Update: Synthetic Household Travel Data from Passive Data Sources

A Proof-of-Concept in Atlanta

ARC Model Users Group
27 February 2015

Josie Kressner, Ph.D.
josie@transportfoundry.com

TRANSPORT FOUNDRY

3423 Piedmont Rd NE • Atlanta, GA 30305

Outline

About Transport Foundry

About NSF SBIR Program

Approach

Proof-of-Concept

Current Steps

Next Steps

About Transport Foundry

How do we take advantage of all the passive data that exists “at the click of a button”?

Submitted an idea; other people liked it

Won a National Science Foundation Small Business Innovation Research (NSF SBIR) grant

Transport Foundry was born

About NSF SBIR Program

Grant to do a proof-of-concept study building synthetic household travel data using an array of passive data sources

July 2014 - March 2015

Picked City of Atlanta as the test case

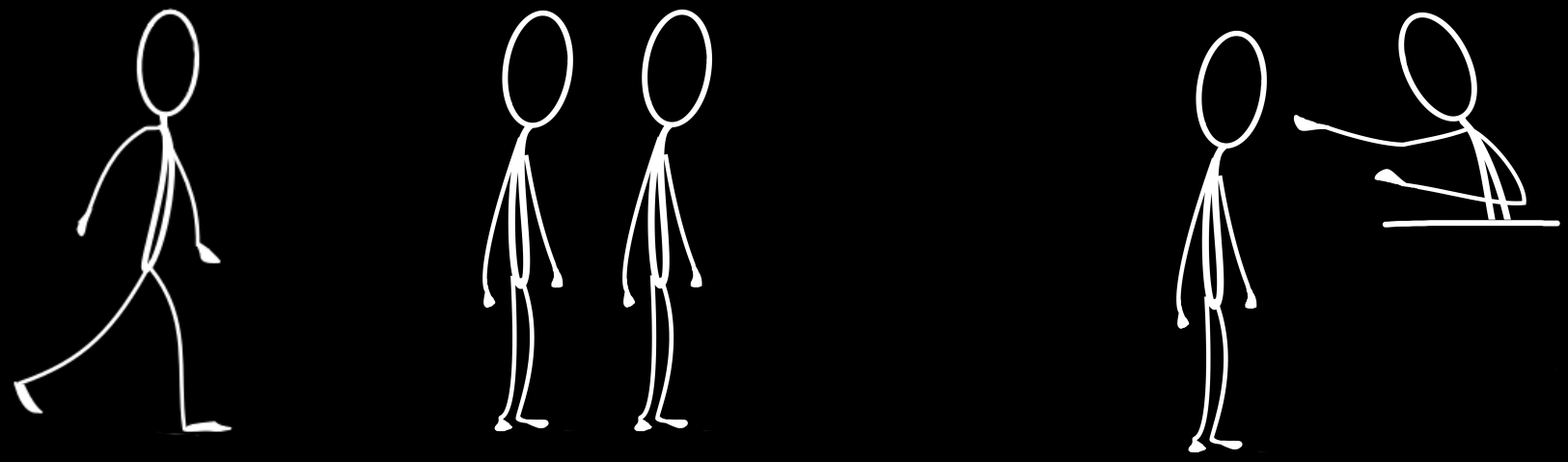
Approach

Generate synthetic travel diaries
using a process that marries concepts from
discrete event simulation
with **population synthesis** and **data fusion**

Use passive data from companies like Google, Inrix, HERE, AirSage, Experian, etc.

