

Advantages and Disadvantages: Longitudinal vs. Repeated Cross-Section Surveys

A Discussion Paper

Julie L. Yee and Debbie Niemeier

Doctoral Candidate, Division of Statistics, University of California, Davis
Asst. Professor, Dept. Of Civil and Environmental Engineering, University of California, Davis, 95616

Introduction

The collection of panel data entails following a cohort of individuals with the purpose of monitoring changes over a period of time. For the Puget Sound Transportation Panel (PSTP) data, households were solicited to provide a two-day travel diary for each survey time period. The survey time period is referred to as a "wave," and a total of five waves of data have now been collected at approximately one year intervals. Ideally, in any longitudinal study, the same group of subjects (individuals) is followed during each wave, thus making possible the observation of any one individual's travel behavior over time. Longitudinal surveys differ greatly from the collection of repeated cross-sectional data in which an independent sample is collected at each wave to represent the population for that time period. With cross-sectional data, the observed trip information is representative of the population at a single period in time and the temporal aspects of a specific individual's travel is not necessarily available. In the following discussion we have focused on several pertinent issues regarding the scope and limits of statistical inferences for the two types of data that result from longitudinal and cross-section surveys.

Statistical Power

The aim in comparison studies is not only to illustrate the differences between populations, but also to establish some measure of significance on the observed difference. The population of interest in the PSTP is the population of residents residing in the Puget Sound area with statistical questions regarding the dynamics of their individual travel behavior at different time points during a five year interval. Either of the above methods, longitudinal or cross-section surveys, may be used to gather data in order to make computational comparisons regarding the travel behavior differences among representative samples of the Puget Sound population.

The statistics typically used to compare travel differences are means and proportions and are reflections of theme differences among the *entire* population, Since the result of any sampling procedure is subject to variation, then the amount of stock that can be placed on any estimate, (e.g., a mean) is controlled by the estimate's standard error. In repeated cross-sectional analyses, standard errors are large whenever large

variations between individuals (not necessarily the same person in each wave) exist, and the power to detect statistically significant differences in the estimates can be undermined. Alternatively, in longitudinal analyses, by identifying those observations that are measured on the same individuals, it is possible to focus on changes occurring within subjects and to make population inferences that are not as sensitive to between-subject variation. A simple illustration helps to elaborate on this point:

Example

1. Suppose three trip lengths in wave 1 and three trip lengths in wave 2 are observed from three subjects in each wave. This is a cross-sectional study and the three individuals may or may not overlap in the two waves. There is no matching of observations.

	Observed Trip Length			
Wave 1:	1	2	3	mean = 2, s.e. = 0.577
Wave 2:	4	3	2	mean = 3, s.e. = 0.577

$$\text{Mean increase} = 1, \text{ s.e.} = \sqrt{0.577^2 + 0.577^2} = 0.816$$

$$\Rightarrow t = 1/0.816 = 1.23$$

\Rightarrow statistically insignificant

There is no significant increase in the mean trip length over the two waves.

2. Alternatively, suppose it was known that the observations were obtained from the same three individuals as shown:

<u>Subject:</u>	<u>1</u>	<u>2</u>	<u>3</u>
Wave 1:	1	2	3
Wave 2:	2	3	4
Increase:	1	1	1

$$\text{Mean increase} = 1, \text{ s.e.} = 0$$

$$\Rightarrow t = +\infty$$

\Rightarrow statistically significant

In the first example, it is not clear that a widespread increase in average trip lengths is occurring based on inspection of the data. This is supported by an insignificant t-statistic. In the second example, where each observed subject's trip length increased uniformly by 1, there is greater reason to believe that this is occurring on a widespread basis; this is supported by a very large t-statistic.

The estimated mean increase is the same for both methods, but the standard error plays a key role in determining statistical significance. The Generalized Estimating Equations (GEE) analysis used by Yee and Niemeier [1] is based on this concept. The analysis adjusts the Generalized Linear Model (GLM), which fits a model of completely

independent observations, to account for correlated observations (e.g., trips made by the same individual). The GEE approach accounts for this individual correlation by computing first, naive standard errors assuming complete independence (the standard GLM) and then computing an adjustment to the naive standard errors. This adjustment results in robust standard errors that are asymptotically equivalent to standard errors based on matching observations. In studies comparing trends with time, longitudinal (panel) data have an advantage over repeated cross-sectional data because it facilitates use of methodologies, such as GEE, that separate the nuisance variation due to population-wide behavior from the variation related to trends with time. The naive variance still has the variation due to population-wide behavior incorporated into it.

