Seasonality of Transportation Data

Emily Parkany
Assistant Professor
Civil and Environmental Engineering
Villanova University
800 Lancaster Avenue
Villanova, PA 19085
Tel: (610)-519-4957
Fax: (610)-519-6754
Email: emily.parkany@villanova.edu

Report Submitted to
Elaine Murakami
September 2004

Seasonality of Transportation Data

Emily Parkany

BRIEF SUMMARY

The ability to collect transportation data over an extended period (i.e. more than a year) is quite appealing. With a long collection period, we expect to obtain seasonal information such as travel times and work departure times that may vary from winter to summer. Perhaps we can observe trends in the data if the collection period is long enough. But there are disadvantages to continuous measurement, many related to sampling and measurement errors. The results shown here are of special note for the transportation planning community and policymakers who rely on decennial census journey to work data. The decennial census long form will likely be replaced by the American Community Survey (ACS), a continuous measurement survey. The brief literature of seasonal data is reviewed. We discuss the advantages and disadvantages of continuous measurement. Most of the data provided here was obtained in the Research Census Data Center with three years of ACS data for one test site. The visual data suggests that seasonality may not be readily apparent in either monthly or quarterly aggregations. However, close examination may reveal that the quarter that includes April may have lower travel times and work departure times may be earlier than other quarters. This has ramifications for planners accustomed to using decennial census data collected in April. Additionally, the randomness in the graphics suggests unwelcome disparities in continuously measured data.

SEASONALITY OF TRANSPORTATION DATA

The decennial census long form will likely be replaced by the American Community Survey (ACS), a continuous measurement survey. This may provide an opportunity to obtain transportation data collected throughout the year which may reflect seasonal differences. Here we study whether the data reveals variations that can be attributed to seasonality. The results shown here are of special note for the transportation planning community and policymakers who rely on decennial census journey to work data such as the Census Transportation Planning Package (CTPP). After a brief review of the seasonal transportation literature, the advantages and disadvantages of continuous measurement are discussed. Results pertaining to three years of ACS data collected at one test site are presented.

PROJECT OBJECTIVE

The seasonality of transportation variables from three years of microscopic data from one of the 31 American Community Survey test sites, a county in western Massachusetts, is studied. The microscopic data is only available in high security conditions at a Census Bureau Research Data Center (RDC). Data can be manipulated at the RDC but results must go through the Census Bureau Disclosure Review process before dissemination. The journey to work travel time and morning departure time are studied as well as the relationship of these variables with income and mode choice. The data varies greatly from the overall mean each month and from one month to the next month, so statistical analysis is discounted and visualizations of the data are emphasized.

ACS could provide an opportunity, heretofore unavailable, that allows the examination of seasonal trends. If there is any indication of seasonality, then "CTPP-like" products would include new tables that might be useful for modeling (travel demand models, air quality). A related question is whether 60 months of accumulated ACS records be adjusted to "look" more like April 1 data for a particular year (to be analogous to decennial census data). Here, we look at both monthly and quarterly visualizations of the data. It is expected that quarterly aggregates of "CTPP-like" tables will be easier to release than monthly aggregates for most census areas. In addition to determining whether seasonal differences can be observed in the data, we consider whether "calendar quarters" (January, February, March as one quarter; April, May, June as the next; etc.) better depict seasonality compared to "seasonal quarters" (December, January, February as one quarter; March, April, May as the next).

BACKGROUND/TRANSPORTATION DATA SEASONALITY

One would expect that other researchers would have considered seasonality of transportation data used for planning and other purposes. Seasonality is addressed in the macroeconomics literature in terms of employment and other types of time series data. Transportation data used as economic indicators (inland waterway transported tonnage or railroad services, for example) may be adjusted for seasonality (Bureau of Transportation Statistics (BTS), 2004; BTS 2002). But this freight data is not generally used for transportation planning purposes.

One exception is traffic data (mostly volumes). The FHWA monitors traffic volume counts at several thousand stations along different types of roadways in the U.S. This is one of the few sources of monthly transportation data. With the current emphasis on archiving data for intelligent transportation system purposes, one can expect more areas to have traffic volume and speed data archived. Monthly or quarterly aggregations of this data are not readily available in all states and locations, but some states seem to have better accessibility and processing of the data than other states. For example, the Oregon Department of Transportation explicitly describes how to use a counter location to make a seasonal adjustment factor appropriate to use on similar roadways (Oregon DOT, 2001).

Figure 1 shows monthly average daily traffic for one federal monitoring station in Hampden County, MA (the county of the ACS test site data of this analysis). There is some variability from month to month, but the adjacent points seem correlated, and there is a distinct seasonal pattern: volumes for the first half of the year are low (with a large "drop" in January) and volumes rise until they "peak" in August, September and October. One can hypothesize that January lows may be because of winter holiday and vacation periods. Fall peaks may be the result of the start of the school year and New England's foliage season.

The most closely related literature to the project presented here is a recently released report, "Seasonality in Daily Travel Patterns" (Seethaler and Richardson, December 2003). This seems to be the only literature available on this subject except that written by members of this team. This report considers six years of the Victoria University Activity and Travel Survey conducted in Melbourne, Australia. The authors found that monthly segregation of daily distance, daily travel time, daily trip stages for non-motorized, motorized, and public transportation modes ***do not*** show any seasonal patterns that repeat themselves over the six years studied: 1994-1999. They suggest that "each year has its own temporal fluctuation pattern around the annual mean values." They segregate by week and do not find a seasonal pattern. They did not seem to segregate by quarter. They conclude that there is no need to adjust for seasonality when conducting "before" and "after" studies related to policy implementation.

A related paper conducted by Parkany and Madron at Villanova University that was motivated by this project included comparisons of ACS data with contemporary travel data sets. This paper is described in this report and is presented as an appendix.

CONTINUOUS MEASUREMENT

Statistician Leslie Kish advocated the use of rolling samples in cumulating information over space and time in 1958 (Alexander, 2001). He wrote on samples and censuses in 1979 and first started advocating their use for the U.S. Census in 1981 (Alexander, 2001). Continuous sampling has at least two "features": (1) better data quality through maintenance of a permanent enumeration staff and improvement through constant experience; and (2) increased frequency of available data at multiple points through the decade (The National Academies Press, 1993;

Alexander, 1999; Alexander, 1993). Other census advocates explain the benefits as helping to maintain the master address file, potential timeliness of data release, and potential cost savings (Scarr, 1994).

The transportation community has expressed some concern over the implications of continuous measurement in transportation planning (US Department of Transportation, Bureau of Transportation Statistics, 1996). In a recent report, researchers and metropolitan planning organization officials expressed trepidation about costs to local agencies, timeliness of data from the Census Bureau (CB), and similarity to currently-used long-form (Public Use Microdata Samples (PUMS) and CTPP) data. Now that continuous measurement will replace the long form, there is a need for determining whether seasonality effects/files need to be included in CTPP and other census data projects used for transportation planning.

Advantages and disadvantages for the transportation community of using continuous data are presented in Table 1. The advantages include data available more frequently than once every ten years and the likelihood that no data used for policy-making in any given area is older than seven years (five years for release and up to two years for processing). One can also suggest that larger areas (possibly with larger data and planning needs) will receive acceptable data every two to three years (still possibly with a two-year lag). Hopefully, with continuous measurement, one can capture existing seasonal differences in data.

There are many disadvantages to using continuous data. The greatest disadvantage relates to the Census Review Board's disclosure review policies. The Census Bureau (CB) must protect individual confidentiality. One hypothesis is that if a table shows transportation mode for a specific tract-to-tract pair, a specific individual could be identified. In general, the census policy is to release *statewide* data with 75 observations per tabulated "cell" (where a cell is one box on a cross-tabulation). For the CTPP, long negotiations have ensued wherein the CB agrees to release observations of three unweighted records for some of the flow tabulations. [For the data released for this study, cells required two or more unweighted observations.] CB release negotiations have delayed the current release of CTPP 2000. All of these difficulties have occurred with a 17% sample. With full implementation of ACS, the CB anticipates a 12.5% sample at best and only a 2.5% sample annually.

In addition to disclosure-related concerns, there are other disadvantages to continuous measurement. Large sampling errors may occur when dealing with small area estimates and aggregations over many years (Alexander, 2002; Chand and Alexander, 1996). Naysayers suggest that these errors may propagate rather than counteract. The same metropolitan statistical area may have only five-year aggregates for some areas and shorter aggregated data for larger areas. One planning agency may use data from 2004-2008 and another may use 2006-2010 data leading to different baselines and difficult comparisons. The advantages to a permanent data collection staff may be outweighed by the disadvantages in having to secure annual funding from Congress.

AMERICAN COMMUNITY SURVEY

The American Community Survey is intended to replace the decennial census "long form." The long form has traditionally provided additional demographic and socioeconomic data for one out of every six households. The data has been collected with other census data in April of the census year. In contrast, the American Community Survey is intended to survey the same number of households over a five-year period. The decennial census long form sampled 17 million households; ACS will sample fewer than 3 million household units per year. This will potentially result in the annual release of data for areas over 65,000, three-year release for areas over 20,000, and five-year release for smaller areas such as tracts and block groups. Currently the permanent funding of the ACS is in question (U.S. Census Bureau, Demographic Surveys Division, Continuous Measurement Office, 2003). For now, data collection continues at the original test sites, but the full implementation expected in 2003 has not occurred.

Table 2 shows the transportation-related variables available in the census products (either the long form or the new American Community Survey). The products have questions related to the "Journey to Work" such as the location, travel mode, travel time, and departure time.

The American Association of State Highway and Transportation Officials (AASHTO) and the Census Bureau have put significant resources toward creating the Census Transportation Planning Package (CTPP) (Census Transportation Planning Package, 2000) that is used by Metropolitan Planning Organizations (MPOs), state Departments of Transportation, and other planning agencies. There is concern that there will be seasonal variability of the transportation variables including "journey time to work" and "transportation mode selected for commuting" in the continuously measured variables from the American Community Survey compared to the currently-used census long form variables collected in April of the decennial year. Additionally, there are concerns about the compounding of sampling and small-estimate errors when aggregating data over extended enough periods to provide desired origin and destination tables by mode and race such as provided in the current CTPP. This report concentrates on the seasonality concern.

MPOs and other organizations that use CTPP data make policy decisions based on the census data collected in April every ten years. More frequent data release will change long-standing policies and uses of the data. Additionally, we expect that seasonal differences may affect travel times and work departure times so these differences are studied here. We also expect mode choice differences in summer and fall given seasonal transit schedules and that more transportation alternatives (such as walking and biking) are attractive during the warmer summer months. Some areas have changing populations (such as the "snowbirds" in Florida) and others have high tourist seasons. We hope that the seasonal differences can be captured, but normal variability in the data and small sample errors may lead to less understanding of "true" values compared to the values obtained with the full decennial census information.

AMERICAN COMMUNITY SURVEY TEST SITE IN HAMPDEN COUNTY, MASSACHUSETTS (ACS-HMA)

The main data presented here are ACS data collected by mail-out surveys, follow-up phone and follow-up in-person interviews during 1999, 2000, and 2001 (January through December) in the county including Springfield, Massachusetts. This county is part of a metropolitan area that is the fourth largest in New England and includes the region's second largest transit system. Parts of the transit system utilize reduced summer schedules. Additionally, New England winters likely result in different mode choices, travel times, and commute departure times. Access to the microdata, including survey response month, is available only in a Census Bureau Research Data Center (RDC) after many researcher hurdles. [A related FHWA transportation data project was not successful due to difficulties with the RDC system. That project report better documents the difficulties experienced by demographic data users at RDCs (Niemeier, 2004).] Unlike the full implementation of ACS where three percent of households are surveyed each year, five percent of county households in the test sites are surveyed each year. Thus, the three-year aggregates of the data displayed here should be analogous to the decennial census long form data. About 6,500 peak period work trips were collected from 12,000 individuals surveyed. The only direct transportation-related variables are the journey to work variables, as shown in Table 2.

METHODOLOGY

Research Data Center Logistics

Analysis of census data from individual respondents must be conducted as a Census Bureau Research Data Center (RDC). Eight Research Data Centers exist across the country for external researchers to work with confidential census microdata. From the first location of this research, University of Massachusetts (adjacent to Hampden County, Massachusetts), the closest RDC was in Boston about one hour, forty minutes away. The principal investigator switched universities and the closest RDC to her Philadelphia, Pennsylvania location is Upper Marlboro, Maryland (suburban Washington, D.C.) which is about two hours, fifteen minutes away. Using an RDC requires an application process with, among other things, a detailed explanation of how the research benefits the Census Bureau and an administrative payment for using the RDC. The proposal for this project was accepted in its first try but with modifications. Approval of the modifications took another few months. Before one arrives to use the data, the researcher needs to fill out several clearance forms, providing fingerprints and personal references. Additionally personnel at each university sign forms assuring that the data would not be mishandled. After this process, the RDC administrator swears you in to use the data and provides access to workstations. A timetable for transferring the data from Boston to Maryland and getting data ready for disclosure review is provided in Appendix A. It seems difficult to make the process quicker.

At the Research Data Center, researchers are provided with UNIX workstations, raw data, statistical packages like STATA, SAS, and SPSS, and a locker. The only word processing program available is the UNIX-based Star Office. Disclosure documents are prepared with this limited system and research notes were either handwritten or utilized UNIX-based unformatted note utilities. Printing is strictly monitored—approval must be gained before printing out data dictionaries and other documentation and these must remain at the RDC in the project's locker. Disclosure documents were removed from the researcher's directories and printed by the RDC administrator. There is no internet access and the use of another computer or storage device is forbidden. The RDCs are shared by other researchers, most of who use Census Bureau "economic" (business-related) data and tax data, not demographic data like the American Community Survey.

Given the remoteness of the Centers compared to the location of the principal investigator, analysis was conducted sporadically (sometimes every other week, sometimes less or more often) at the RDC for relatively long periods of time (five to ten hour work sessions). All data manipulations must be conducted at the RDC. Preparation of the disclosure documents including descriptions of the tables desired for release was conducted at the RDC. The initial steps in the disclosure review process are provided in the timeline in Appendix A. The final data was released in November 2003 after values based on two or fewer respondents were deleted. This data is presented in raw form in Appendix B.

Preliminary Analyses

Initial data analysis included determining the best variables to use to date the time each respondent filled out the survey, which journey to work variables were possible to compare seasonally, and determining the best limits to use to capture peak hour travellers. Two date values were provided for each respondent. These corresponded to the date that that respondent initially filled out the survey and the date that the Census Bureau received the survey. In approximately 10% of the cases, these cases were separated by a month, in some cases more. The transportation/journey to work questions are at the end of the American Community Survey form so it is not clear when respondents really answered the journey to work questions. The receipt date was used as the date for each survey and plans to distinguish weekly data and perhaps match up weeks with local weather conditions were scrapped.

Some preliminary analyses were performed with the journey to work mode choice frequencies and it was determined that the numbers varied greatly month to month. For this reason, the results in this report mainly consider journey to work travel times and monthly and quarterly aggregated travel times instead of depicting monthly mode choices.

Some investigation of seasonality by detailed geographic zone was completed but for the Hampden County data, the smallest geographic areas were the 14 "places" of more than 2500 people in the county. There are no census tracts, census block groups, combined zones, or traffic analysis zones in the census-related data collected in Massachusetts. Approximately 40% of respondents live or work in the largest place, the city of Springfield. Respondents that live and work in Springfield are analyzed in terms of travel time and morning departure time by mode later in this report.

During the preliminary analyses, ANOVA and Bonferroni analyses were conducted. ANOVA (Analysis of Variance) tests were conducted on similar data values to determine whether each value differs significantly from the group mean. In this case, monthly aggregates (of travel time, for example) were compared with the overall average annual travel time and it was found that most months varied significantly from the mean. This result suggests the value of displaying the differences visually as occurs in this report.

When considering quarterly aggregations of the data, there are fewer comparisons with the overall mean. The quarterly aggregates still varied significantly from the mean according to preliminary ANOVA testing. As a further test, the researcher used a post-hoc statistical test, a Bonferroni test, to determine how each quarter's value varied from the other quarters (not just in comparison to the sample mean). The Bonferroni results showed significant differences among most of the values for the quarters.

Thus, ANOVA and Bonferroni tests can statistically show that most of the monthly and quarterly means are significantly different from the annual means or the three-year aggregated mean. To relate how different the monthly and quarterly values are, graphical results are presented in this report. One can learn more from visualizing the differences and attempting to observe seasonal trends than by trying to observe trends in tables of numbers (The raw data released by the Census Bureau is presented in Appendix B.) or from the basic statistical result provided by the ANOVA and Bonferroni tests. Visually, one looks for similar annual patterns in the data to help determine if there seem to be seasonality effects.