Take advantage of what each company is good at

Discrete Event Simulation Example

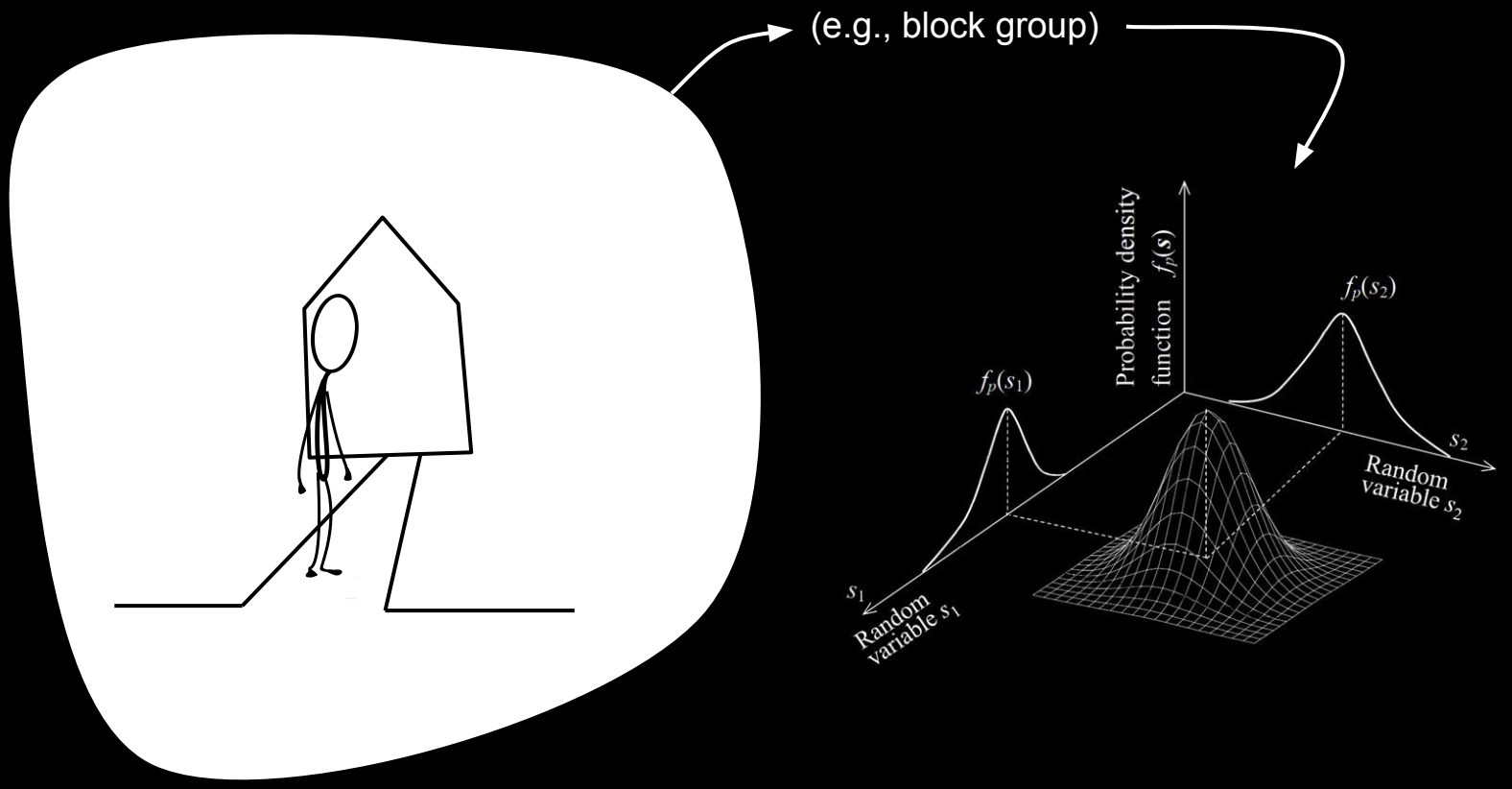


Proof-of-Concept

Simplify to start

Just home-to-work and work-to-home trips for one adult per household

The Process


(e.g., block group)
Probability density function $f_p(s)$
$f_p(s_1)$
$f_p(s_2)$
s_1
Random variable s_1
s_2
Random variable s_2

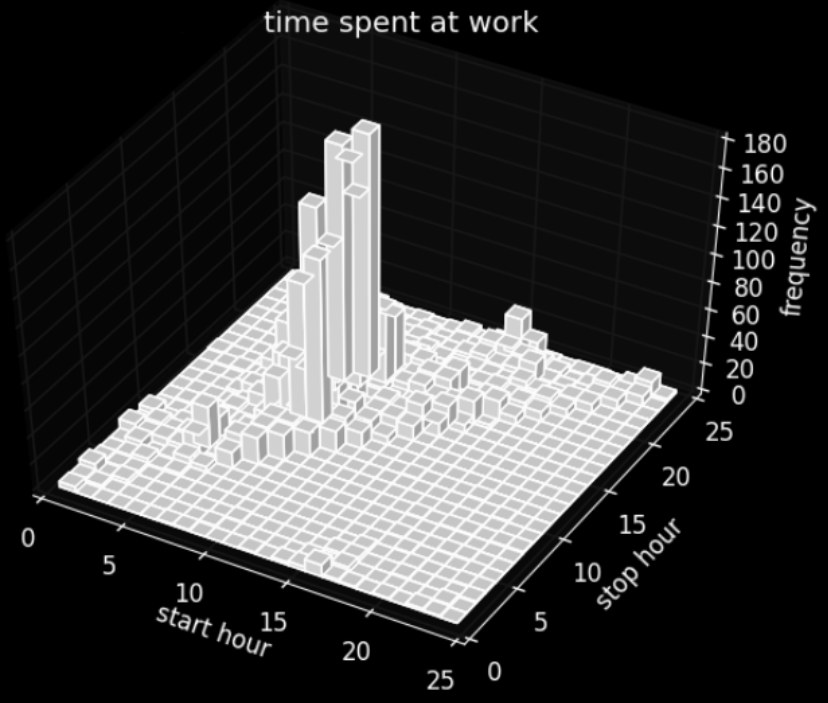
Passive Data Inputs

1. Household-level data (50% sample)
2. Origin-destination (OD) matrix for home-work travel
3. Travel time statistics in real traffic conditions between OD pairs
4. Activity durations for home and work

