

**Data preparation**
- Seed Sample
- Marginal Controls
- Geographic crosswalk

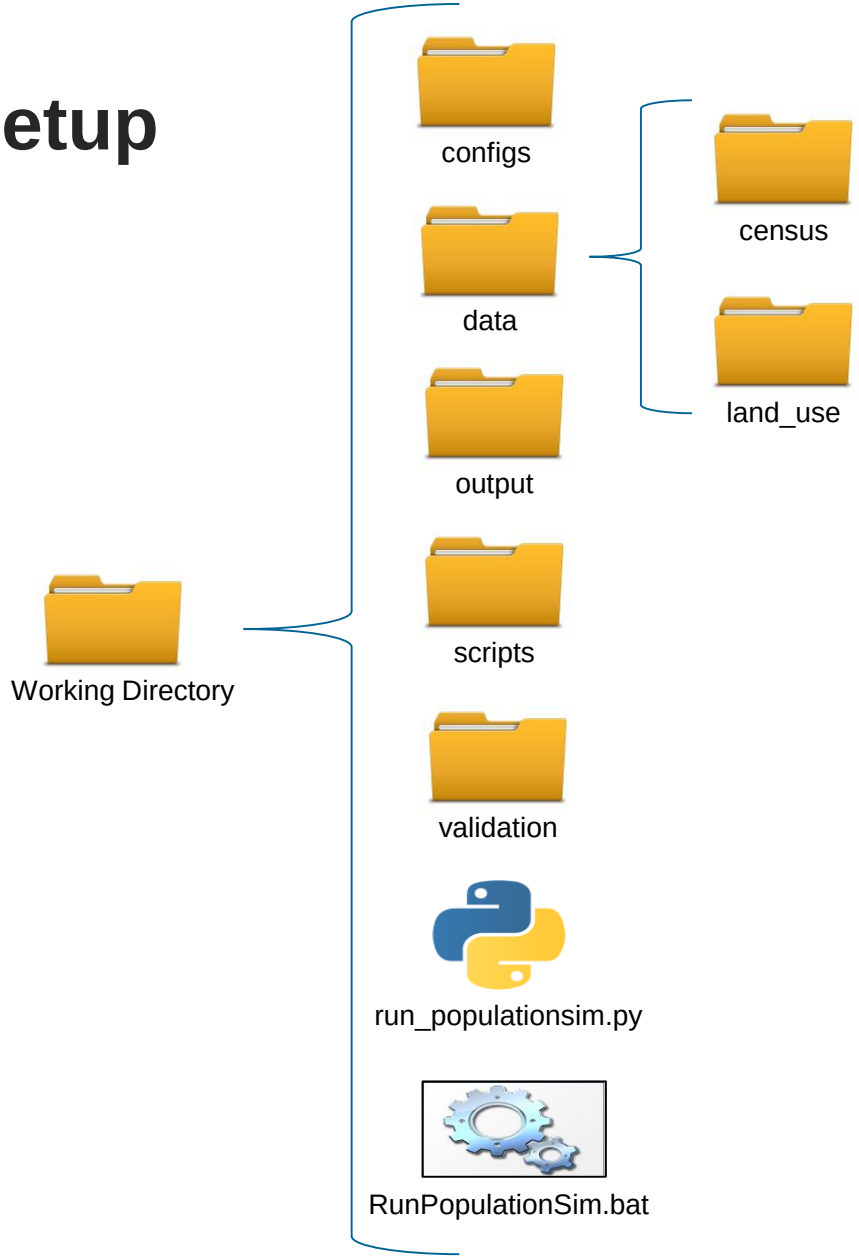
**Run PopulationSim**
- User specified controls and settings
- Separate household and GQ runs

**Post-processing**
- Combine GQ and residential outputs

**Validation**
- Run standard PopulationSim validation script
- Mean percentage difference, STDEV, uniformity distribution