Data limits and weights

All of the results presented here are from respondents who left for work between 5 a.m. and 10 a. m. and who travelled less than 120 minutes in their journey to work. All of the results use the person weight provided in the data set. This is a frequency weight. It is intended to be used with each observation to represent that person's actual proportion in the whole population. By using the weights in the data analysis, the results (monthly travel time aggregations, for example) should be representative of the entire county. Most of the results use data (from all respondents) representing the whole county. The mode-related results use only the respondents that live and work in the city of Springfield as described below.

Process

For the study to determine whether there are seasonal effects in the American Community Survey transportation variables, we reviewed several variables and combinations of variables. First, we considered employment in the area. Then we specifically evaluated the journey to work travel time variable monthly, in relationship to the mean, and both calendar and seasonal quarters. We consider journey to work departure time in the same way. Then we consider the seasonal effect of different income quartiles and the two variables: travel time and departure time. Finally, we consider quarterly results of the two variables by mode (drive alone, carpool, and public transit) for the respondents that live and work in Springfield. Further subdivisions of the variables, specifically travel time and departure time by mode choice and income of ACS respondents who live and work in Springfield, are presented in Appendix C as described below.

EMPLOYMENT

Before we look at the transportation data for Hampden County and Springfield, Massachusetts, we look at employment in the area and try to observe seasonal trends. We hypothesize that journey to work travel time may be reduced during periods of less employment.

Figure 2 shows Springfield, Massachusetts employment taken from the Bureau of Labor Statistics (BLS). Here we observe more jobs in 2001 compared to 2000 (until September 2001) and more jobs in 2000 compared to 1999, as expected. But a seasonal trend is also easily discernable. Each year, job numbers rise until a peak in June and then they drop for July and August. Employment rises almost as steeply from August to September as it drops from June to July. And, with the exception of 2001, jobs increase from September to December and the highest employment occurs in December. The average of the three years emphasizes this obvious trend of "dip" in July and August and the subsequent December jobs peak.

But when the respondents who are employed are taken from the American Community Survey for Hampden County, the employment trend is not obvious. Figure 3 shows monthly employment values (the weighted number of people who responded that they worked last week—including serving in the armed forces and working without pay for a family business) for 1999, 2000, 2001 and the weighted average of the three years. Here, employment for the three years cross and only 1999 exhibits a peak in December. None of the years show a dip in July and only year 2000 shows a dip in August, and in contrast to the BLS data, also in September. There does seem to be correlation in the ACS data in that adjacent months follow the same tendency (positive or negative employment gains).

TRAVEL TIME TO WORK

In this section we look at the Journey to Work travel time variable from the Hampden County ACS. The data points obtained in this figure and all of the figures in this report are weighted with the person-weights provided in the data set. The average line in each figure is weighted by the monthly number of respondents; it is not simply the average of the 1999, 2000, and 2001 values.

Monthly Aggregations
Figure 4a shows weighted monthly average values of the journey to work travel time variable for all respondents of the American Community Survey in Hampden County, Massachusetts from 1999 to 2001. The average value shown for each month is based on approximately 200 respondents from the county for the month. Each respondent's value for travel time is weighted with the person weights provided in the data set. The three year aggregate line is weighted by the number of respondents each year. This set of travel time to work figures (Figures 4a, 5, and 8) are shown with a 15 to 30 minute scale so that they can be compared with each other.

The monthly data shown in Figure 4a shows that the data varies from month to month. Two of the years (2000 and 2001) show the longest travel time in February followed closely by a high point in all three years in August. March and April reveal the lowest travel times except for an overall low in December 1999. The high journey to work travel time value in August contradicts our hypothesis that lower employment in August would result in lower travel times. The larger August value may be related to the people who need to travel to work in August having longer travel times and that transit schedules are reduced in the summer. That March and April travel times are consistently lower than other months is an interesting result given that transportation planners have typically used long form census data and the Census Transportation Planning Package to validate their travel models. The decennial census data is collected in April of the census year. Thus, these April values may be lower than the true (or annual) average travel time.

Figure 4b gives a pictorial representation of how each month varies from the weighted average travel time obtained by aggregating all three years of data into one value. Each month varies from the overall mean by 1 ½ minutes or less. Eight of the twelve months vary by less than a half-minute more or less than the overall mean of XX minutes. As stated above, February and August are about a minute higher than the overall mean. The March and April means are more than half a minute less than the overall mean.

Quarterly Aggregations

It has been difficult to get all of the desired data used for transportation planning purposes in the CTPP released using decennial census data collected in one month from 1/6 of all households. Even over three and five year periods, fewer households will be sampled with the American Community Survey. So this report may reveal that monthly or quarterly aggregations of transportation data may be desirable, but for small census areas, smaller aggregations may not be releasable. Instead of monthly values, it is more likely that quarterly aggregations may be released. Here we consider quarterly aggregations of the data in terms of both "calendar" quarters and "seasonal" quarters.

Figure 4 depicts monthly values for the journey to work travel time. Figure 5 shows quarterly values of the travel time to work. In Figure 5a, we see the twelve travel time minute points for the "calendar" quarters. The first calendar quarter is January, February, and March; the second point is for the quarter including April, May, and June; the third quarter represents July, August, and September, and the fourth quarter is October, November and December. As can be expected with any aggregation, fewer points on the graph (quarters) reveal much less variation than the monthly aggregations in the previous graphs. Compared to the variation observed in the monthly graphs in Figure 3, Figure 5a is fairly flat. In fact, the second and third calendar quarter points show little annual variation.

In contrast, the "seasonal" quarters displayed in Figure 5b fluctuate and there is more variation among the values from the three years. "Seasonal" quarters consider December, January, and February as the "winter" quarter represented by the point over "January" in Figure 5b. Similarly, "spring" is represented by March, April, and May; "summer" by June, July, and August; and "fall" by September, October, and November. Figure 5b reveals the shortest journey to work travel time variables in the "spring" (March, April, May). This confirms the lower travel time variables revealed by the monthly data. Thus, seasonal quarters do show a small effect of lower travel time values in spring compared to a "peak" in the summer and higher values in fall and winter. Figure 5c provides the three-year calendar and seasonal quarter trends in the same figure.

DEPARTURE TIME

Monthly Aggregations

Figures 6 and 7 show data for one of the other main journey to work transportation variables: departure time. Here the data from each month is aggregated, weighted with the provided person-weight variable, and one value for departure time is provided. One may hypothesize that people leave for work later in the winter to take better advantage of daylight. In the summer, people may leave earlier for work because there is more daylight and there are more activities to undertake during daylight hours after work. These figures (Figures 6a, 7, and 9) all use a 6:43 a.m. to 7:55 a.m. scale so that they can be compared with one another.

The monthly journey to work departure time data is shown in Figure 6a. The data varies from month to month and the values from year to year cross each other frequently. The three year aggregate line reveals earlier departure times in May, August, and November and later departure times in January, March, and September. This is confirmed in Figure 6b which shows monthly variation from the mean of the journey to work departure times and how six months stay within a minute of the overall average (leaving about 7:18 in the morning) and the average departure time in the other months varies by about two, three or four minutes in either direction. One may not have good hypotheses about the variation in the other months, but one can imagine that leaving four minutes later in September may be tied to school starting and families relearning the routine of having to get children ready.

Quarterly Aggregations

Calendar and seasonal quarters are exhibited in Figure 7. Figure 7a shows journey to work departure time aggregated by calendar quarter and Figure 7b shows journey to work departure time aggregated by seasonal quarter. In contrast to the journey to work travel times, the seasonal quarters for departure time graph shows less variation between the quarters and among the years of data. As shown in Figure 7c which has both the calendar and seasonal three-year aggregations, the calendar quarter aggregation reveals a slightly earlier departure time in the second quarter (represented by a point over May) and a later departure time in the third quarter (represented by a point over August). This may be intuitively satisfying and represent an eagerness to take advantage of daylight and outdoor activities (after work) in late spring/early summer and less of a need to get to work earlier at the end of the summer.

INCOME

After observing seasonality in the county-wide aggregations of journey to work travel time and departure time, we subdivide the data by income quartile, three modes, and even by considering only journeys to work within Springfield, Massachusetts city limits. There is some concern that considering smaller numbers of observations may lead to increased variability.

Journey to Work Travel Time by Income Quartiles

For the income-related graphs shown in Figures 8 and 9, income was divided into four quartiles with approximately 25% of the respondents in each household income group. For the Springfield area with approximate median income of $31,000 for the three year period survey, the first quartile of respondents had household incomes less than $20,000, the next quartile had incomes between $20,000 and $31,000, the next quartile had incomes between $31,000 and $44,500 and the top 25% of households in the county had incomes over $44,500. The Total Income Average line depicted in each survey is the same three-year aggregate line shown in Figure 4a.

In Figures 8a through 8d, the journey to work travel times for each income quartile is displayed. The journey to work travel times for respondents from below average income households for the most part had travel times that were less than the average. For low income households, some monthly travel times were almost five minutes or 20% less than the total monthly average time. Although there seems to be quite a bit of variance from month to month, the respondents in the third quartile had average travel times closest to the overall monthly average times. Respondents in the highest income quartile had travel times greater than the monthly average—five minutes or 25% higher in some cases.

We look at the weighted three year average lines for each quartile in Figure 8e to observe seasonal trends. Except for low incomes which exhibit a bit less variability than the other three quartiles, the income lines seem to follow the same positive and negative pattern of the aggregated line, but with a bit more variation. For example, the highest income quartile shows not only the highest travel times, but the highest peaks (representing longest travel times) in both January and August. Relatively low travel times for all four quartiles are found during March and only the lowest income does not show another low point for April. These are similar to our county-wide results but the income distinctions further illustrate the monthly variability in the data and extremes in the data given subdivisions (25% of the county-wide respondents to aggregate) by income quartile.

Figure 8f further illustrates the variability in the data when each income quartile is depicted over the 36 month period (January 1999 to December 2001). Here the month to month variation is striking and frustratingly, peaks and valleys don't occur in the same months for the same quartile of data. This figure is most similar to the results depicted in Seethaler and Richardson (2003) who conclude (as described above in the literature review) that they did not observe seasonal trends or monthly repeatability in travel survey data.

Journey To Work Departure Time By Income Quartiles

The journey to work departure times separated by income quartiles are found in Figures 9a through 9d. Respondents with the lowest incomes left for work later than respondents in the second, third and high income quartiles. Respondents in the second quartile left for work both above and below the monthly mean departure time values. Respondents in the third and fourth quartiles left for work earlier than the mean in most cases. This may be counter-intuitive to a hypothesis that suggests respondents with higher incomes have more flexible work departure time choice and would be more likely to leave later. But the longer travel times exhibited by the respondents in the higher income quartile (Figure 8d) do suggest having to leave earlier as shown in Figure 9d.

Again, Figure 9e showing weighted three-year average departure times for each quartile compared to the overall departure time mean demonstrates that there is more variability in the data corresponding to smaller numbers of respondents. Yet, each data point reflects approximately 150 respondents—suggesting a large enough group that aggregation would discount much of the variability caused by randomness in the data. There appears to be non-random variability in the data but seasonal hypotheses do not seem to explain much of this variability.

Similar to the results described in Figure 6a, all four income quartiles in Figure 9e show a later departure time in September relative to most of the other months. Two of the quartiles have relatively earlier departure times in March and April, but this is definitely not the case with the third quartile group that exhibits a low (earliest departure time) in June. The highest quartile shows the earlier departure times in February, May and July and the latest in June.

Figure 9f is analogous to Figure 8f and shows the trends by quartile over the 36 month period of data collection. Again, the month to month variation is striking with few similarities between adjacent points. For the most part the data fluctuates with adjacent points varying by almost thirty minutes in the low income case, for example. Twenty minute variations occur in adjacent months in the third quartile case.

MODE

One expects seasonal variations in the transportation mode chosen for the work trip, but finds that the unweighted and weighted data oscillates, especially in monthly aggregates. This is likely due to the small frequencies in alternate (non-single occupancy vehicle) modes. Bus ridership fluctuates wildly in all three data sets; even the number of respondents that carpool varies significantly from month to month. So the numbers of respondents (frequencies) that choose either drive alone, carpool or transit (bus) are not provided in this report. It seems more likely that the frequency variation is due to sampling randomness rather than seasonality.

But one can expect to capture seasonal differences by considering the journey to work travel time and departure time variables segmented by mode. As one way to capture mode variability within reason, Figures 10-13 depict travel time and departure time by mode for respondents that live AND work within Springfield, Massachusetts, the largest city in Hampden County, Massachusetts. One hopes that mode differences are reasonable for the relatively homogeneous group (about 40% of the county-wide respondents) that live and work in the county's largest city.

Monthly Journey To Work Travel Times By Mode

The weighted monthly minutes traveled to work by solo drivers who live and work in Springfield, Massachusetts are shown in Figure 10a. As can be expected, their travel times are shorter than the average travel times for the whole sample that lives and works in Springfield (shorter still than the county-wide sample depicted in Figure 4a). There are minimal differences in travel times from month to month and between years of data with the exception of noticeably longer auto travel times in October 2001.

Figure 10b shows the weighted monthly travel times of carpoolers (about 11% of the respondents that live and work in Springfield). These people all indicated that at least one other person accompanied them on their journey to work. As expected, most of these points are lower than the average travel time for respondents that live and work in Springfield. However, there is more variability from month to month. The shortest travel time for this group is in February and the longest travel time is in November.

The weighted monthly travel times of bus riders that live and work in Springfield are depicted in Figure 10c. The bus system including Springfield is the second largest in New England, but less than five percent of these urban respondents indicated that they use the bus for their journey to work. There are no other forms of public transit in the area. Travel time by bus is much longer than for drive alone cars or carpools. The bus rider travel time average is close to 27 minutes compared to the Springfield average of 19 minutes and the drive alone average of 13.5 minutes. Also, as expected with the smaller numbers of people that contribute to the bus travel time data points, there is great variability in the travel times from month to month. Knowing that many of the routes in this region have reduced (less frequent) summer schedules, it is surprising that the three year aggregate of bus travel times has the lowest value in June and highest values in both August and September (when more frequent service resumes). This may suggest that longer trips are taken during the school year months and these options may not be available during the summer (reduced bus schedules).

The three year average for each mode along with the average travel time for respondents that live and work in Springfield is provided in Figure 10d.

Quarterly Aggregations of Mode Travel Time

Figures 11a, 11b, and 11c show the calendar and seasonal quarterly aggregations of travel time by three mode categories for people that live and work in Springfield, Massachusetts. The solo driver and carpooler lines are relatively flat for both calendar and seasonal quarters. But, frustratingly, we learn different things from the calendar and seasonal aggregations of the bus transit trend line. The calendar quarter line depicted in Figure 11a and 11c shows the longest travel time in the 3rd quarter represented by July. In contrast, the longest travel time for the seasonal quarters occurs in the spring quarter (May) and the shortest bus travel time occurs in the summer (August) quarter. The different quarter aggregations for bus transit seem to give us conflicting results for quarterly travel time differences. This is likely further indication of the volatility of the bus transit monthly values.

Monthly Journey to Work Departure Times By Mode

The journey to work departure times by mode for respondents that live and work in Springfield are shown in Figures 12a, 12b, 12c, and 12d. Compared to journey to work, there is even more variability in the monthly departure times. Even with weighting the values in the total average by the number of respondents in each mode, the total average is affected by outliers in the data. Of note, all solo drivers leave much later than the average in April. Carpoolers leave earlier than the average value in March. Bus transit riders leave before and after the mean departure time values as indicated in Figure 12c, but most bus departures are earlier. Figure 12d shows the three-year average lines for each mode. In general, solo drivers and carpoolers leave later in the morning for work compared to bus transit riders and the average departure times.

Quarterly Aggregations of Mode Departure Time
Mode departure time by quarter is shown in Figure 13. Similar to the other sets of graphs, the scale on these three graphs is the same to allow for easy comparisons. Here, the seasonal quarterly lines are relatively flat. Aggregated Springfield workers leave at 7:20am each day using seasonal quarters (Figure 13b). The aggregated line in Figure 13a showing departure times for calendar quarters is not flat—in the first quarter, Springfield workers leave at approximately 7:28 p.m. in the first quarter and at 7:16 a.m. in the second quarter. Strikingly, bus transit riders leave later in the first quarter than the average and much earlier than the average in the second and third quarters. In contrast, carpoolers leave much later than average in the third quarter.

APPENDIX C

In Figures 4-13 we had many opportunities to visualize journey to work travel time and departure time from the 1999-2001 American Community Survey test site in Hampden County, Massachusetts. Additional data covering additional scenarios were analyzed in the Research Data Center and released by the Census Bureau in the disclosure process in November 2003. Figures related to the additional data are provided in Appendix C. These include monthly travel time and departure time by income for respondents that live and work in Springfield and quarterly travel time by mode and income for both county respondents and respondents that live and work in Springfield.

