Summary of Q&A Pod:

Q: Can you say something about the Univ of MN Metropolitan Travel Survey Archive and ask people to send copies of their data to David Levinson
A: http://www.surveyarchive.org/

Q: Problems with missing or incomplete GPS data?
A: Generally, it's pretty good.   The units we used went to stand-by when not used, so generally the cold-start problem was only encountered on the first day.   We also made several assumptions for cold starts, such as if it were the first trip of the day, then it started from the home.   If it were lost in the middle of a trip, then the origin was assumed to be the destination from the previous trip, assuming that the times were not excessive.   Other than that, if a person noted that their unit was forgotten part of the day, then that day was not countable toward the "complete" definition.

One thing that we did notice was that the household would note that their units were not forgotten, but on trips where the entire family went, there were times when only 1 household member brought their unit.

C: The "keychain" sized GPS units that are becoming available eliminate the "tech savvy" consideration.   Just attach this "thing" to your keys and send it back in a few days.

Q: How does quality of data collected via web compared to CATI retrieved data? Seems inability to prompt would have an effect...
A: We have not evaluated this.

C: User smartphone use is a challenge as most will drain the battery fast, leading the user to turn it off.

Q: As much as new technology may increase participation from younger people, are you concerned about a loss from less tech savvy people?
A: Although we controlled for Retirees, we still over-recruited Retired Households by 3.7%.   However, the response to the computer only based Prompted Recall survey, Retiree Households were under-represented by 6.4%.

Q: Did anybody consider using the 2009 NHTS in the design of their survey?
A: No.   We started with the 1999 Columbus and 2001 Ohio Statewide datasets for compatibility.

C: Link to Met Council video: http://www.metrocouncil.org/planning/transportation/TBI/index.htm

Q: Any reason why the survey is not being conducted in any of the South Asian languages like Urdu or Bengali or Punjabi?   I would think there is a sufficiently large number of non-English speaking South Asians in the NYMTC region to justify this.
A: There is limited information to delineate the linguistically Isolated Household in specific language level in current Census data. Under Census 2000 Summary File 3 (SF 3), "P20_Household Language by Linguistic Isolation" only provides aggregate level of languages such as Asian and Pacific Island, and Other Indo-European languages. Only "PCT10_Age by Language Spoken at Home for the Population 5+ Years" includes detail level of HH language. Based on the analysis of both dataset, we decided top four languages in 28 our survey counties, which are likely to be languages for linguistically isolated Households.

Q: Thank you all, we had the problem of over-representation of 60+ age group 30% in the sample against 22% in the census.   I assume that this still works, not?
A: Typically, an over-representation of the elderly results in a suppressed trip rate.   The weighting should correct for that, but you might want to benchmark your weighted results by age against other sources such as similar regional surveys and/or the NHTS data.

Q: Can you clarify the "matched" (raffle) vs "unmatched" ($/hh) definitions for incentives?
A: Match the addresses to landline telephone numbers for necessary proactive contacting of prospective respondents.
- Matched Sample: sample that could be matched to a landline telephone number
- Unmatched Sample: sample that could not be matched to a landline telephone number. This sample that cannot be matched to a landline telephone number represents non-telephone and cell-only households.
  o Will be offered a monetary incentive of $50 to participate – paid upon verification that data from the household was complete and accurate.
- GPS households will also be offered a monetary incentive of $25 (payable upon completion verification and return of the GPS equipment)
  o Incentive for matched GPS household: $25
  o Incentive for unmatched GPS household: $50+$25 = $75

Q: what were the major advantages of these special GPS units compared to the commercially available ones?
A: We leased the units.   It seems that each of the consultants has a certain unit that they are familiar with and prefer.   (i.e. their data processing software is optimized for a certain unit.)   So, we preferred to use a cheaper unit, but as we were only leasing them, we deferred to their judgment.

Q: Do you think you got a much better sample of hard to reach households by the address based vs. phone based sampling frame?
A: Yes, our current matched and non-matched samples are based on availability of landline telephone or not. Hard to reach household does not only mean a minority and an environmental justice group, but include a young professional group. They are likely to have only one telephone line because of their budget restrain or preference.   In order words, many households no longer keep their landline phone, but use cell-phone only. Address sample design is less likely to miss contacting those groups.   FYI, cost for the cell-phone based HH sample is higher than one for landline-phone based HH sample.

Q: What type of GPS device? "Keychain" sized?
A: It was about 2.5" by 1.5".   So about pager sized.   It was developed exclusively by PlanTrans, and is not available commercially.

Q: How much can you balance differences between census and sample through weighting?
A: Ideally all weights across stratifications would be equal.   For small demographic segments, we are concerned when weights reach into the thousands.   It's not so much a concern when the demographic segment is large.

Q: Why are cell-only households hard to get?
A: several reasons, especially because people, when they move, often keep their old cell.   So if you are using local telephone area codes as the pool from which you obtain your phone numbers to call, you won't find those people.

Q: Becky mentioned a learning curve with incentives. Can she expand on what was learned?
A: We learned after a little time that we were offering incentives too early for some groups and too late for other groups. Once data collection begins, close monitoring is necessary to monitor the demographic retrieval bins. So, we added steps to the quality control plan to monitor progress in this area closer. We focused incentives on demographic groups we know are hard to get everywhere - such as large households, single-worker households.   I can provide more information on this particular aspect of the survey if you wish.

Q: 1% sample size
A: The best case would be to collect at least 2% of the population, but we knew our budget would not cover that. Thus, we spent extra time and effort focused on obtaining a truly random and representative sample. The use of incentive payments was a new part of this effort and helped a lot.

Q: average cost per sample
A: We estimate the overall cost per household will be $161. However, we knew the costs would be higher at first and decline as we perfected the data collection process. We are about half-way through data collection, so I am not sure how close we are to our original estimate yet.

Q: In your opinion, how reliable is the % sample size? What would be the margin of error with this sample size?
A: Please refer to answer to your next question.

Q: Did you use consultants or in-house staff for your data collection?
A: consultants.

Q: Why would weighting the data over-represent minorities?
A: one household's unique response can be repeated thousands of times...thereby biasing their individual choices

Q: What are the first three criteria that should be used to validate a travel demand model?
A: You might want to check the Model Validation and Reasonableness Checking Manual (2nd edition)
http://tmip.fhwa.dot.gov/resources/clearinghouse/1397