Python-based implementation

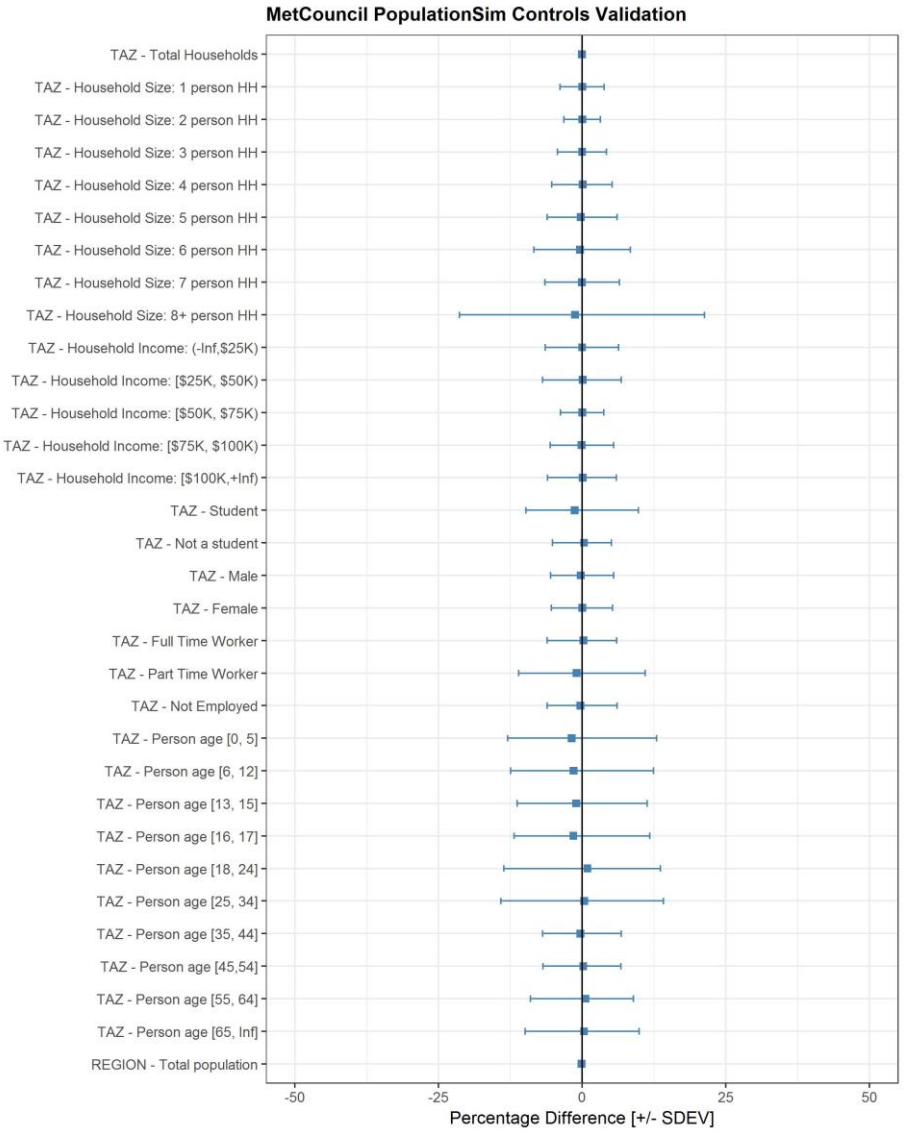
# Directory Setup



# Validation Techniques

- Synthetic Population vs Marginal Controls
  - Average percentage difference
  - Standard deviation
- Expansion factor distribution
  - Distribution of ratio of initial and final weight
- Uncontrolled variable distributions
- Data checks
  - Sample vs controls
  - Consistency among controls

# Matching Controls



# On-Board Surveys

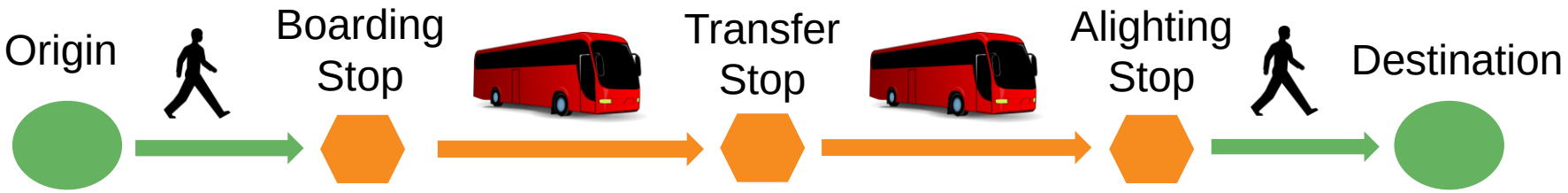
# Uses of on-board survey for model development

- Understand transit markets:
  - Origin and destination location and purpose (code trip purpose, impute tour purpose)
  - Access and egress mode (walk, bike, PNR, KNR, other)
  - Socio-demographic: autos owned/auto sufficiency, income, person type (worker/student), etc.
- Understand transfer rates
  - Within and between modes/operators
  - Convert boardings to ‘linked trips’
- Test transit network & path-building parameters
- Model estimation, calibration and validation targets
- Other
  - Understand “external” transit trips, fare, visitor trips, etc.

# The ideal on-board survey…

- Has a sample plan based on observed boardings by route, direction, and time-of-day
- Collects the following data
  - Origin and destination location, first boarding location
  - Origin and destination purpose
  - Access and egress mode
  - Route sequence from first boarding to final alighting
  - Relevant household and person information
  - Fare paid
- Implemented with in-person, random interviewing
  - using tablet computers with integrated mapping functionality
- Is carefully expanded to boardings
  - by station on/off for fixed guideway, boarding stop group for high volume routes, total boardings for other routes