THREE ADDITIONAL DATA SETS/APPENDICES D AND E

As one way to study seasonality results in the American Community Survey, we examined contemporary surveys containing the same transportation variables over periods of at least a year. The National Household Transportation Survey (NHTS) collected in 2001 and 2002 and the Bay Area Transportation Survey collected in 2000 and 2001 fit the criteria and are readily available to researchers. We also use the Swiss Microcenzus survey collected in 2000. Similar to the procedure described above, we look at the weighted journey to work for all non-rural workers (with commutes less than two hours) who leave their homes between 5:00 and 10:00 a.m. We are looking at weighted data using frequency weights provided with each survey as described above. Thus, differences from the number of observations from survey to survey and variances in sampling rates should be mitigated so that seasonality effects can be observed.

Several graphics including data from three of the surveys were included in the poster presentation at the Transportation Research Board Annual Meeting in January 2004. The pages of this poster are presented in Appendix D. Similar figures including data from four different surveys are part of Parkany and Madron (2004) presented at the Progress in Activity-Based Analysis Conference in Maastricht, the Netherlands, May 2004. This paper is being considered for a book related to the conference and is included as Appendix E to this report. More details about the other data sets and the data comparisons can be found in the posters and in the paper (appendices to this report).

CONCLUSIONS AND IMPLICATIONS

Is there seasonality? For travel time, the seasonal differences that we may have expected are not obvious. Except for "spring", the analyzed data do not seem to exhibit similar patterns for either the same months or the same quarters. The three years of ACS data do not show a yearly pattern, except for possibly the "spring quarters." A close look at the departure time data may reveal a pattern, but we do not have *a priori* hypotheses related to why spring work departures may be earlier than summer departures. We expect seasonal mode differences, but if the data with several hundred observations varies so wildly, we understand that data with an average of only ten or twenty observations per month (work trip bus riders, for example), will likely be highly random in monthly or quarterly depictions.

What are the implications for future "CTPP-like" data products and transportation planning and policies? Little observed seasonal variation may be a good result. But the wildly fluctuating data streams observed here may also cause alarm about the variability in the data and the usefulness of aggregating data collected at disparate times. Further, many planners currently calibrate most of their models with the census long form data collected in April. These results show that travel times are shorter and departure times are earlier in the spring. So the models may not represent typical conditions. Separately, we may use the results here to recommend that "seasonal" quarters be used in data releases as a way to capture possible seasonal effects. We can also use these results to suggest that "calendar" quarters be used to ensure more consistent data.

The biggest lesson from these figures and the textual statistical results may be that seasonality is only one part of the concern. The "randomness" of the monthly and quarterly aggregates may be largely due to sampling differences. These differences and errors may lead to large aggregation errors.

ACKNOWLDGEMENTS

The author is extremely grateful to undergraduate assistants, Courtney Meade, Whitney Madron and Jordan Boticello, and graduate assistant, Ryan Gallagher, for their valiant help with this project. There would not have been a report without them. This project has been funded and supported by the Federal Highway Administration and Elaine Murakami. Her comments are incorporated here.

REQUIRED CES DISCLAIMER

Part of the research in this report was conducted while the author was a research associate at the Center for Economic Studies, U.S. Bureau of the Census. Research results and conclusions expressed are those of the author and do not necessarily indicate concurrence by the Bureau of the Census or the Center for Economic Studies.

REFERENCES

Alexander, C.H. "A Discussion of the Quality of Estimates from the American Community Survey For Small Population Groups." Draft presented at the Census Advisory Committee Meetings, September 2002. Accessible at http://www.census.gov/acs/www/SBasics/op_plan.htm. Accessed July 30, 2003.

Alexander, C.H. "Still Rolling: Leslie Kish's 'Rolling Samples' and the American Community Survey." Invited paper session at Statistics Canada's Symposium 2001. To appear in proceedings and be submitted to Survey Methodology. Accessible at http://www.census.gov/acs/www/SBasics/op_plan.htm. Accessed July 30, 2003.

Alexander, C.H. "A Rolling Sample Survey for Yearly and Decennial Uses." Presented at a contributed session on combining populations, organized by Leslie Kish at the meetings of the International Statistical Institute, in Helsinki, Finland (1999). Accessible at http://www.census.gov/acs/www/SBasics/op_plan.htm. Accessed July 30, 2003.

Alexander, C.H. "A Continuous Measurement Alternative for the U.S. Census." Proceedings of the American Statistical Association Survey Research Methods Section, pp. 486-491 (1993).
Accessible at http://www.census.gov/acs/www/SBasics/op_plan.htm. Accessed July 30, 2003.

Bureau of Transportation Statistics (BTS), "A Time Series Analysis of U.S. Inland Waterways Trade" in BTS Transportation Indicators, Dec 2002. Found at http://www.bts.gov/publications/transportation_indicators/december_2002/Special/html/A_Time_Series_Analysis_of_US_Inland_Waterways_Trade.html. Accessed June 30, 2004.

Bureau of Transportation Statistics (BTS), "Transportation Services Index" and "Terminology: Transportation Services Index", 2004. Found at http://www.bts.gov/xml/tsi/src/index.xml. Accessed June 30, 2004.

Census Transportation Planning Package 2000 (CTTP) web site: http://www.fhwa.dot.gov/ctpp/about.htm Accessed September 13, 2004.

Chand, N. and Alexander, C.H. "Small Area Estimation with Administrative Records and Continuous Measurement." Proceedings of the American Statistical Association Survey Research Methods Section, pp. 870-875 (1996). Accessible at http://www.census.gov/acs/www/SBasics/op_plan.htm. Accessed July 30, 2003.

Christopher, E., E. Murakami, S. Riklin, and N. Srinivasan, "The Long Form/American Community Survey—Questions and Their Relevance to Transportation" Long Form Content Justification prepared by the U.S. Department of Transportation and accessed at http://www.fhwa.dot.gov/ctpp/omb2.htm. Accessed Aug. 1, 2003.

The National Academies Press. "A Census that Mirrors America: Interim Report." 1993. http://books.nap.edu/books/0309049792/html/81.html. Accessed July 29, 2003.

Niemeier, Debbie, report submitted to Federal Highway Administration related to using demographic data at the Census Bureau Research Data Centers, 2004.

Oregon Department of Transportation, "Developing Design Hour Volumes" in the Transportation Planning Analysis Unit Training Manual, 2001. Accessible at www.odot.state.or.us/tddtpau/SysAnalysis.html. Accessed at June 29, 2004.

Parkany, Emily, "Seasonality of Transportation Data," poster presented at Transportation Research Board Annual Meeting, January 2004.

Parkany, E. and W. Madron, "Seasonality of Transportation Data: A Comparison of Four Data Sets," paper presented at Progress in Activity-Based Analysis, Maastricht, the Netherlands, May 2004.

Quesinberry, Elaine, "American Community Survey Alert Number 10", email newsletter: ACS Alert, February 20, 2003.

Scarr, Harry A. "Continuous Measurement – Harry A. Scarr to Association of Public Data." Speech presented to the Association of Public Data Users (APDU) at Washington, DC. October 1994. Accessible at http://www.census.gov/acs/www/AdvMeth/Papers/ACS/Paper1.htm. Accessed July 29, 2003.

Seethaler, R.K. and Richardson, A.J. (December, 2003). "Seasonality in Daily Travel Patterns". TUTI Report TR 24-2003, a report to the Victorian Department of Infrastructure (pdf, 1.3mb) found at: http://www.tuti.com.au/PUBLICATIONS/2003/2003papers.htm. Accessed June 29, 2004.

U.S. Census Bureau, Demographic Surveys Division, Continuous Measurement Office. What is the American Community Survey? April 07, 2003. http://www.census.gov/acs/www/SBasics/What/What1.htm. Accessed July 30, 2003.

U.S. Census Bureau. "American Community Survey Operations Plan." March 2003. http://www.census.gov/acs/www/Downloads/OpsPlanfinal.pdf. Accessed July 29, 2003.

US Department of Transportation, Bureau of Transportation Statistics, *Implications of Continuous Measurement for the Uses of Census Data in Transportation Planning*, Washington D.C., April 1996.

List of Figures

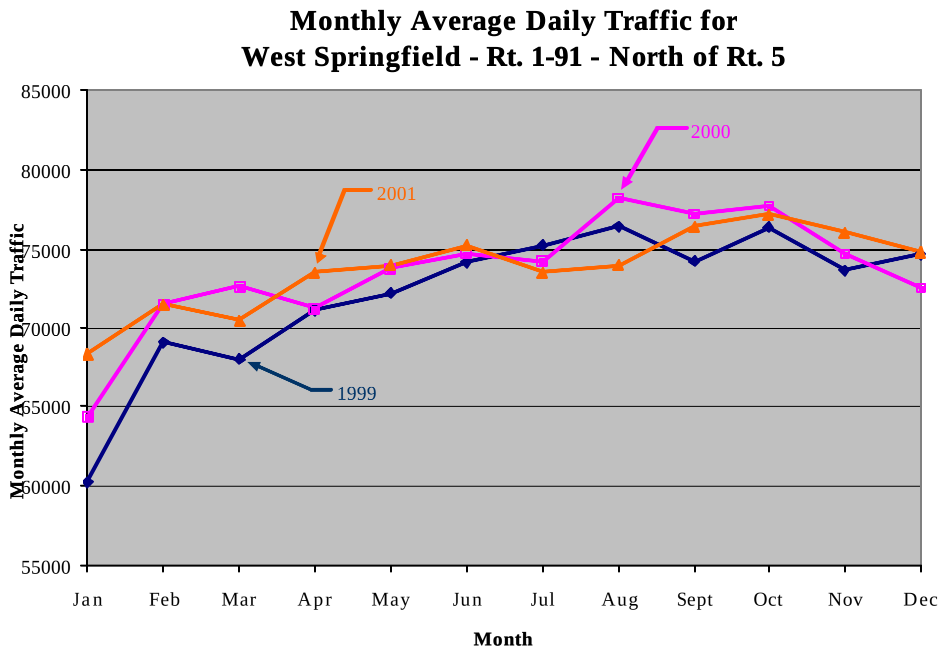

FIGURE 1 – Monthly Average Traffic Volumes at a Federal Monitoring Site in Hampden County, MA

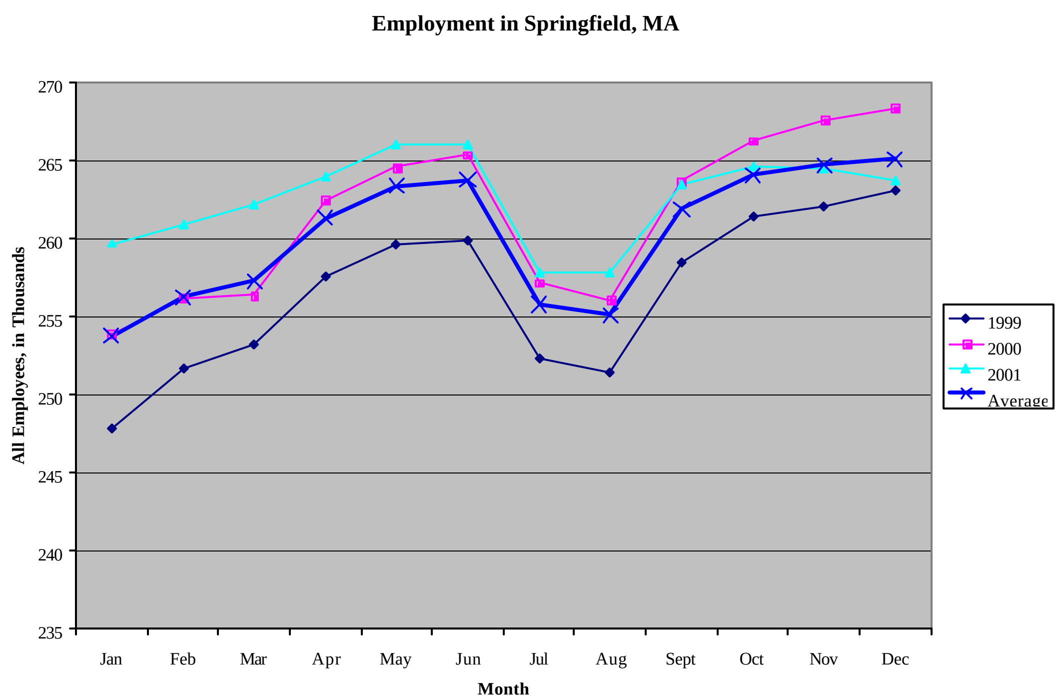

FIGURE 2 – Springfield, MA Employment (from Bureau of Labor Statistics accessed June, 2004)

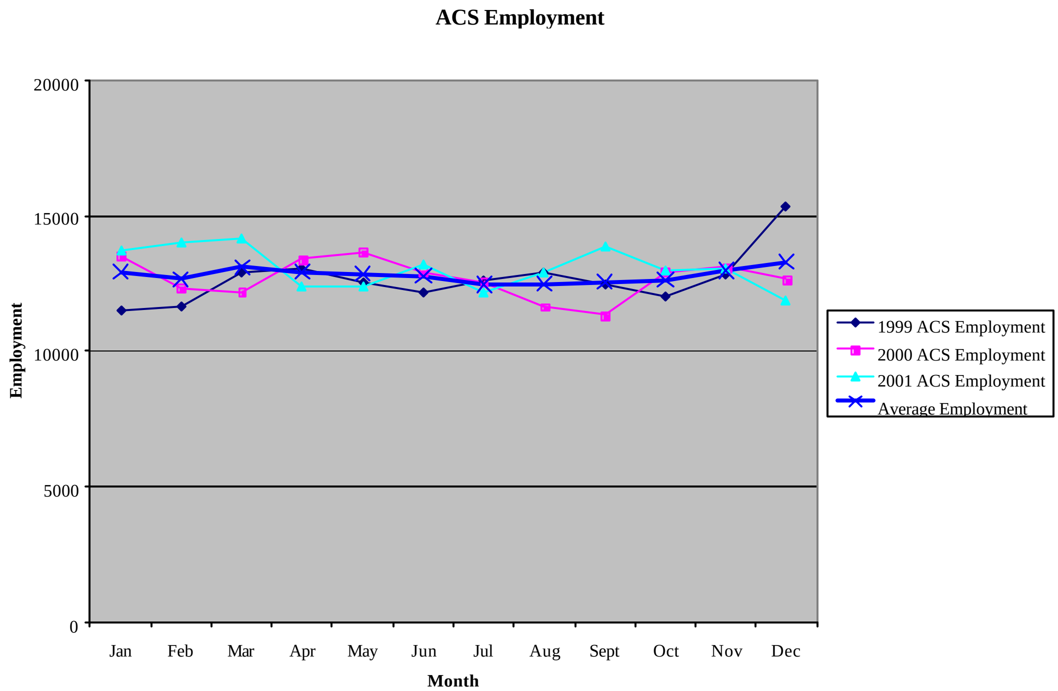

FIGURE 3 – ACS Employment in Hampden County, MA

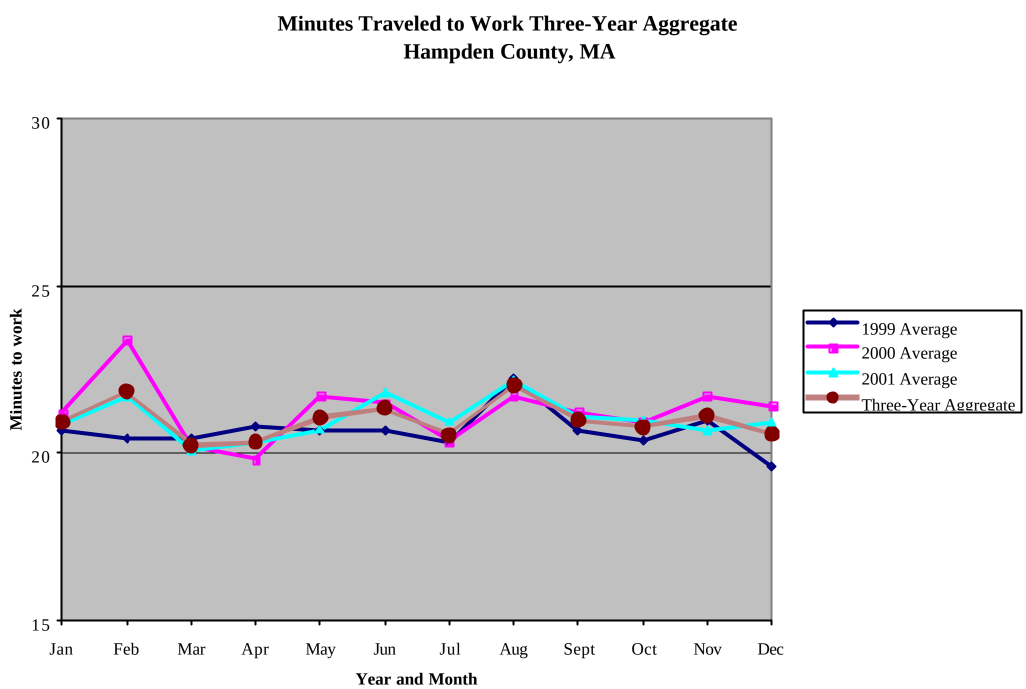

FIGURE 4a – Three Year Averages and Weighted Three Year Aggregate Minutes Traveled to Work