Passive Data Inputs

1. Household-level data (50% sample)
2. Origin-destination (OD) matrix for home-work travel
3. Travel time statistics in real traffic conditions between OD pairs
4. ~~Activity durations for home and work~~
   American Time Use Survey (temporarily)

Example Distribution



Output

Trip Table

HH_ID	Person_ID	Trip_ID	Trip_Strt	Trip_Dur	O_ID	O_Type	D_ID	D_Type
012948	01	01	07:53	00:15	131210095002	Home	132231206052	Work
012948	01	02	19:32	00:21	132231206052	Work	131210095002	Home
093410	01	01	08:40	00:24	131210035001	Home	130670301014	Work
093410	01	02	18:31	00:18	130670301014	Work	131210035001	Home

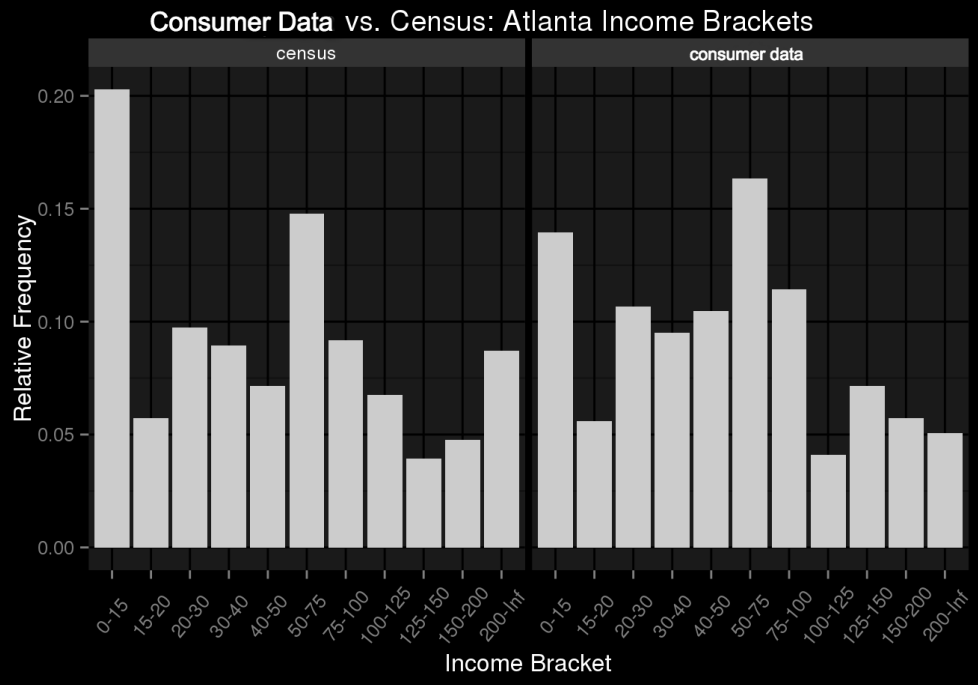
Household Table

HH_ID	Income	Num_Veh	...
012948	$30,000 - $39,999	2	
093410	$50,000 - $74,999	1	

Person Table

HH_ID	Person_ID	Gender	Age	...
012948	01	Female	42	
093410	01	Male	28	

Gut-Check Validation: Example 1



ACS 5-Year Estimates have such high margins of error at the block group level that QA with it is questionable

Gut-Check Validation: Example 2

Percent of adults who make a trip to work on a given weekday

American Time Use Survey = 33%
(output trip table does match input)

National Household Travel Survey = 44.3%

Gallup = 45%

Current Steps

1. QA using margins of error in ACS 5-Year Estimates

2. Tour characterizations from NHTS to output tour tables

Tour Table

HH_ID	Per_ID	Tour_ID	Trip_ID	Trip_Strt	Trip_Dur	O_ID	O_Type	D_ID	D_Type
012948	01	01	01	07:53	00:11	131210095002	Home	132231206052	Other
012948	01	01	02	08:03	00:04	132231206052	Other	132231206052	Work

Next Steps

1. More validation
2. Test new, detailed passive data sources again NHTS tour characterizations
3. Implement in Seattle with larger area to see if it transfers well

Questions and Comments

Thanks!

TRANSPORT
FOUNDRY

3423 Piedmont Rd NE • Atlanta, GA 30305

josie@transportfoundry.com