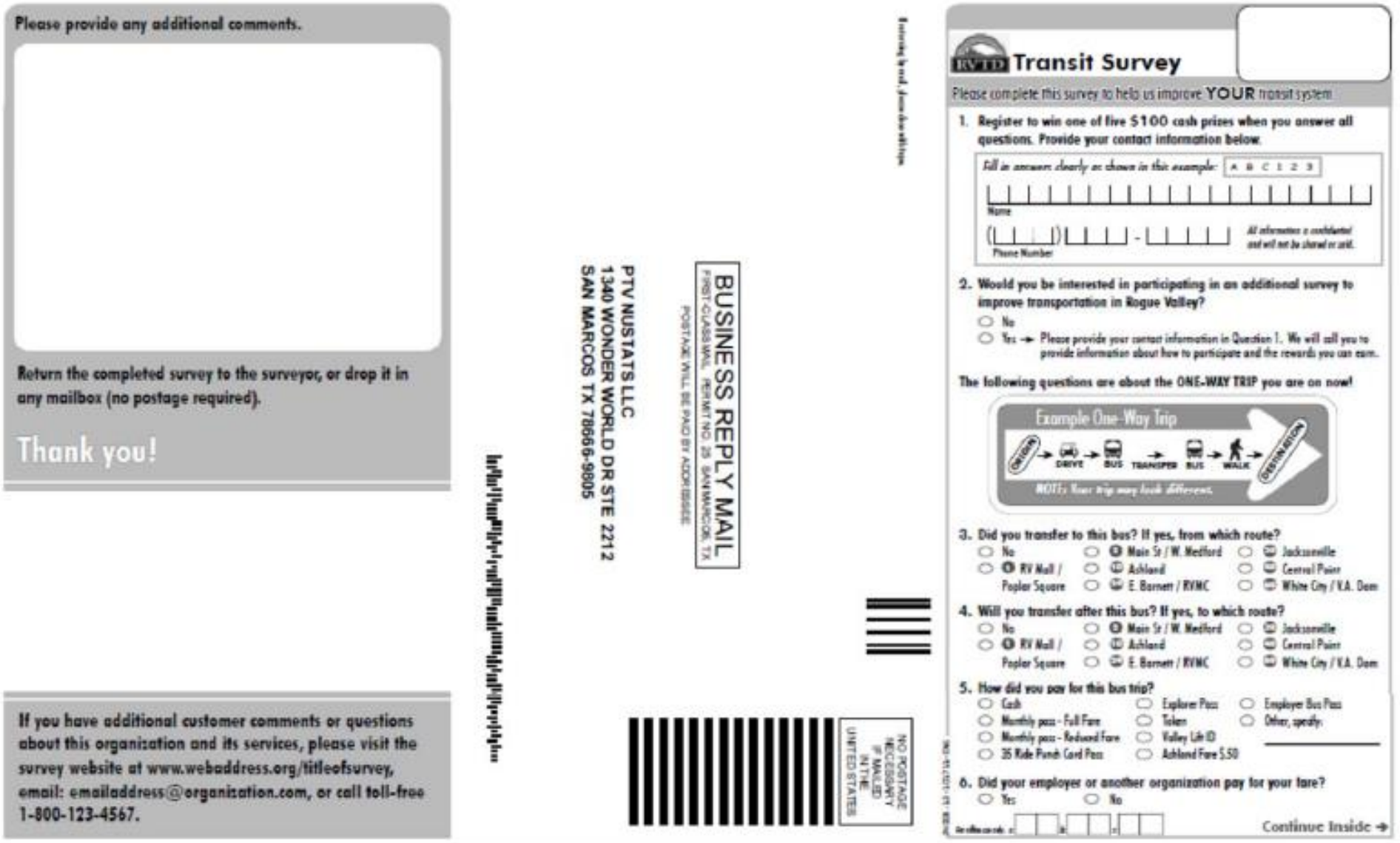
# Transit Trip Data




Origin purpose
Origin place name
Origin location
Origin departure time
Access mode
Access distance
First boarding stop location
Route\run\direction surveyed on
Boarding stop on surveyed route
Alighting stop off surveyed route

Route number transferred from
Route number transferred to
Total number of routes
Last alighting stop location
Egress mode
Egress distance
Destination purpose
Destination place name
Destination location

# RVTD On-Board Survey (2011)



Please provide any additional comments.

Return the completed survey to the surveyor, or drop it in any mailbox (no postage required).

Thank you!

If you have additional customer comments or questions about this organization and its services, please visit the survey website at www.webaddress.org/titleofsurvey, email: emailaddress@organization.com, or call toll-free 1-800-123-4567.

PTV NUSTATS LLC
1340 WONDER WORLD DR STE 2212
SAN MARCOS TX 78666-9805

BUSINESS REPLY MAIL
FIRST-CLASS MAIL PERMIT NO. 25 SAN MARCOS, TX
POSTAGE WILL BE PAID BY ADDRESSEE

NO POSTAGE NECESSARY IF MAILED IN THE UNITED STATES

RVTD Transit Survey

Please complete this survey to help us improve YOUR transit system.

1. Register to win one of five $100 cash prizes when you answer all questions. Provide your contact information below.

Fill in answers clearly as shown in this example: A B C 1 2 3

Name

Phone Number

All information is confidential and will not be shared or sold.

2. Would you be interested in participating in an additional survey to improve transportation in Rogue Valley?

○ No
○ Yes → Please provide your contact information in Question 1. We will call you to provide information about how to participate and the rewards you can earn.

The following questions are about the ONE-WAY TRIP you are on now!

Example One-Way Trip

ORIGIN → DRIVE → BUS → TRANSFER → BUS → WALK → DESTINATION

NOTE: Your trip may look different.

3. Did you transfer to this bus? If yes, from which route?

○ No
○ RV Mall / Poplar Square
○ Main St / W. Medford
○ Ashland
○ E. Barnett / RVMC
○ Jacksonville
○ Central Point
○ White City / V.A. Dom

4. Will you transfer after this bus? If yes, to which route?

○ No
○ RV Mall / Poplar Square
○ Main St / W. Medford
○ Ashland
○ E. Barnett / RVMC
○ Jacksonville
○ Central Point
○ White City / V.A. Dom

5. How did you pay for this bus trip?

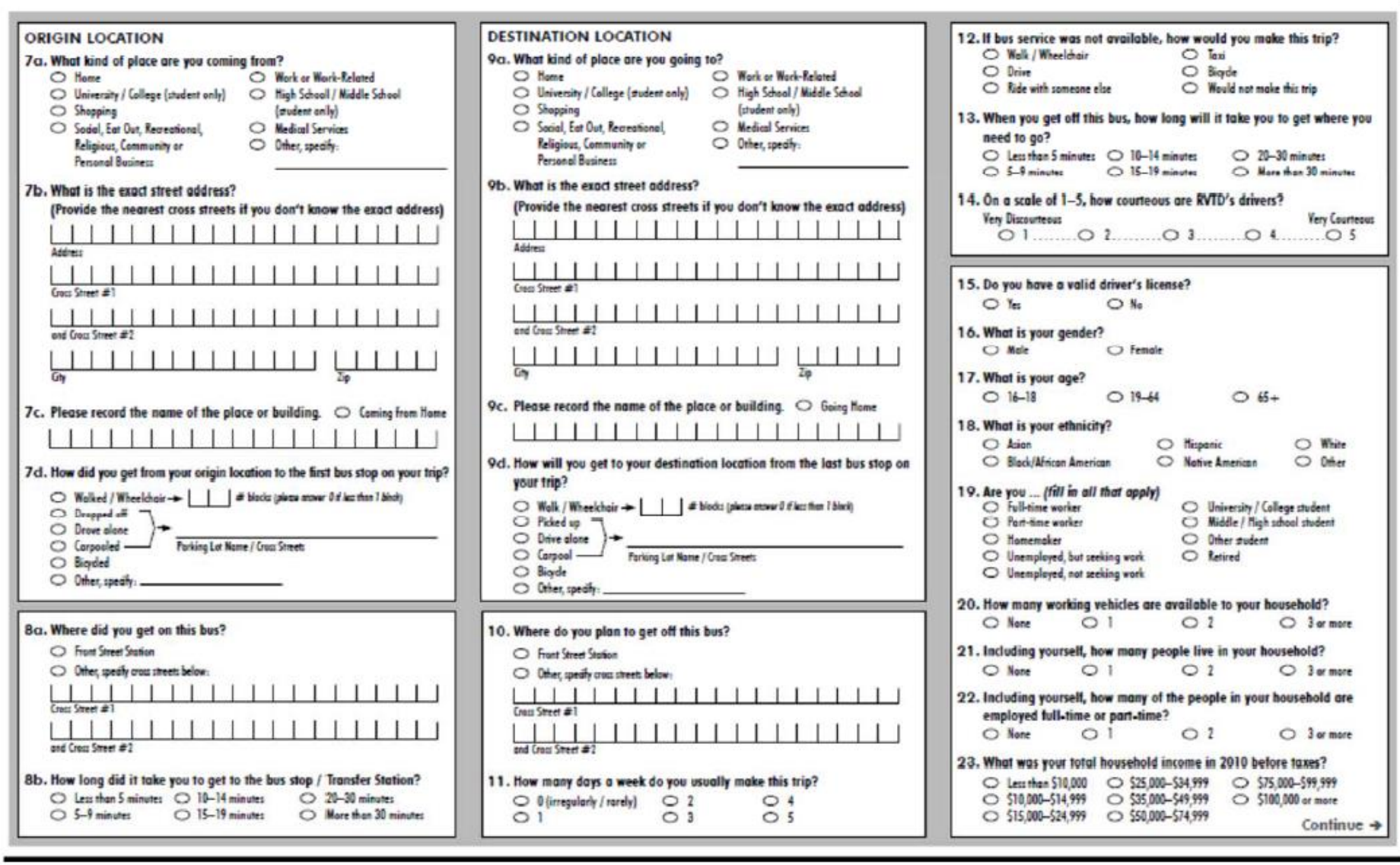
○ Cash
○ Monthly pass - Full Fare
○ Monthly pass - Reduced Fare
○ 25 Ride Punch Card Pass
○ Explorer Pass
○ Token
○ Valley Life ID
○ Ashland Fare $.50
○ Employer Bus Pass
○ Other, specify: _____