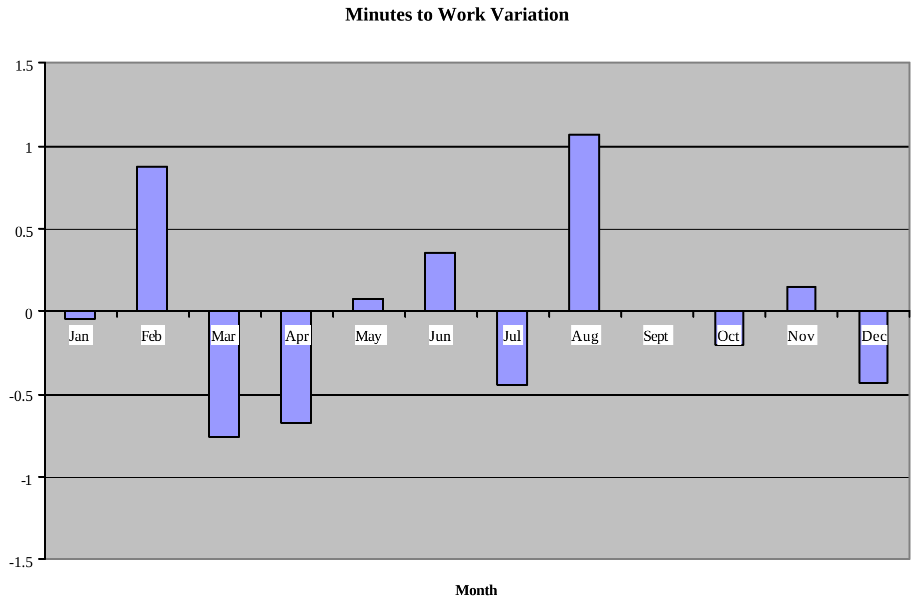

FIGURE 4b – Monthly Variation from the Mean of Journey to Work Travel Times

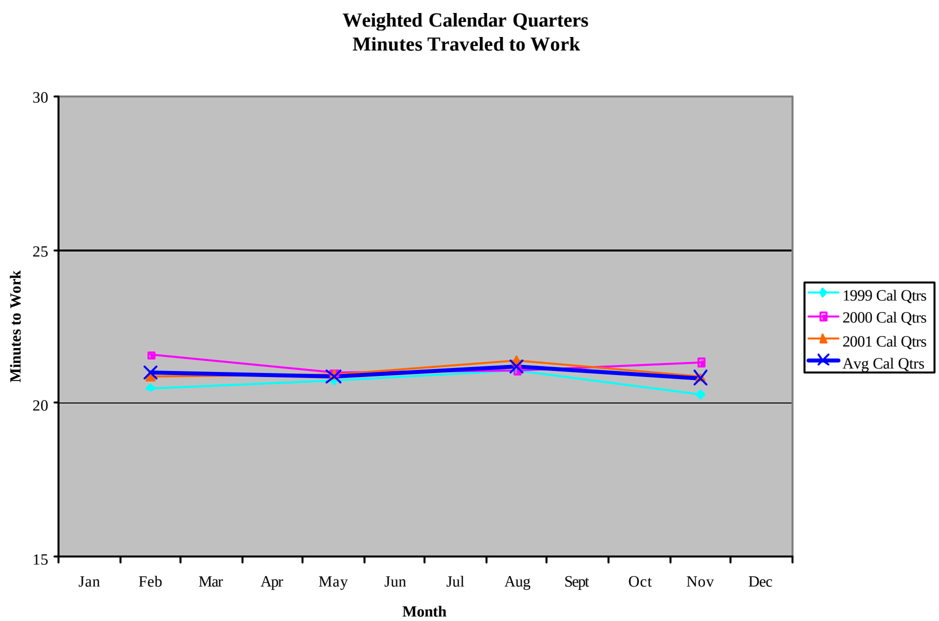

FIGURE 5a – Weighted Calendar Quarters Minutes Traveled to Work

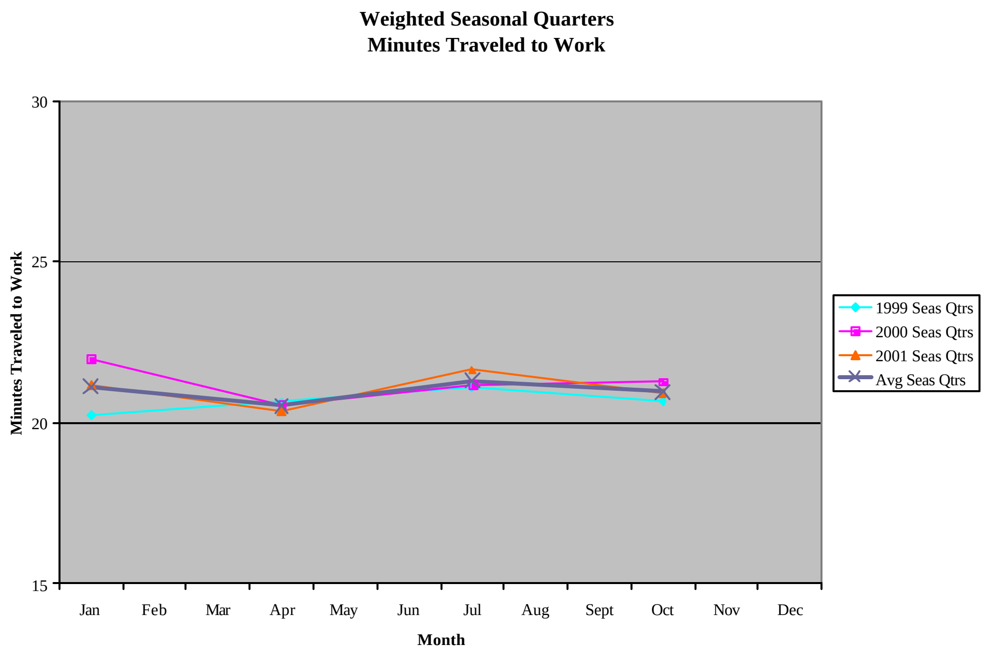

FIGURE 5b – Weighted Seasonal Quarters Minutes Traveled to Work

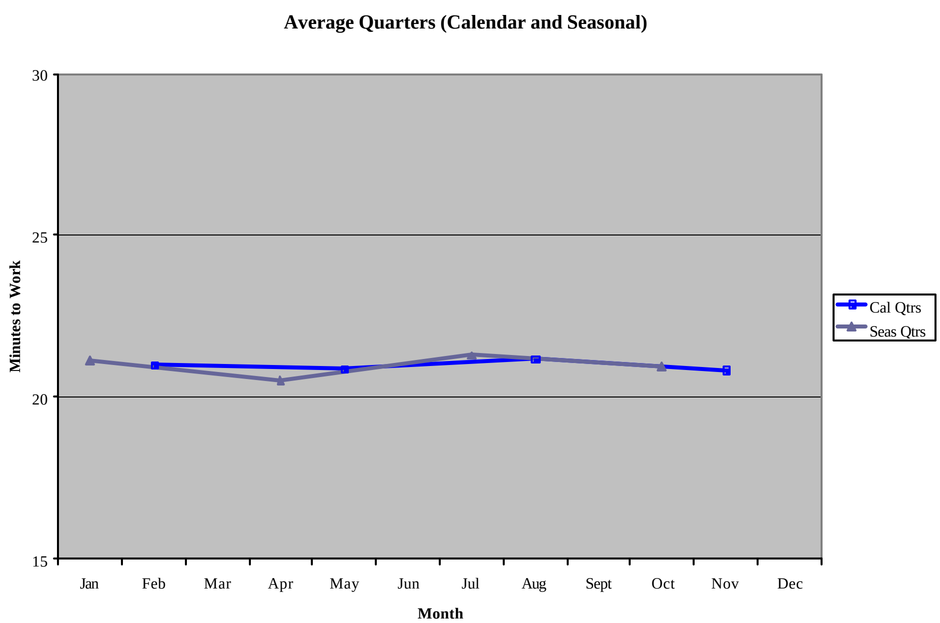

FIGURE 5c – Average Calendar and Seasonal Quarter Journey to Work Travel Time

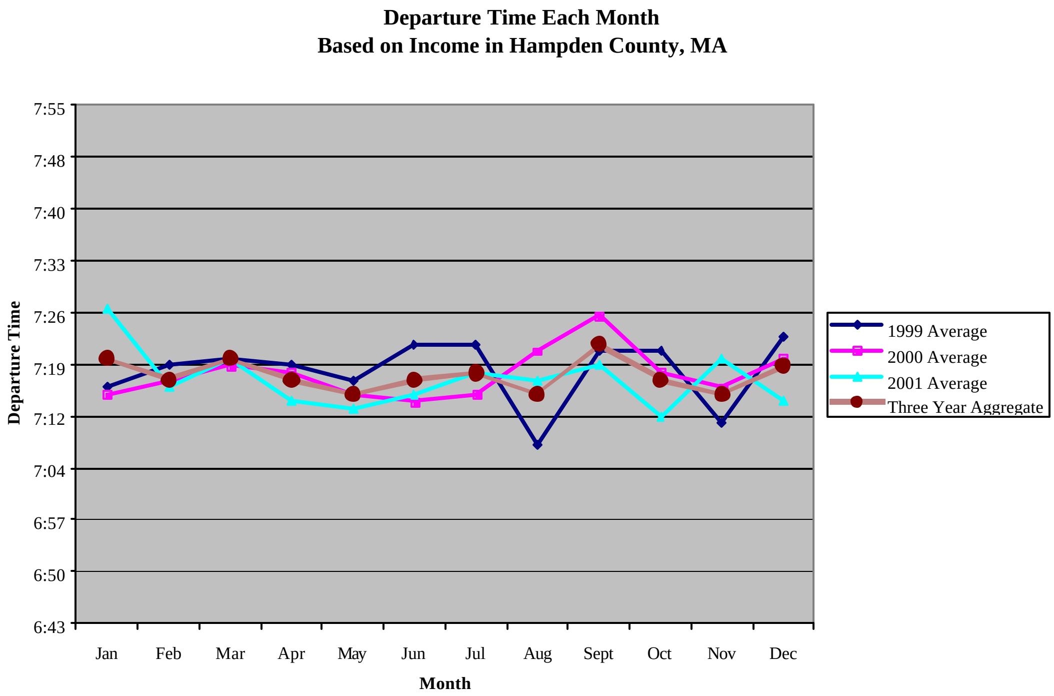

FIGURE 6a – Three Year Averages and Three Year Aggregated Monthly Journey to Work Departure Time