In the Yee and Niemeier [1] analysis of the PSTP, using the first four waves of data, it was found that people were making more SOV trips and fewer HOV-transit trips to work. Had similar data been collected in a repeated cross-sectional design, the naive standard errors would have been applied to compute statistical significance and fewer significant results would probably have been established. This effect can be simulated by comparing a GLM analysis performed on the same PSTP data to the GEE results reported in Yee and Niemeier. As Table 1 illustrates, the GLM analysis would have detected that the frequency of HOV worktrips had increased over time, but would not have detected that HOV-transit worktrip frequency had declined.

Table 1. Comparison between GLM and GEE results (using Yee and Niemeier)

<i>Difference between 1st and 4th wave frequency for</i>	<i>Result using GEE (matching of individuals)</i>	<i>Result using GLM (no matching)</i>
<i>complete data : SOV</i>	significant increase	significant increase
<i>complete data : HOV-transit</i>	significant decrease	not significant
<i>income \$0-35K : SOV</i>	significant increase	significant increase
<i>income \$35+K : SOV</i>	significant increase	not significant
<i>lifecycle 1: HOV-pool</i>	significant increase	not significant
<i>lifecycle 2: SOV</i>	significant increase	not significant
<i>lifecycle 3: non-motor</i>	significant decrease	not significant
<i>lifecycle 6: SOV</i>	significant increase	significant increase
<i>lifecycle 7: HOV-transit</i>	significant decrease	significant decrease

Both the GEE and the GLM analysis used to obtain the results shown in Table 1 were performed on the restricted data; defined as those subjects with complete worktrip data using the four waves of available data. This is only a subsample (519 subjects) of the larger data set (1527 subjects) in which subjects completed travel diaries in all of the first four waves; in the larger data set, trips were made in every wave but they may or may not have been worktrips. Since the GEE method does not require all data come in matched sets then either of the GEE or GLM methods can be used on the full data set; the GEE analysis would still be more likely to detect significant changes whenever they did exist.

Coverage

The drawbacks of using longitudinal data have largely to do with coverage problems. Coverage refers to “the set of units constituting the target population” [2] and includes issues associated with both selecting and tracking individual sample respondents. The prominent coverage limitations are:

1. After the first wave’s recruitment, the study is restricted to the members of that sample although changes in the population may occur.
2. Despite attempts to locate households from wave to wave, there is invariably a fair amount of attrition.

The design of longitudinal data is particularly well suited for stationary populations. In regionwide transportation studies, this limits the inference to subjects residing long-term in a closed region. However, most regionwide populations, like Puget Sound, are not closed systems. By keeping the sample population fixed, there is a risk of making inaccurate conclusions about the true population which may have changed as a result of influx or outflux of residents with different behavioral characteristics than the indigenous population.

This raises the question of what conclusions can be made with longitudinal data and this question is, in turn, related to understanding the primary goals of the PSTP study. For example, is the primary purpose of the PSTP to monitor changes in travel activity in the Puget Sound area as an aggregate of effects ranging from demographic changes and group dynamics.³ Or is it to monitor changes in travel activity due to individual attributes? The former goal, measuring the aggregate effects by demographic group, can be addressed with independent repeated cross-sectional sampling; each wave’s sample of subjects is allowed to vary with time and with any changing population dynamics. The latter question is probably better addressed with longitudinal data since this data collection method provides more detailed insight on behavior at the person level.

Solon notes that marginal probabilities may be estimated using either longitudinal or repeated cross-sectional data, but certain conditional probabilities may be estimated only with longitudinal data [3]. For example, it is possible to use either longitudinal or repeated cross-sectional data to study whether there was an increase in SOV commute activity between wave 2 and wave 1; but to assess whether a person who uses HOV-transit in wave 1 is more likely to use SOV or HOV in wave 2 requires the use of longitudinal data. Similarly, PSTP reveals that there were greater incidences of HOV-transit commuters transitioning to HOV-pool commuting than to SOV commuting [4]. Likewise, SOV commuters were more likely to become HOV-pool commuters than HOV-transit commuters in later waves.