6. Did your employer or another organization pay for your fare?

○ Yes ○ No

Continue Inside →

# RVTD On-Board Survey (2011)



**ORIGIN LOCATION**

**7a. What kind of place are you coming from?**

- ○ Home
- ○ University / College (student only)
- ○ Shopping
- ○ Social, Eat Out, Recreational, Religious, Community or Personal Business
- ○ Work or Work-Related
- ○ High School / Middle School (student only)
- ○ Medical Services
- ○ Other, specify: _____

**7b. What is the exact street address?**
**(Provide the nearest cross streets if you don't know the exact address)**

Address

Cross Street #1

and Cross Street #2

City | Zip

**7c. Please record the name of the place or building.** ○ Coming from Home

**7d. How did you get from your origin location to the first bus stop on your trip?**

- ○ Walked / Wheelchair → [ ] # blocks (please answer 0 if less than 1 block)
- ○ Dropped off
- ○ Drove alone
- ○ Carpooled
- ○ Bicycled
- ○ Other, specify: _____

→ Parking Lot Name / Cross Streets

**8a. Where did you get on this bus?**

- ○ Front Street Station
- ○ Other, specify cross streets below:

Cross Street #1

and Cross Street #2

**8b. How long did it take you to get to the bus stop / Transfer Station?**

- ○ Less than 5 minutes
- ○ 5–9 minutes
- ○ 10–14 minutes
- ○ 15–19 minutes
- ○ 20–30 minutes
- ○ More than 30 minutes

**DESTINATION LOCATION**

**9a. What kind of place are you going to?**

- ○ Home
- ○ University / College (student only)
- ○ Shopping
- ○ Social, Eat Out, Recreational, Religious, Community or Personal Business
- ○ Work or Work-Related
- ○ High School / Middle School (student only)
- ○ Medical Services
- ○ Other, specify: _____

**9b. What is the exact street address?**
**(Provide the nearest cross streets if you don't know the exact address)**

Address

Cross Street #1

and Cross Street #2

City | Zip

**9c. Please record the name of the place or building.** ○ Going Home

**9d. How will you get to your destination location from the last bus stop on your trip?**

- ○ Walk / Wheelchair → [ ] # blocks (please answer 0 if less than 1 block)
- ○ Picked up
- ○ Drive alone
- ○ Carpool
- ○ Bicycle
- ○ Other, specify: _____

→ Parking Lot Name / Cross Streets

**10. Where do you plan to get off this bus?**

- ○ Front Street Station
- ○ Other, specify cross streets below:

Cross Street #1

and Cross Street #2

**11. How many days a week do you usually make this trip?**

- ○ 0 (irregularly / rarely)
- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5

**12. If bus service was not available, how would you make this trip?**

- ○ Walk / Wheelchair
- ○ Drive
- ○ Ride with someone else
- ○ Taxi
- ○ Bicycle
- ○ Would not make this trip

**13. When you get off this bus, how long will it take you to get where you need to go?**

- ○ Less than 5 minutes
- ○ 5–9 minutes
- ○ 10–14 minutes
- ○ 15–19 minutes
- ○ 20–30 minutes
- ○ More than 30 minutes

**14. On a scale of 1–5, how courteous are RVTD's drivers?**

Very Discourteous ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 Very Courteous

**15. Do you have a valid driver's license?**

- ○ Yes
- ○ No

**16. What is your gender?**

- ○ Male
- ○ Female

**17. What is your age?**

- ○ 16–18
- ○ 19–64
- ○ 65+

**18. What is your ethnicity?**

- ○ Asian
- ○ Black/African American
- ○ Hispanic
- ○ Native American
- ○ White
- ○ Other

**19. Are you ... *(fill in all that apply)***

- ○ Full-time worker
- ○ Part-time worker
- ○ Homemaker
- ○ Unemployed, but seeking work
- ○ Unemployed, not seeking work
- ○ University / College student
- ○ Middle / High school student
- ○ Other student
- ○ Retired