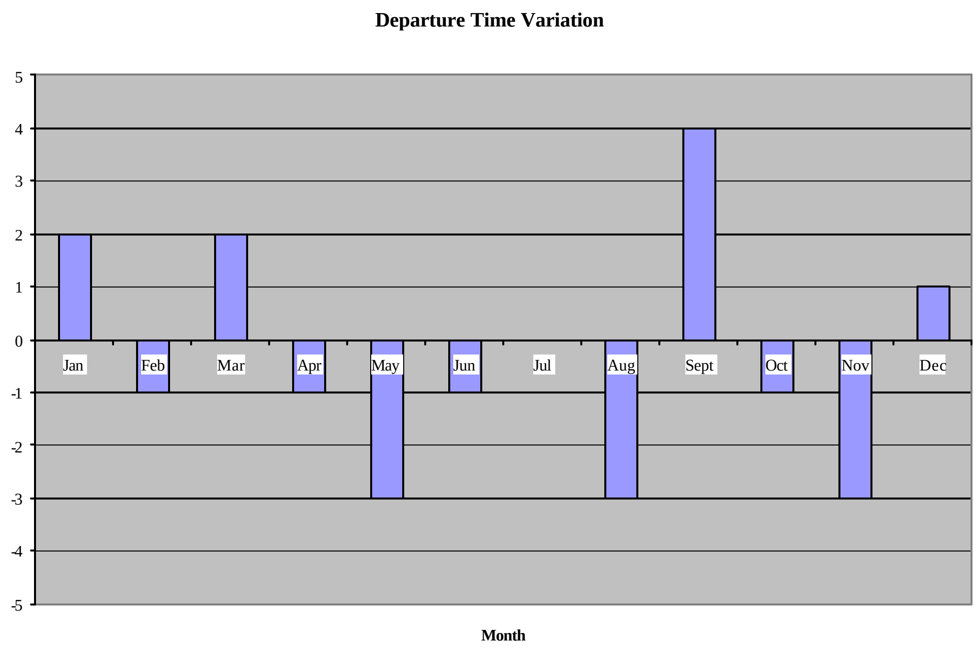

FIGURE 6b – Monthly Variation from the Mean of Journey to Work Departure Time

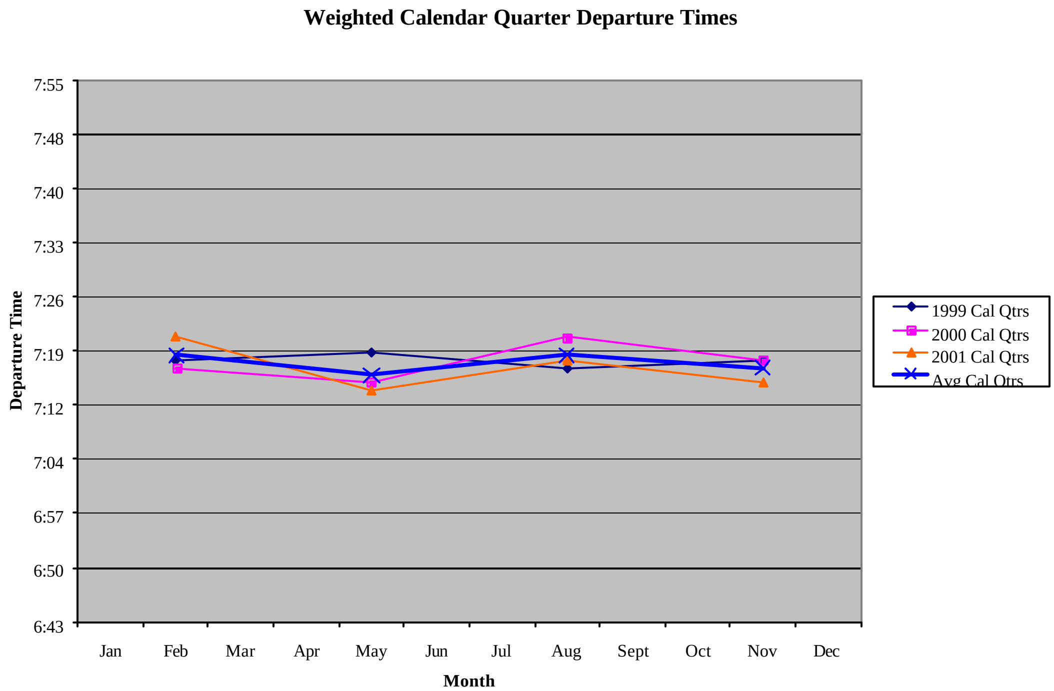

FIGURE 7a – Weighted Calendar Quarter Journey to Work Departure Times

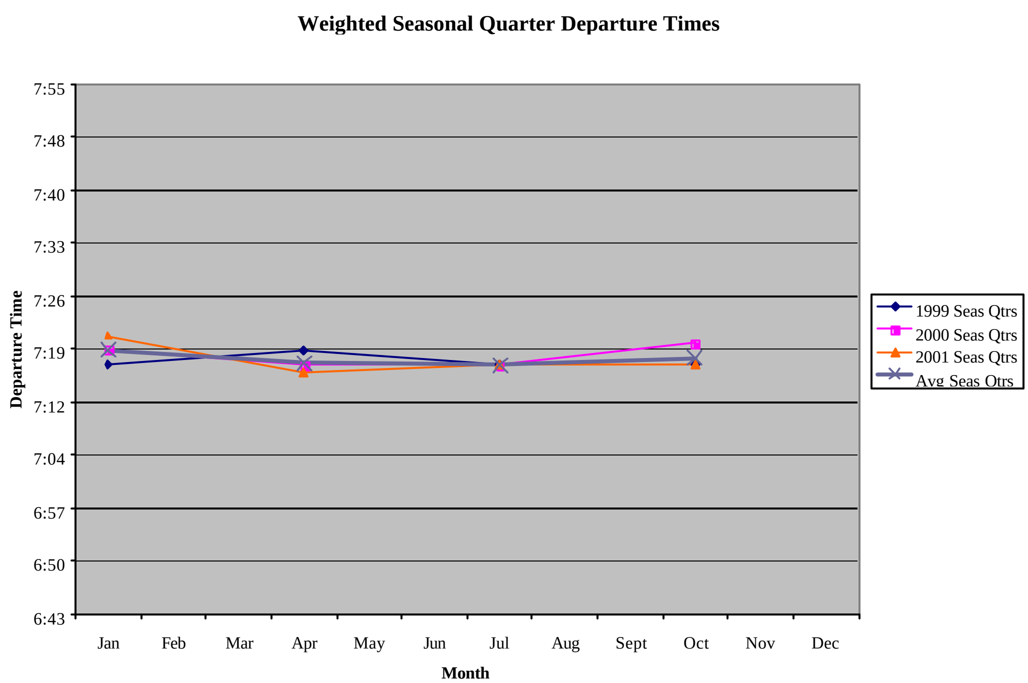

FIGURE 7b – Weighted Seasonal Quarter Journey to Work Departure Times

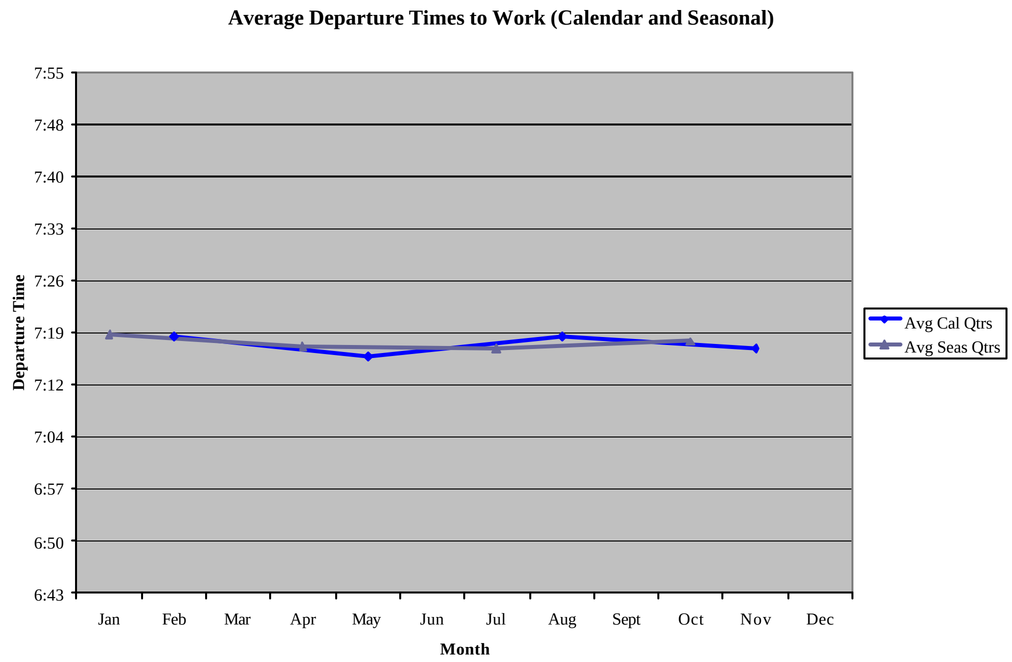

FIGURE 7c – Average Calendar and Seasonal Quarter Journey to Work Departure Time

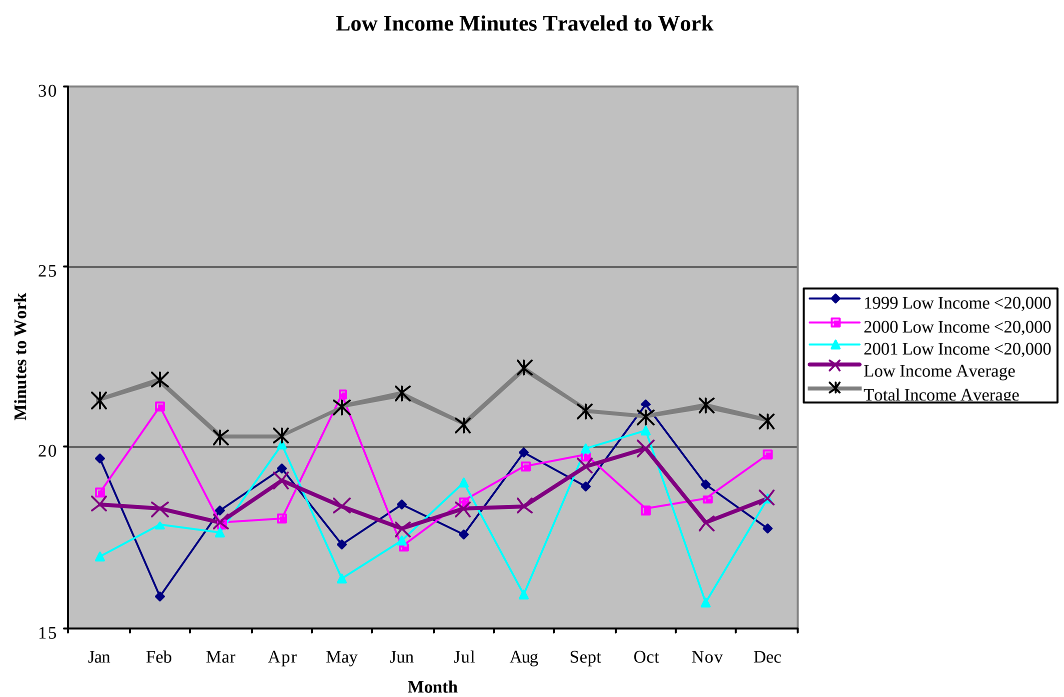

FIGURE 8a – Low Income Minutes Traveled to Work

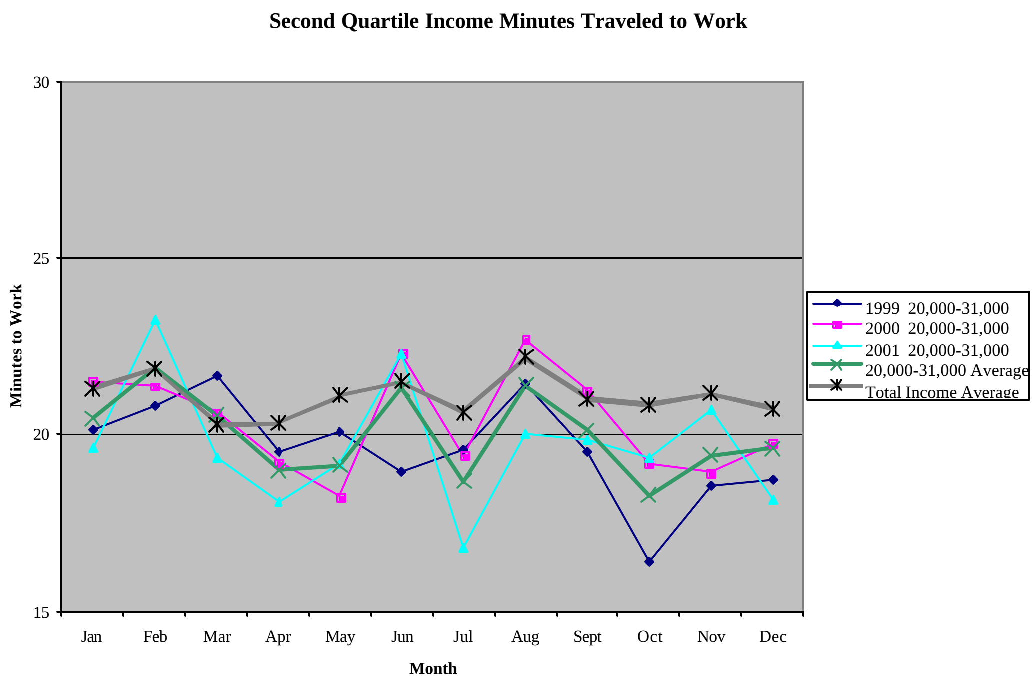

FIGURE 8b – Second Quartile Income Minutes Traveled to Work

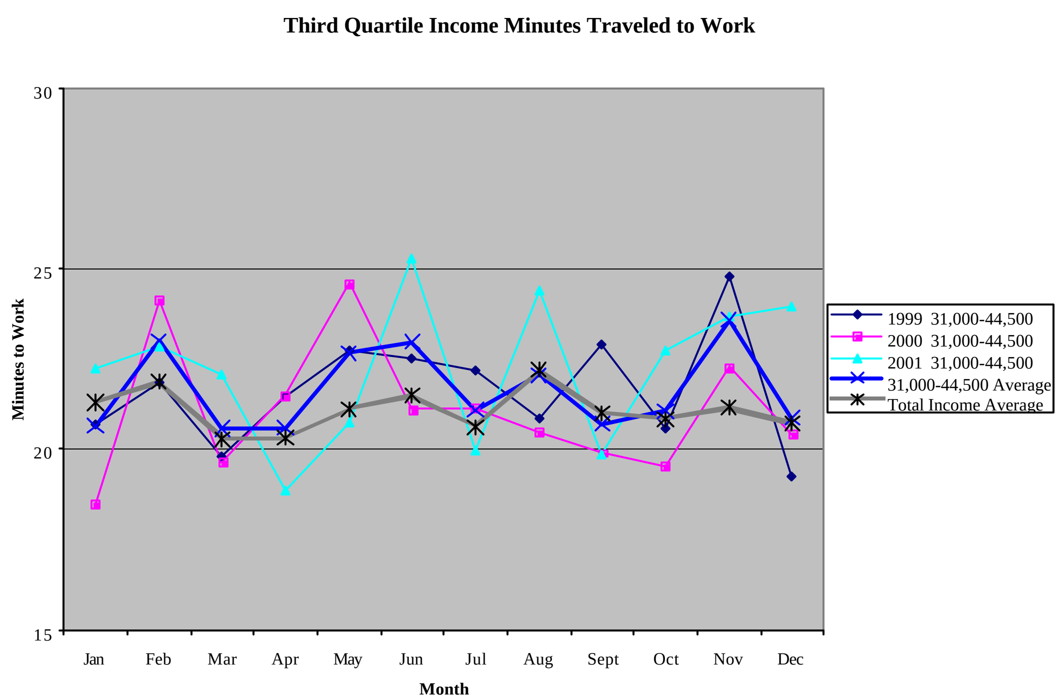

FIGURE 8c – Third Quartile Income Minutes Traveled to Work

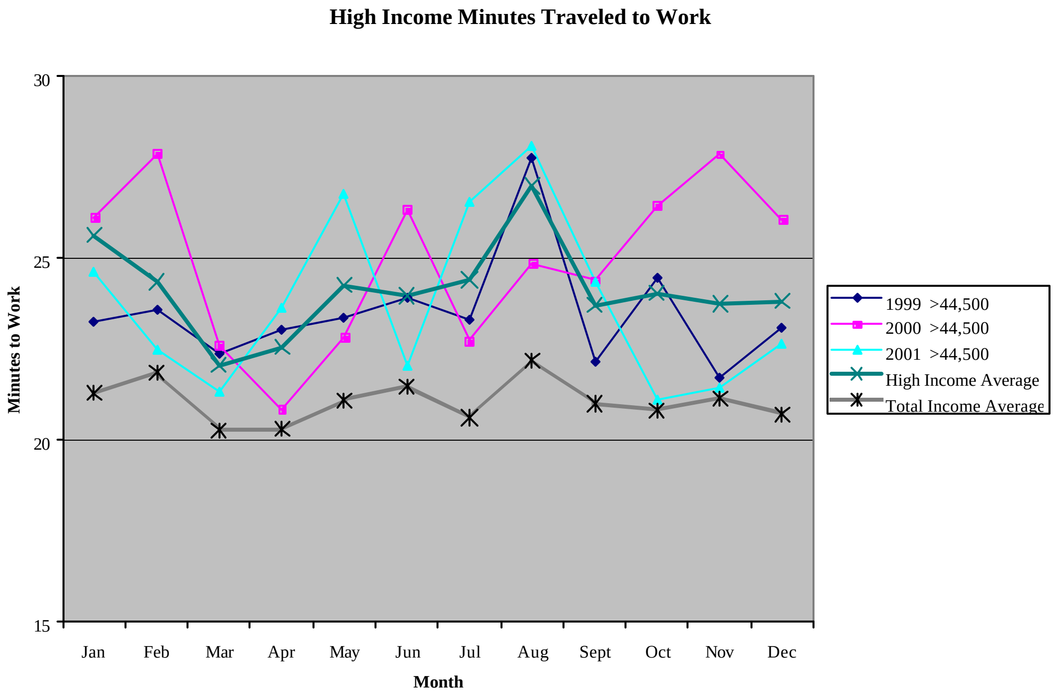

FIGURE 8d – High Income Minutes Traveled to Work

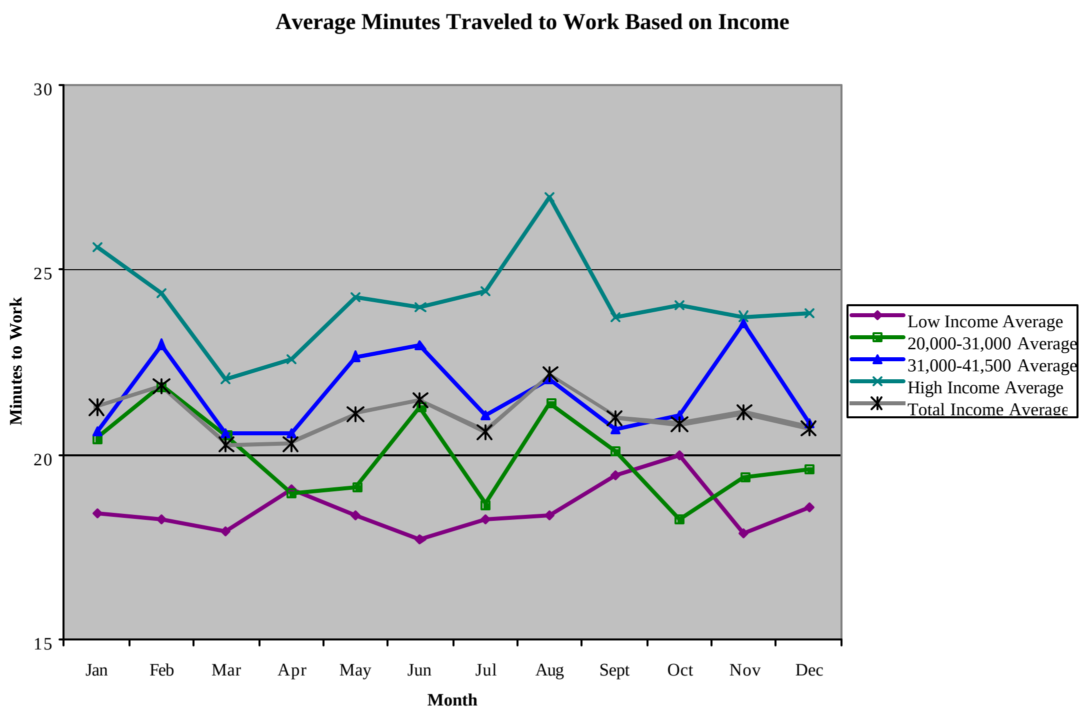

FIGURE 8e – Average Minutes Traveled to Work Based on Income

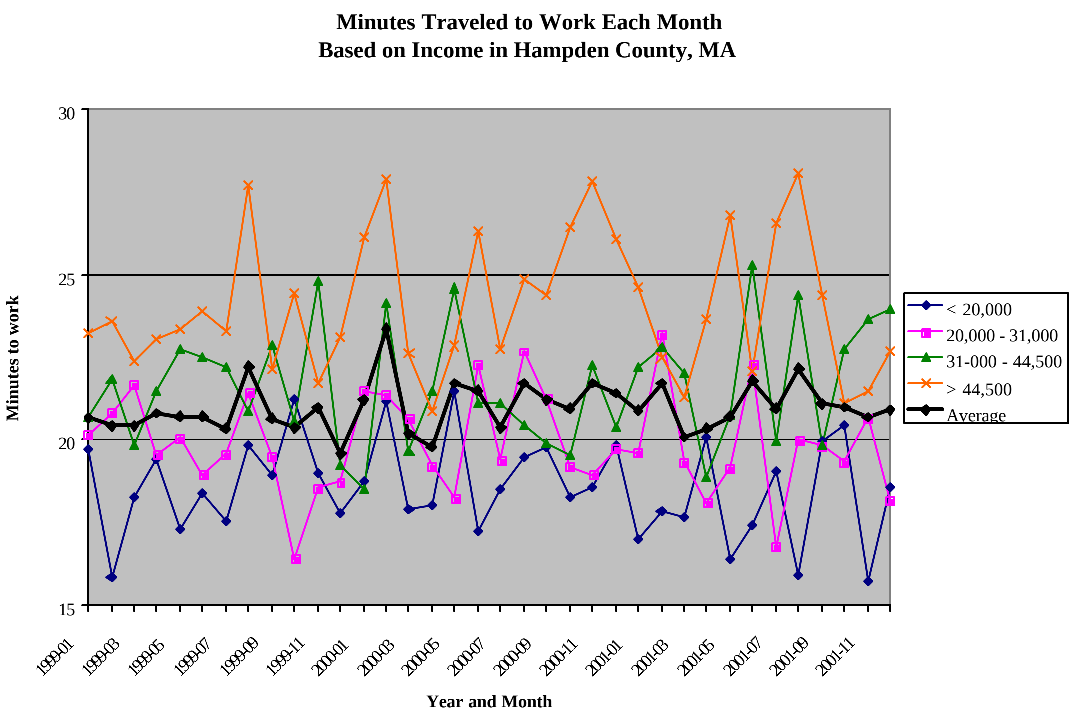

FIGURE 8f – Minutes Traveled to Work Each Month Based on Income in Hampden County, MA