The optimal survey design would combine the positive attributes of both sampling methods. For example, longitudinal sampling with rotation allows the entry of new subjects which helps to capture any dynamic changes in population composition due to immigration while still retaining a portion of the earlier sample groups to represent any person level changes among the stable population.

Each method, either the standard longitudinal or rotating method of survey, still face possible attrition problems. Attrition embodies the combined sampling problems due to out-migration and non-migratory dropouts of survey respondents and is a potentially serious source for bias. The analysis of incomplete data has been widely studied and methods are continually being developed to handle data when some observations are missing at random. However attrition is known to occur more with some groups of people than with others. In particular, subjects with lower income, lower ages, and lower trip frequencies have a greater tendency to drop out, and this is likely to create an offset in statistical results [5].

Independent repeated cross-sectional sampling can be advantageous in this regard since complete sets of new respondents are continually selected, thus ensuring a steady level of reliability for each successive sample when under stable sampling conditions. Alternatively, the practice of “refreshing” the sample of a longitudinal study with new recruits who fit the profiles of those who drop out has much of the appeal of weighted and stratified repeated cross-sectional sampling without compromising the advantages of repeatedly sampling from participants who do not drop out. This is discussed in more detail in the next section.

Sampling Weights

The role of weights serves to counter the biases that may result from disproportionately sized sampling strata. The PSTP was collected by a stratified sampling protocol based on mode use proportions derived from previous research. Respondents were recruited using three methods: random telephone digit dialing, contacting prior participants in the Seattle Metro transit surveys, and solicitation of volunteers on randomly selected bus routes. The random telephone digit dialing method was the primary way of collecting participants who drive alone or carpool. Since transit users comprise a very small portion of the population, the latter two methods were used to supplement the random telephone digit dialing in soliciting transit users into the survey, and sampling weights were developed to help control for the disproportionate sizes of the mode groups.

Using weights, however, binds sample mode proportions to prescribed values. Under an independent repeated cross-sectional sampling strategy, this weighted and stratified sampling protocol could fail to reveal changes in mode choice behavior over time. This is less of an impediment in longitudinal studies where it is the sample that remains fixed and mode choices are observed to vary. Unfortunately, attrition in

longitudinal studies remains a problem and seems to be best treated by the replacement of lost subjects. This effectively leads to a partially repeated stratified cross-sectional sampling of new subjects. Replacements recruited to counter the effects of attrition in the PSTP were selected so that their profile were similar to those who dropped off from the study. Household travel mode was considered one of the main criterion by which the refreshment sample was selected. The integrity of the study on mode choice trends is potentially weakened if the lost households had also changed their travel mode behavior at the time they left the study. Nevertheless, this deficiency of repeated weighted resampling casts the longitudinal analysis with modest attrition in a more favorable light than the repeated cross-sectional analysis.

Conclusion

The benefits of a longitudinal analysis over a repeated cross-sectional study include increased statistical power and the capability to estimate a greater range of conditional probabilities. With the PSTP, and any study where weighted stratified sampling is employed, the benefits extend to include the capability to make appropriate inferences regarding changes in strata proportions with time. In the PSTP study, this would include changes in mode choice proportions.

The repeated cross-sectional analysis is sometimes more desirable. This can occur when:

1. It is more cost-effective to resample many new subjects than to repeatedly sample from the same group of subjects; moreover, enough new subjects are recruited to compensate for the loss in statistical power when using a repeated cross-sectional analysis;
2. only marginal probabilities are of interest; and
3. either unweighted and unstratified sampling is used and justified, or weights are used to control for some attributes but estimates of the values or proportions of these attributes are not made.

References

1. Yee, J.L. & Niemeier, D.A., "Travel Trends Using the Puget Sound Panel Survey: A Generalized Estimating Equations Approach"
2. Colledge, M. "Coverage and Classification Maintenance Issues in Economic **Surveys,**" **Panel** Surveys, Wiley 1989
3. Solon, G. "The Value of Panel Data in Economic Research," **Panel** Surveys, Wiley 1989
4. Yee, J.L. & Niemeier, D.A., "Task E - Volume 2: Specialized Travel Trends in the Puget Sound Panel Data 1989-1993"
5. Murakami, E. and Watterson, W.T. "Attrition and Replacement Issues in the Puget Sound Transportation Panel," 70th Annual Meeting of the Transportation Research Board 1991