**20. How many working vehicles are available to your household?**

- ○ None
- ○ 1
- ○ 2
- ○ 3 or more

**21. Including yourself, how many people live in your household?**

- ○ None
- ○ 1
- ○ 2
- ○ 3 or more

**22. Including yourself, how many of the people in your household are employed full-time or part-time?**

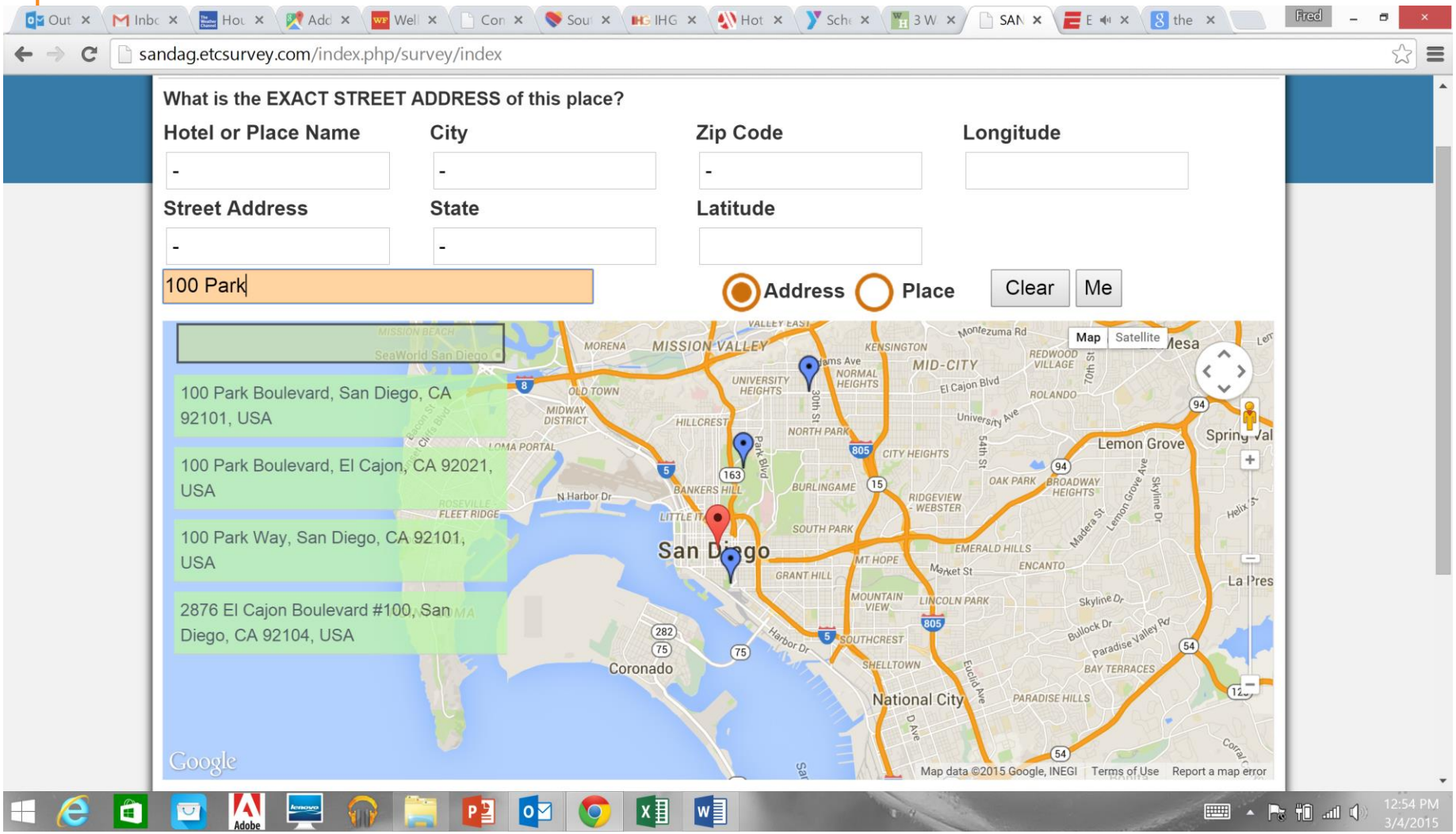
- ○ None
- ○ 1
- ○ 2
- ○ 3 or more

**23. What was your total household income in 2010 before taxes?**

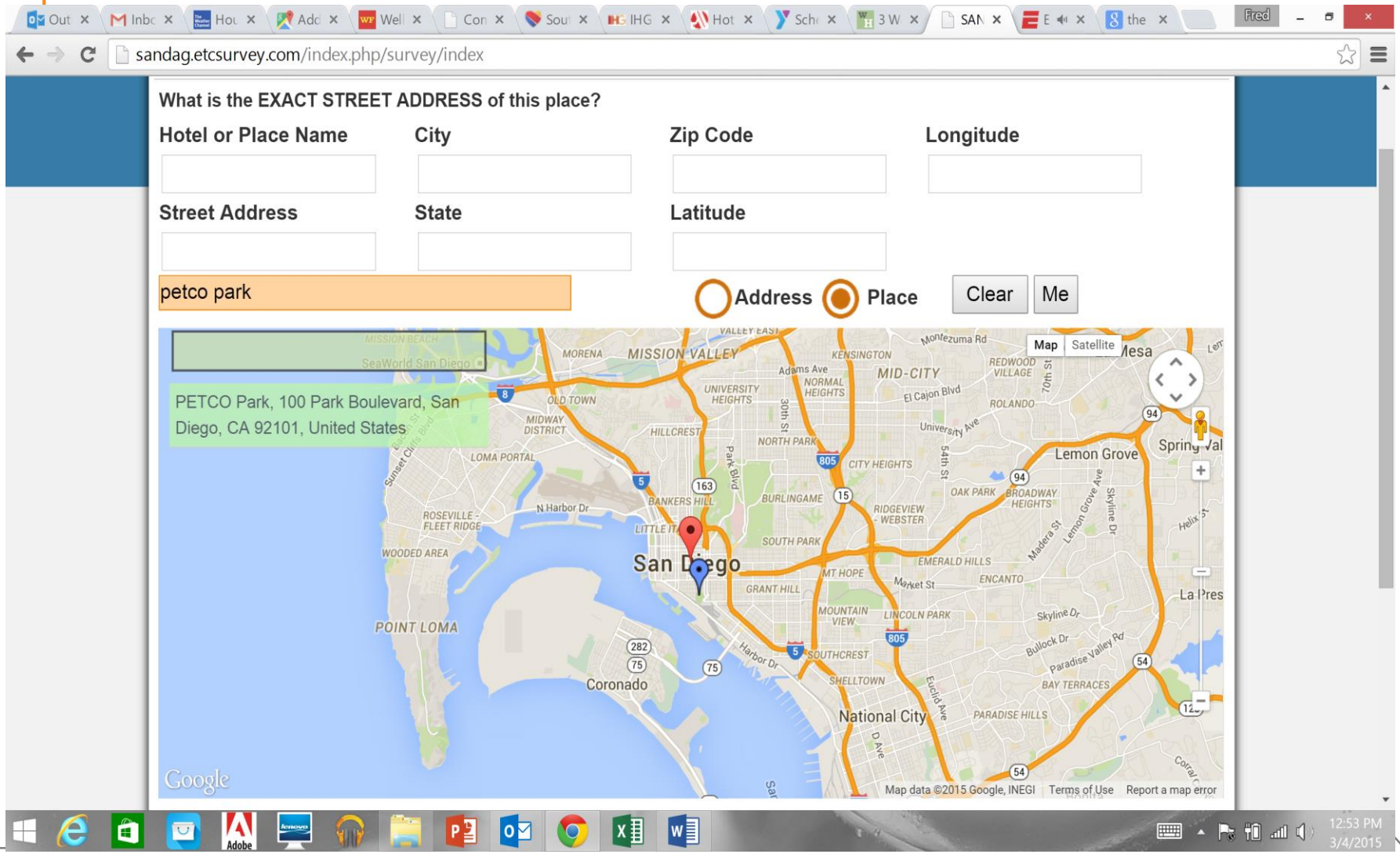
- ○ Less than $10,000
- ○ $10,000–$14,999
- ○ $15,000–$24,999
- ○ $25,000–$34,999
- ○ $35,000–$49,999
- ○ $50,000–$74,999
- ○ $75,000–$99,999
- ○ $100,000 or more