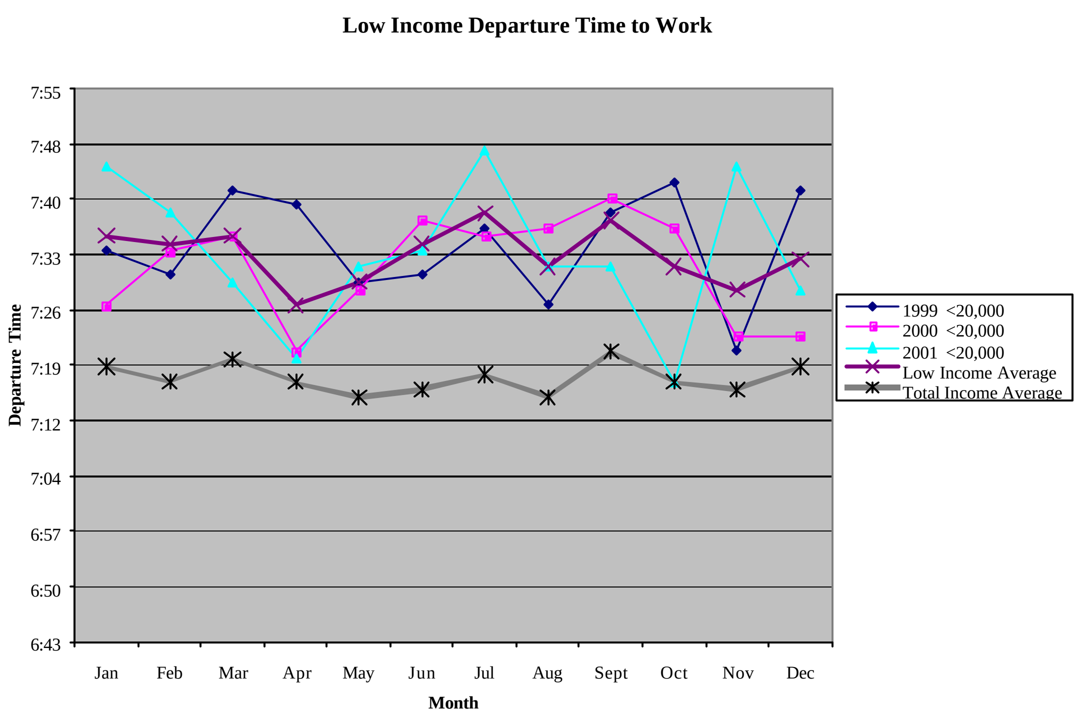

FIGURE 9a – Low Income Departure Times to Work

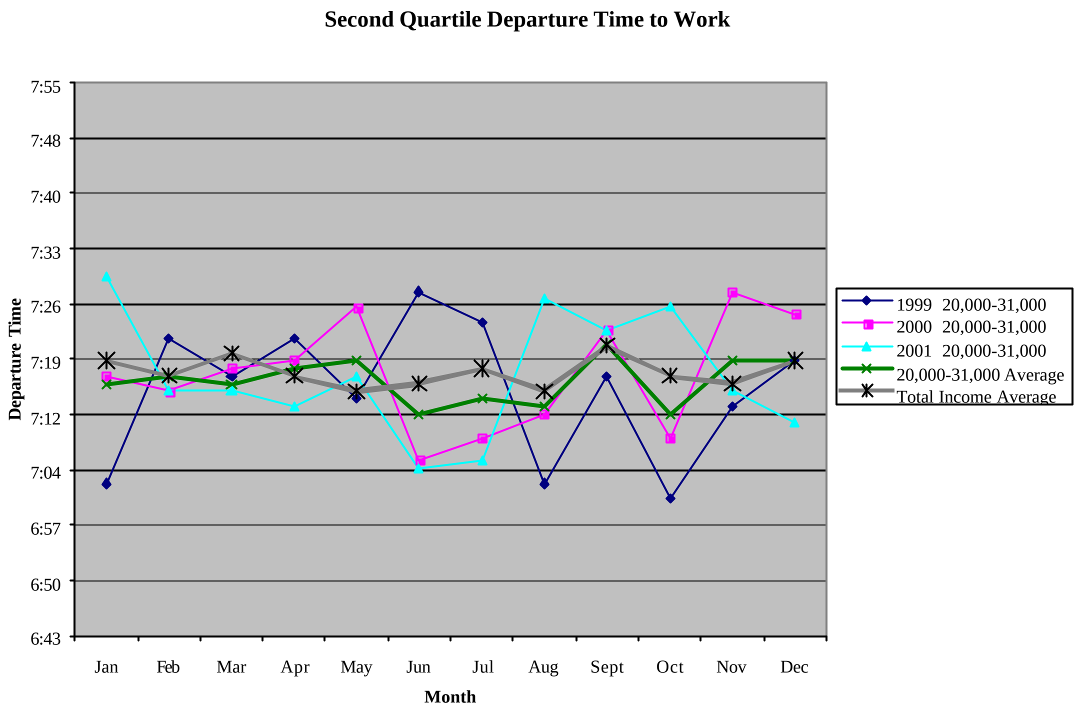

FIGURE 9b – Second Quartile Departure Times to Work

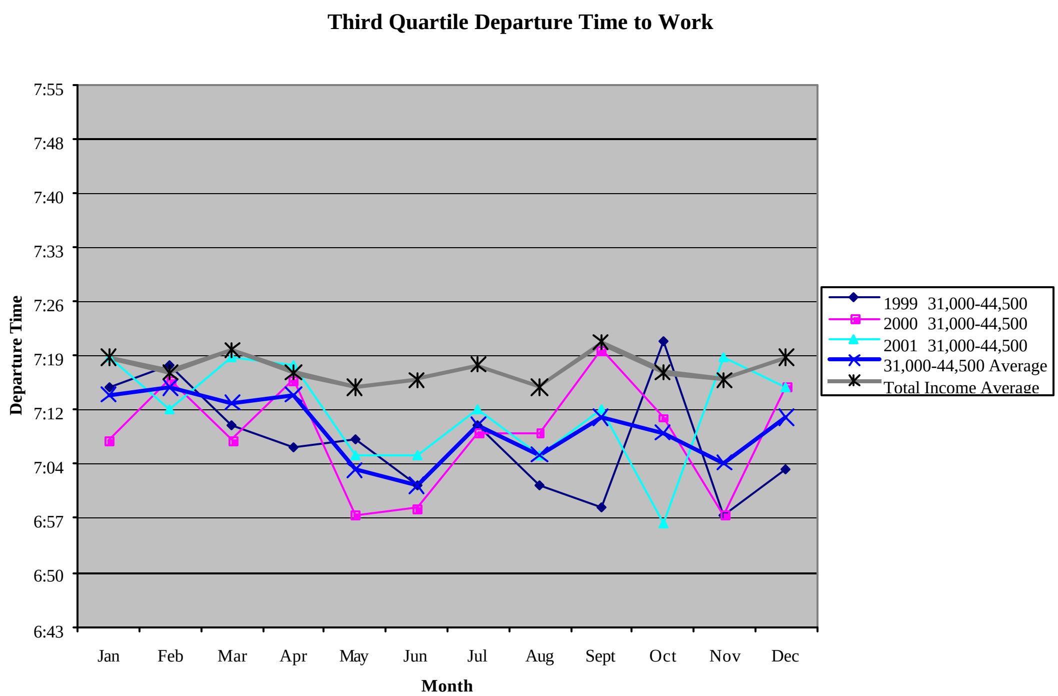

FIGURE 9c – Third Quartile Departure Times to Work

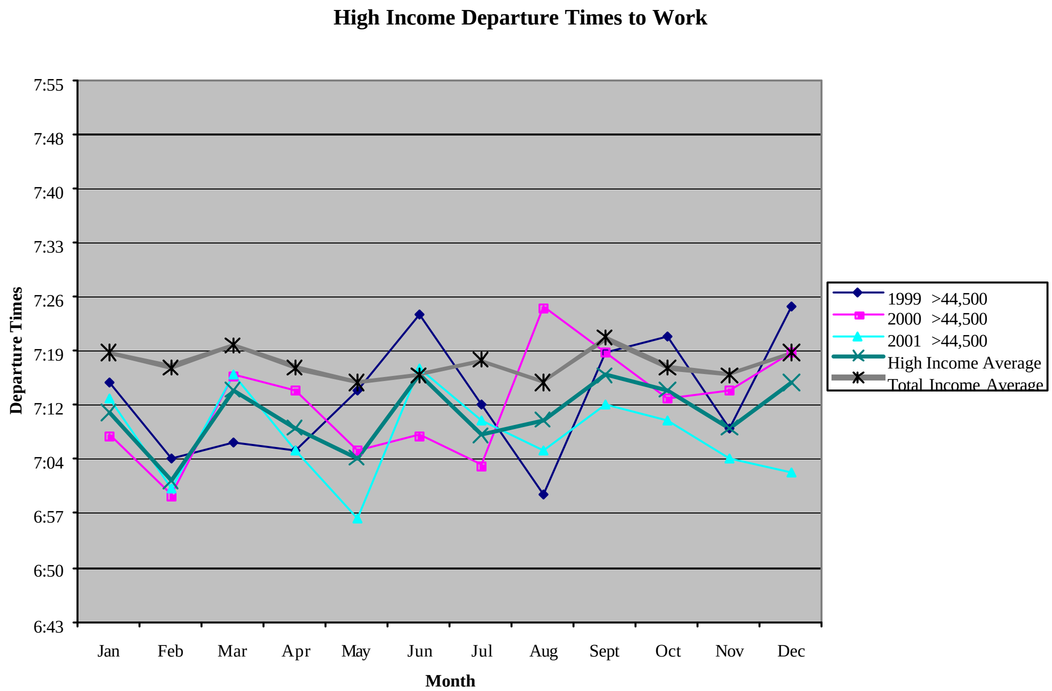

FIGURE 9d – High Income Departure Times to Work

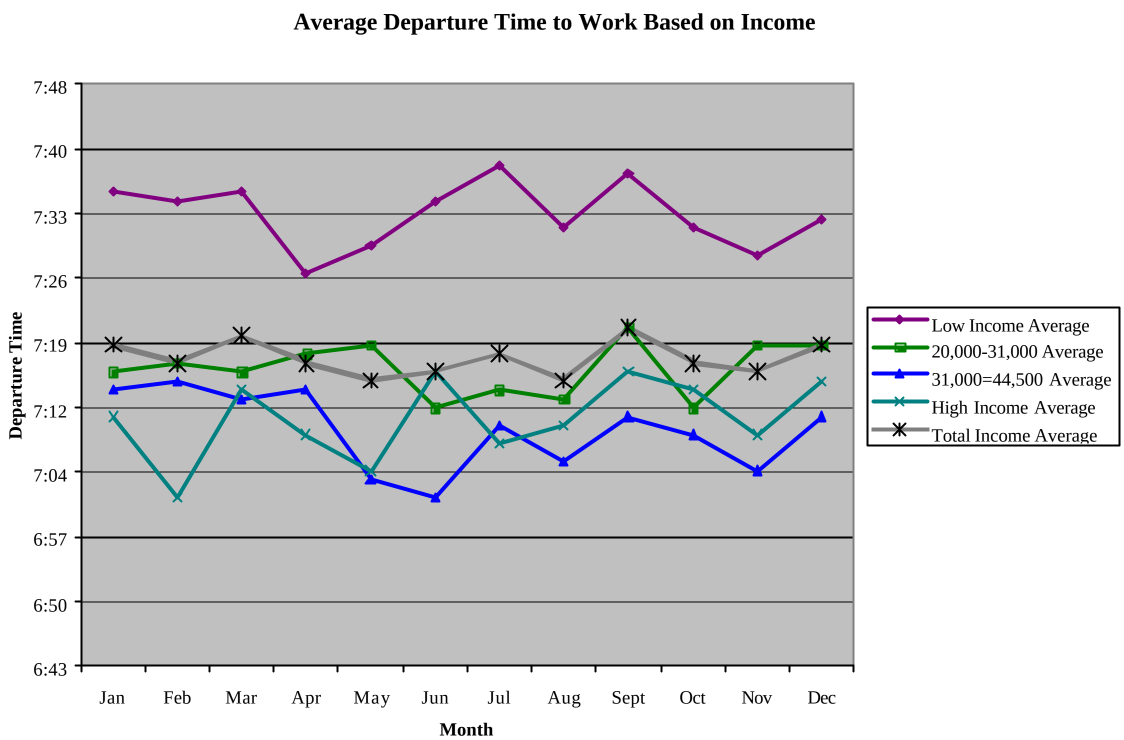

FIGURE 9e – Average Departure Times to Work Based on Income

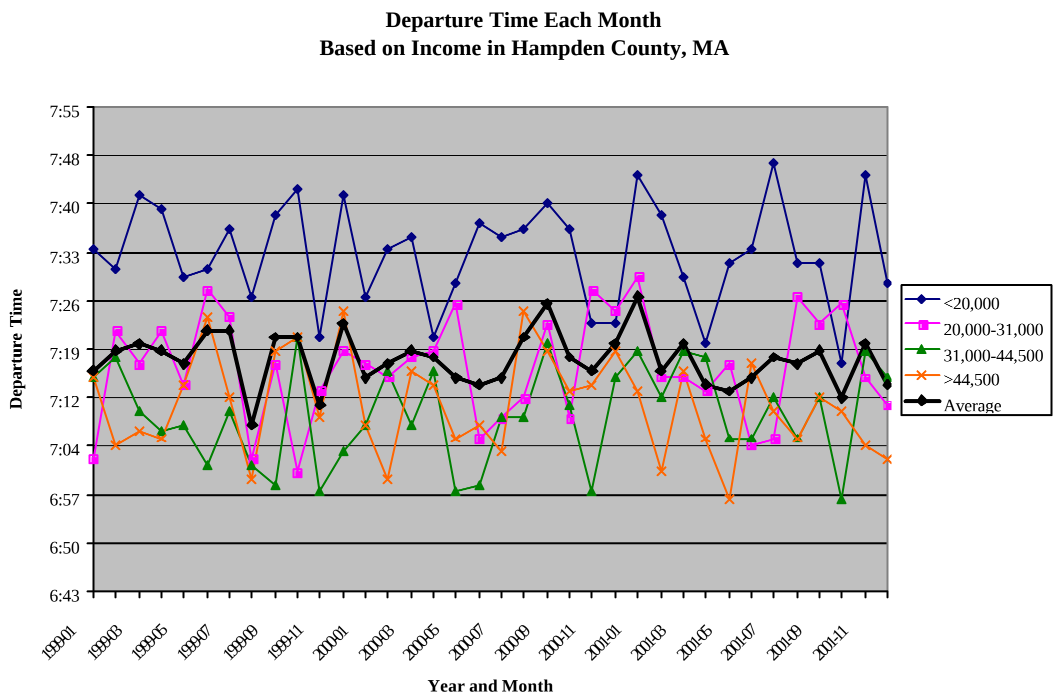

FIGURE 9f – Departure Time to Work Each Month Based on Income in Hampden County, MA