Continue →

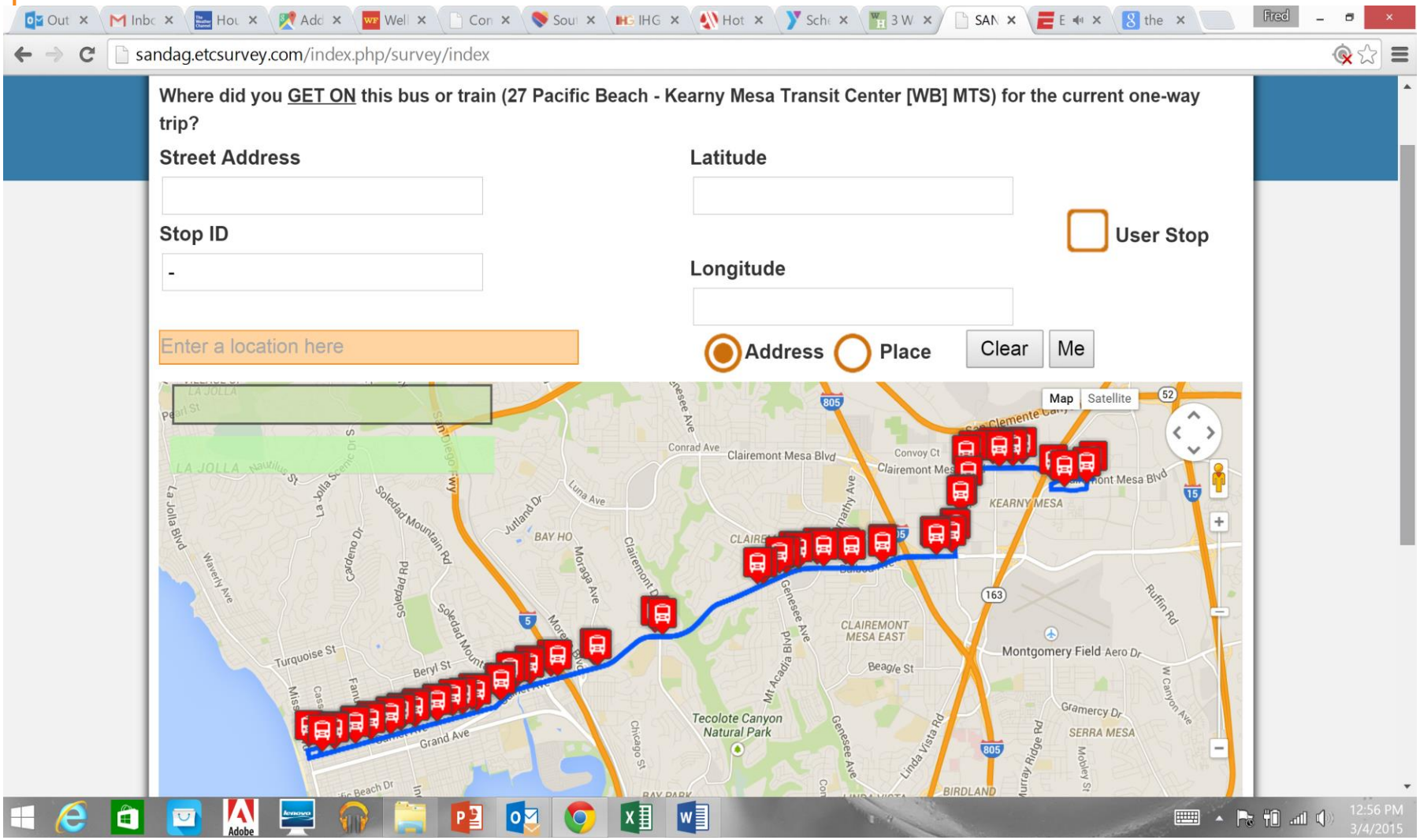
# Tablet Survey Screenshots



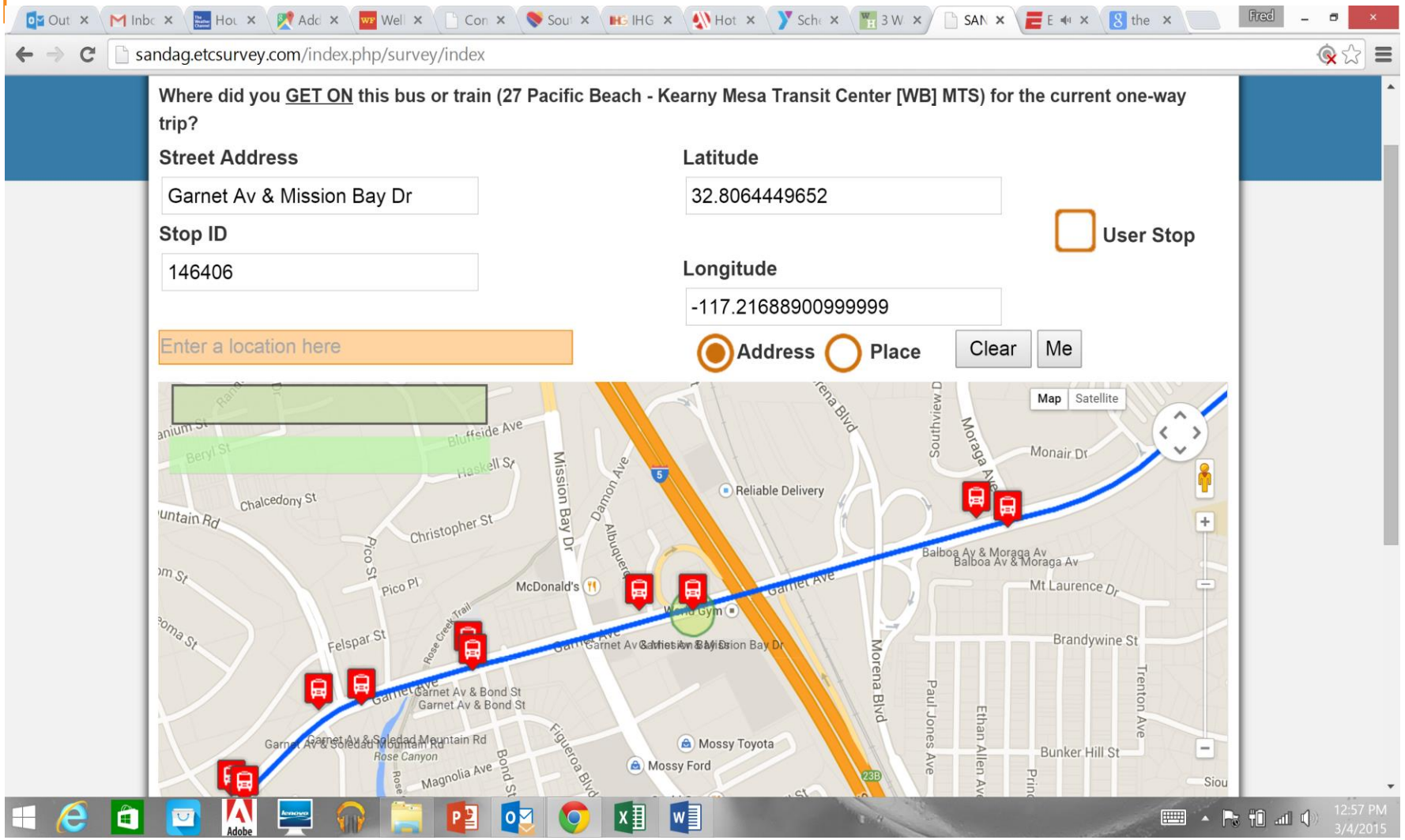
# Tablet Survey Screenshots



# Tablet Survey Screenshots



# Tablet Survey Screenshots



# CTA OD (Routes Used)

**Which other CTA buses or trains did you ride on your trip?**

First, I rode the... 12 - Roosevelt