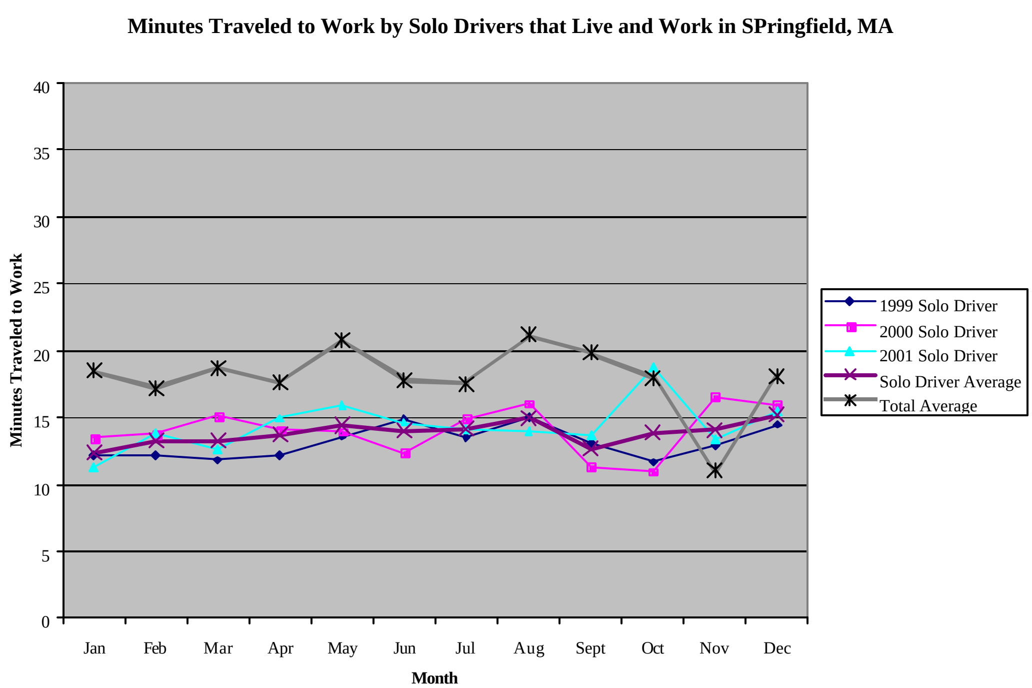

FIGURE 10a – Minutes Traveled to Work by Solo Drivers that Live and Work in Springfield, MA

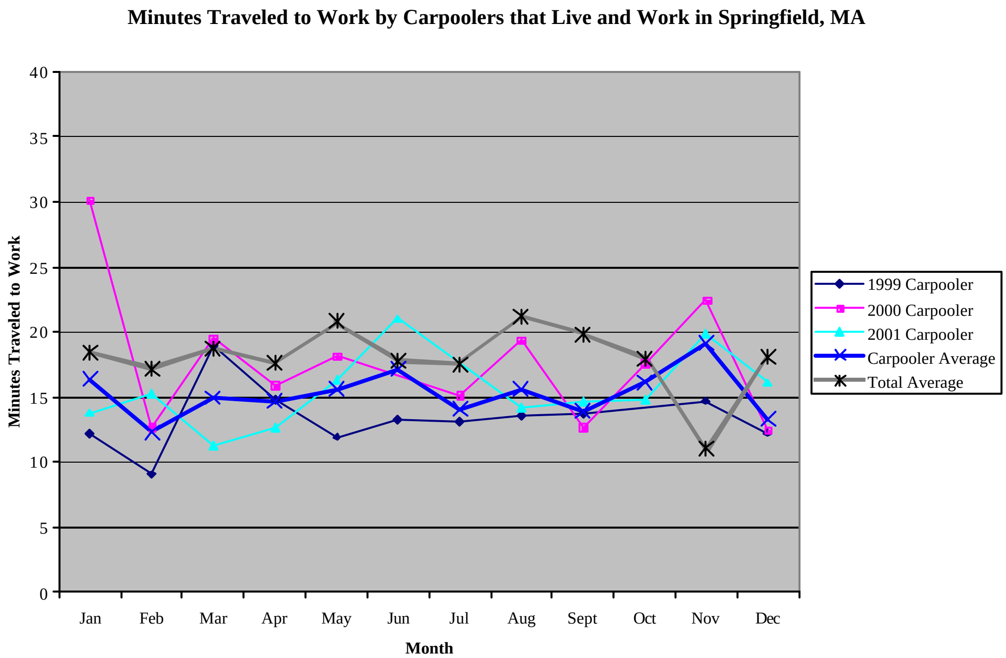

FIGURE 10b – Minutes Traveled to Work by Carpoolers that Live and Work in Springfield, MA

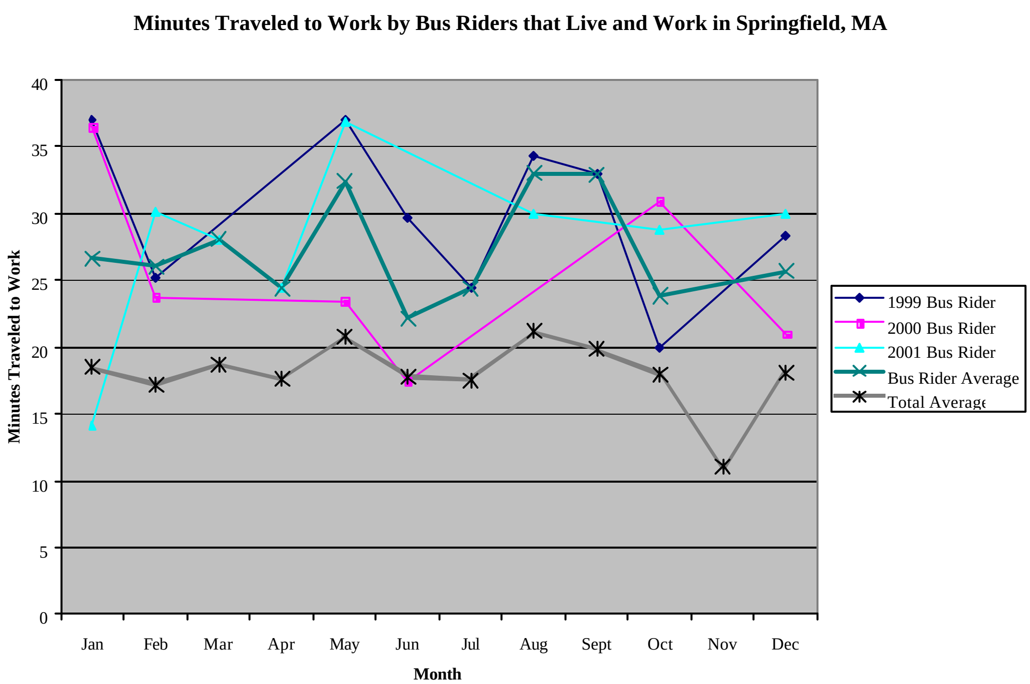

FIGURE 10c – Minutes Traveled to Work by Bus Riders that Live and Work in Springfield, MA

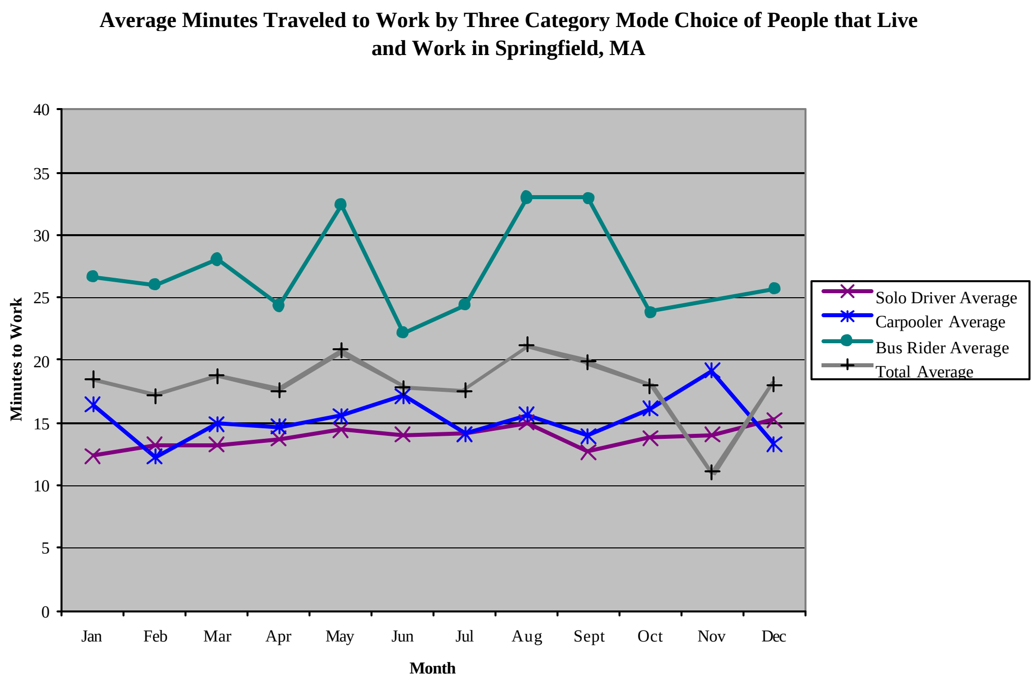

FIGURE 10d – Average Minutes Traveled to Work by Three Category Mode Choice of People that Live and Work in Springfield, MA

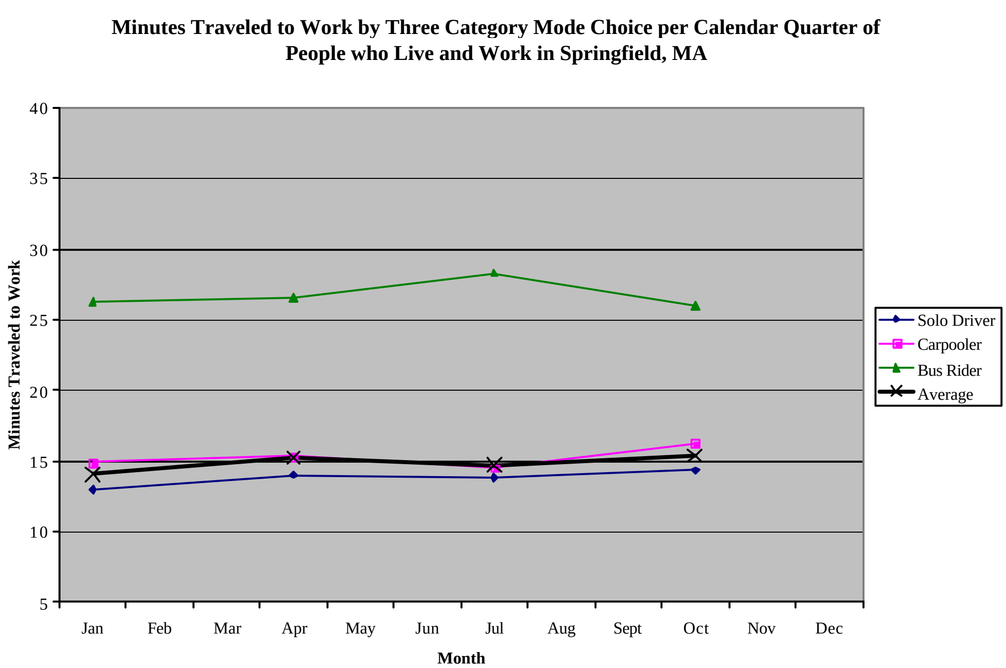

FIGURE 11a – Minutes Traveled to Work by Three Category Mode Choice per Calendar Quarter of People who Live and Work in Springfield, MA

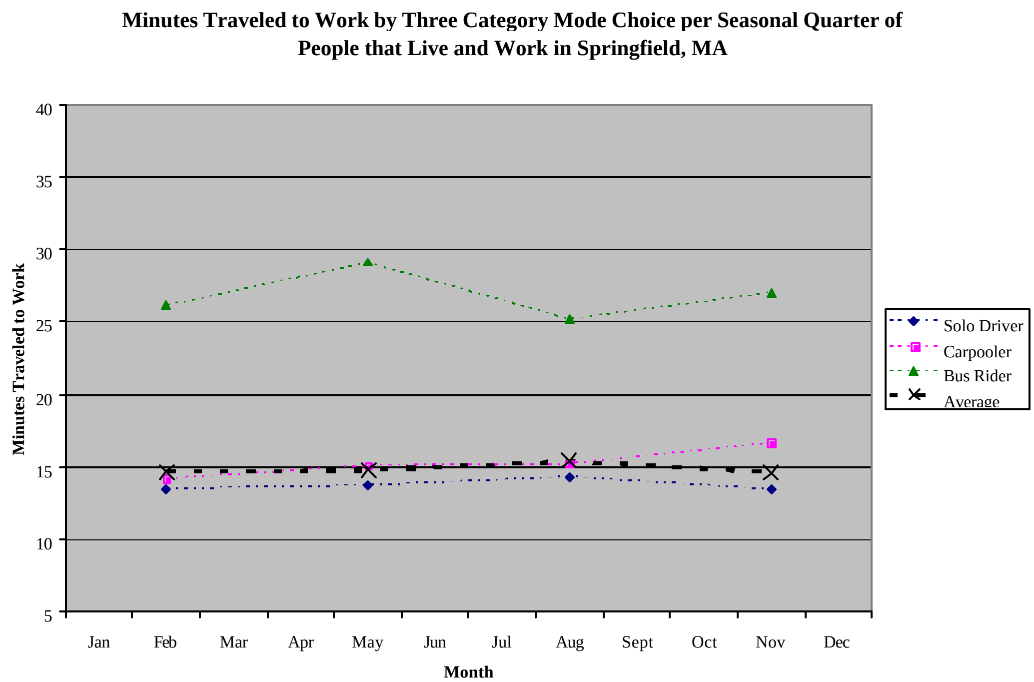

FIGURE 11b - Minutes Traveled to Work by Three Category Mode Choice per Seasonal Quarter of People who Live and Work in Springfield, MA

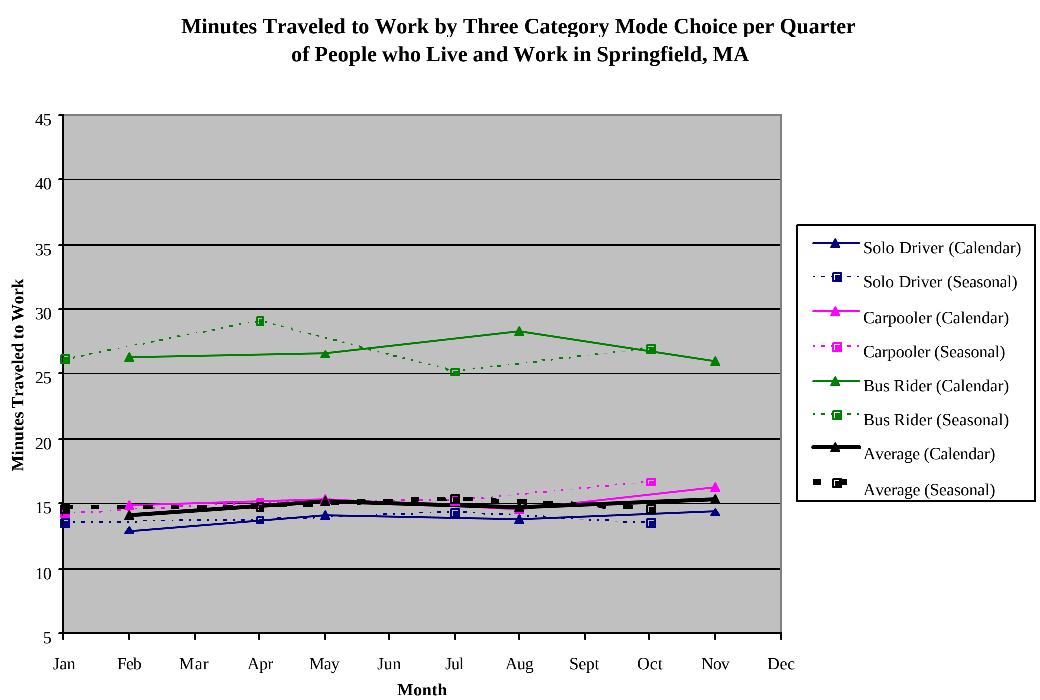

FIGURE 11c – Minutes Traveled to Work by Three Category Mode Choice per Quarter of People who Live and Work in Springfield, MA