Then, I rode the... Red Line

Then, I rode the... 66 - Chicago

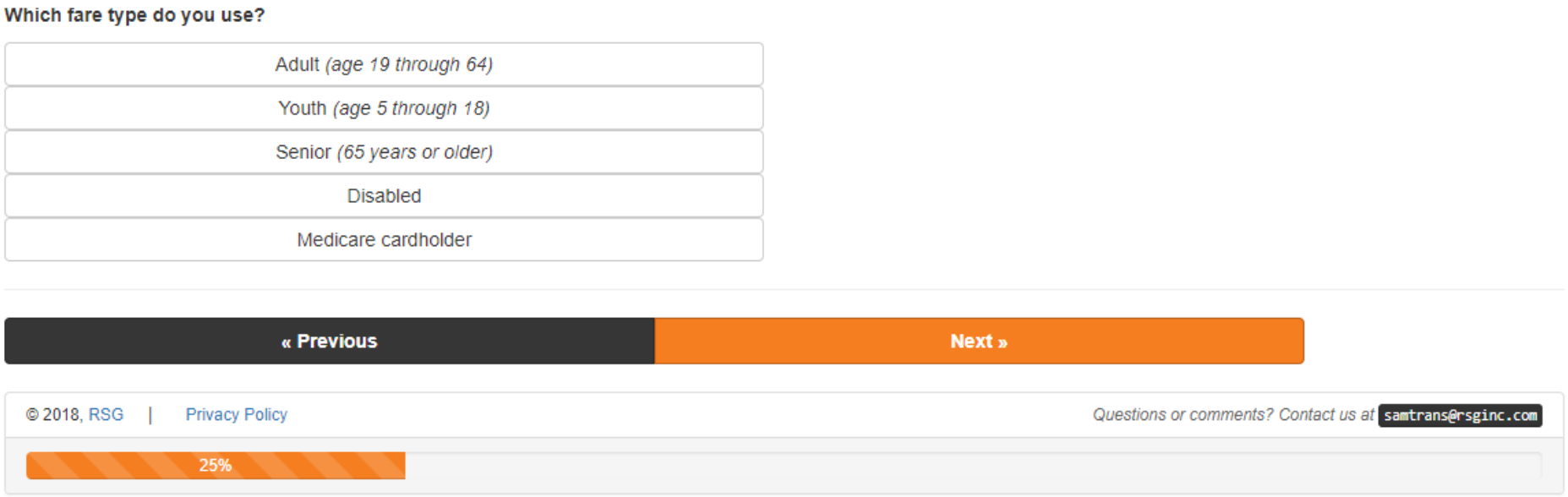
63 - 63rd
63W - West 63rd
65 - Grand
66 - Chicago
67 - 67th-69th-71st
68 - Northwest Highway
70 - Division
71 - 71st/South Shore
72 - North
73 - Armitage

« Previous

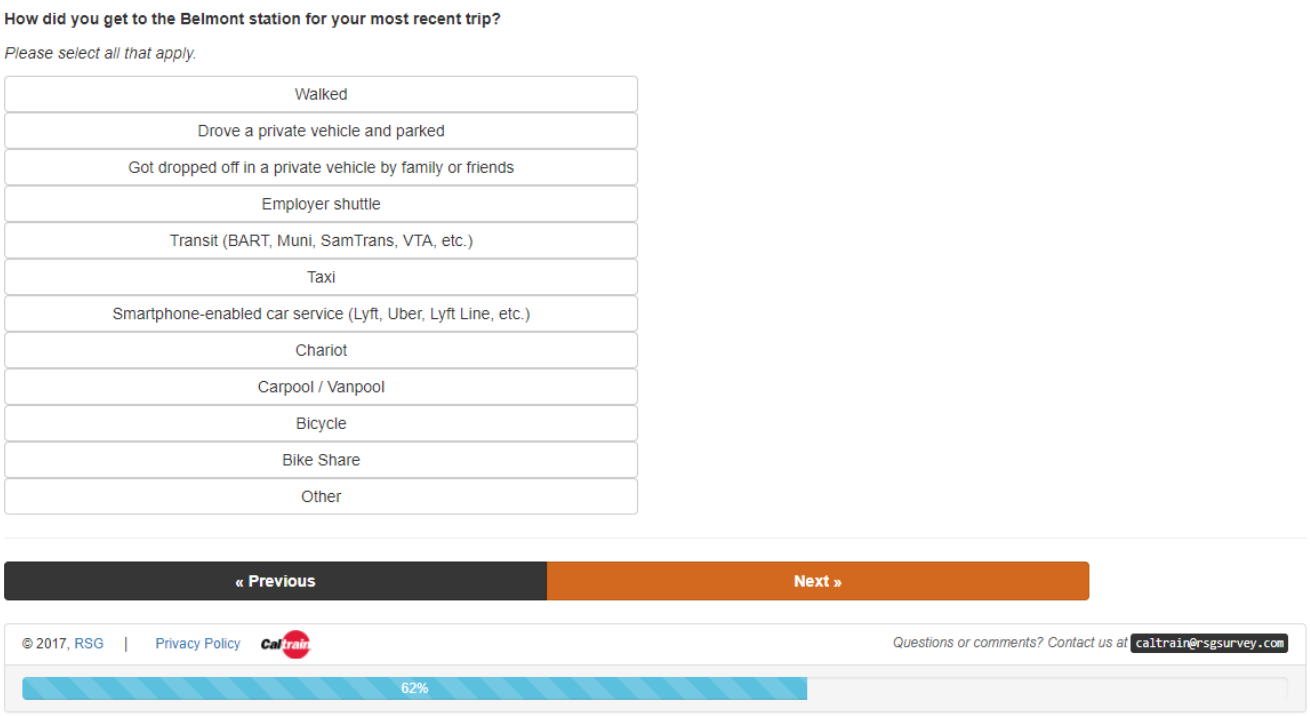
© 2017, RSG | Privacy Policy

*Questions or comments? Contact us at* CTAtravelsurvey@rsginc.com

# Fare Type Question (SamTrans Survey)



# Access Mode (Caltrain Survey)



# Survey Instruments/Documentation Reviewed

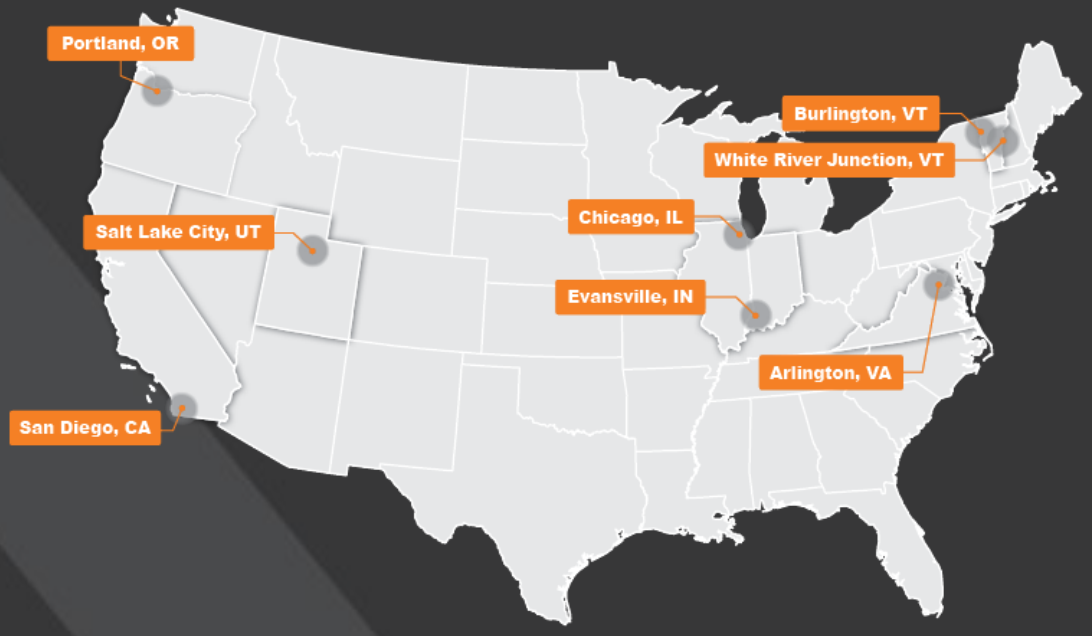
- 2016 Metrobus and Metrorail Surveys
- 2016 and 2018 Virginia Railways Express Surveys
- 2016 Maryland Area Regional Commuter (MARC) Rail Survey
- 2018 Ride On Title VI Survey
- 2018 DC Circulator Customer Satisfaction Survey
- OmniRide Westerly Subsidy Study
- 2017 Potomac and Rappahannock Transportation Commission (PRTC) On-Board Survey
- 2016 TransIT customer satisfaction survey
- Carrol Transit customer satisfaction survey
- Calvert County Public Transportation survey

# Summary of Survey Data

- Many surveys collected mode of access and mode of egress, fare paid, and limited socio-demographic variables
  - Minimum information required for Title VI reporting
- Most surveys did not collect origin and destination location
  - MARC collected most of the data required for modeling
  - Metrorail collected origin/destination station
  - VRE collected home address and destination address
- Route sequence unknown
- Most surveys used simple trip purposes
- Smaller operators tend to have less useful data for modeling

# Conclusions

- May be able to glean some useful ridership characteristics from surveys and compare to transit trips from household survey
  - Data could be used to create or adjust calibration targets
- Not possible to assign trips to network or attach skims to trips except in limited cases
  - Rail trips – boarding/alighting station
  - Trip segment between home and boarding/alighting
- Datasets may require re-expansion
- Suggest MWCOG initiate regional coordination on transit on-board survey design and implementation




RSG
the science of insight

**Joel Freedman**
SENIOR DIRECTOR
Joel.Freedman@rsginc.com