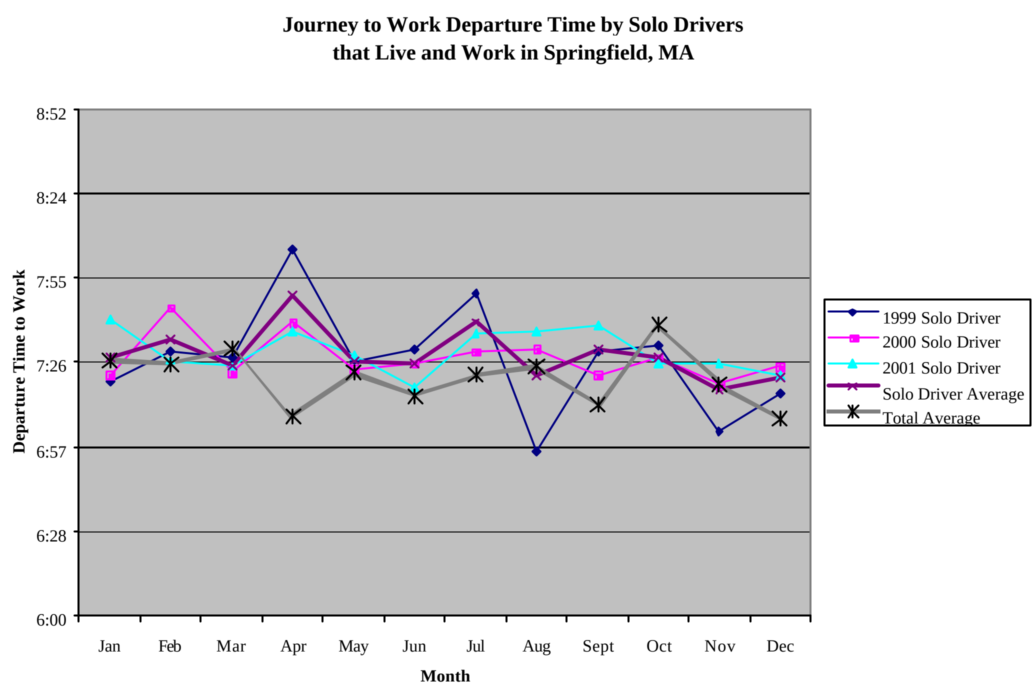

FIGURE 12a – Journey to Work Departure Time by Solo Drivers that Live and Work in Springfield, MA

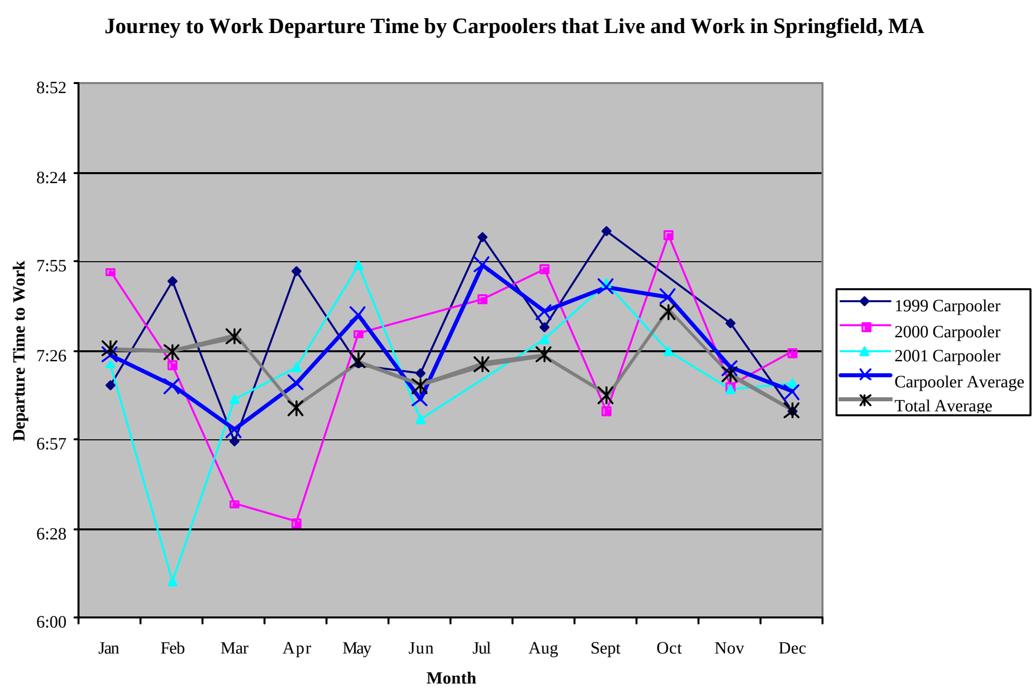

FIGURE 12b – Journey to Work Departure Time by Carpoolers that Live and Work in Springfield, MA

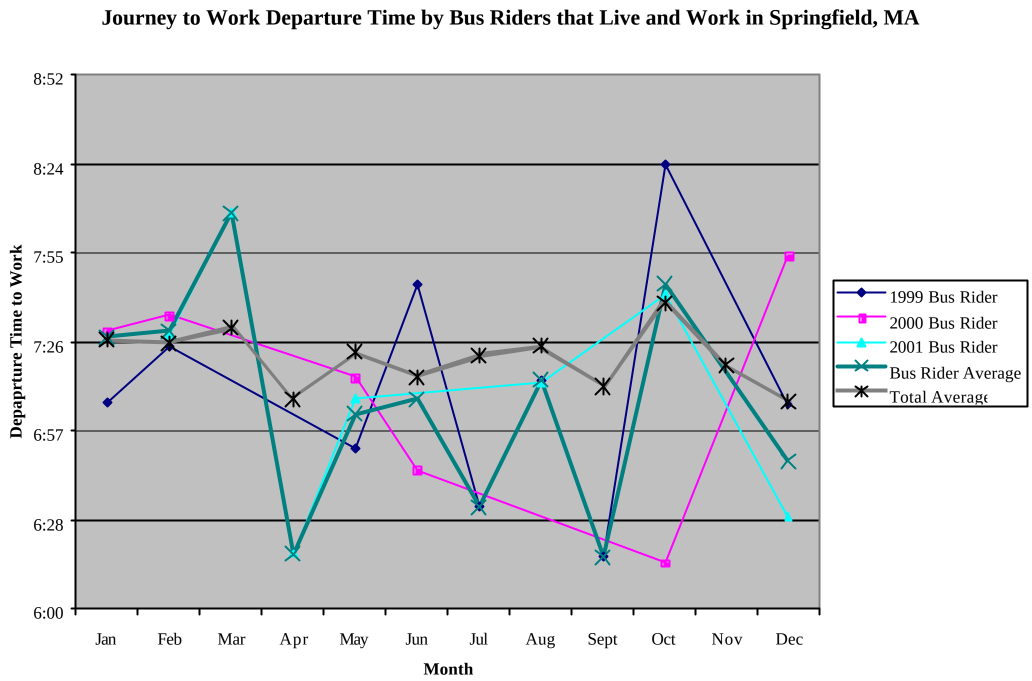

FIGURE 12c – Journey to Work Departure Time by Bus Riders that Live and Work in Springfield, MA

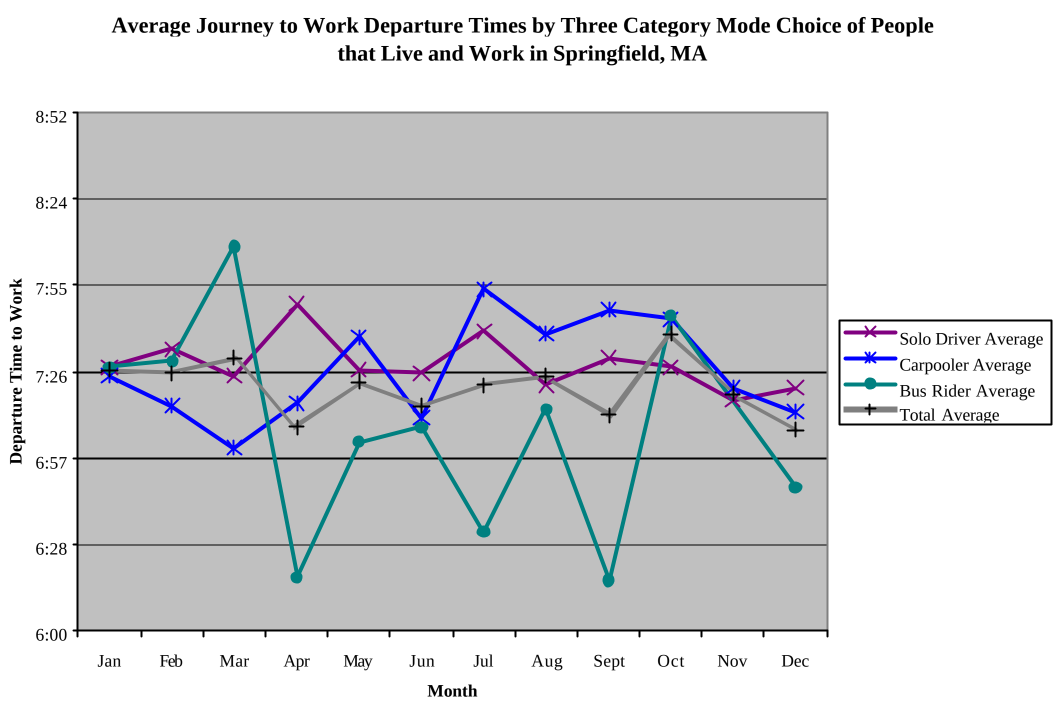

FIGURE 12d – Average Journey to Work Departure Times by Three Mode Choice of People that Live and Work in Springfield, MA

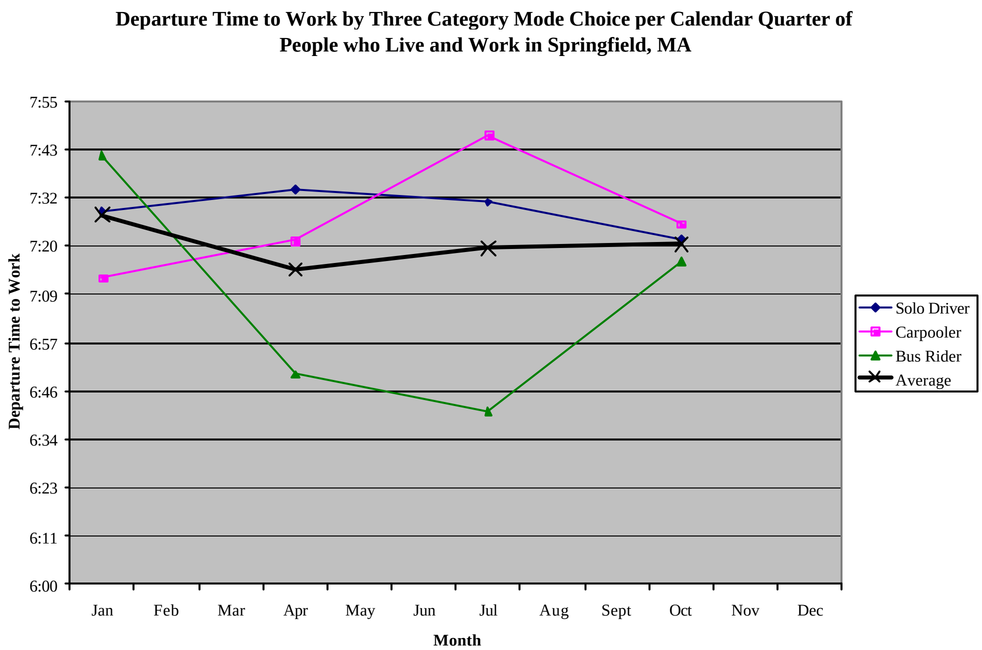

FIGURE 13a – Departure Time to Work by Three Category Mode Choice per Calendar Quarter of People that Live and Work in Springfield, MA

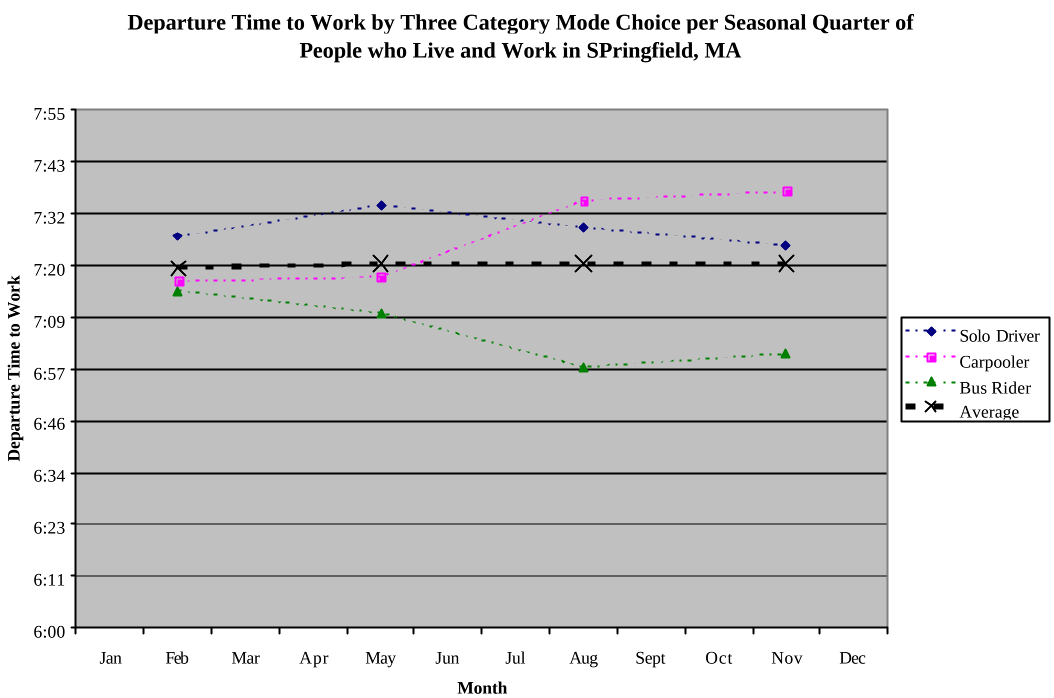

FIGURE 13b – Departure Time to Work by Three Category Mode Choice per Seasonal Quarter of People that Live and Work in Springfield, MA

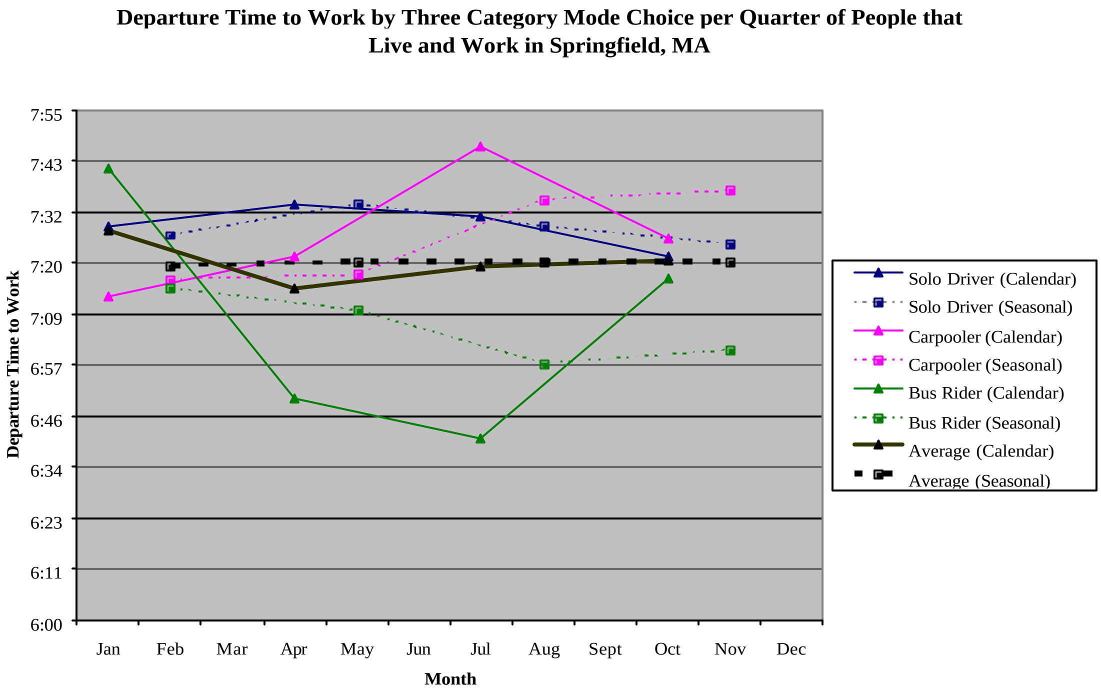

FIGURE 13c – Departure Time to Work by Three Category Mode Choice per Quarter of People that Live and Work in Springfield, MA

List of Tables

TABLE 1 Advantages and Disadvantages of Continuous Measurement for the Transportation Community

Advantages
➲ Agencies will not have to use twelve-year-old data in their planning activities (decennial census + at least two year data release time) ➲ Can argue that larger areas (with larger data and planning needs) will receive acceptable data every year or three years (still with two-year lag, possibly) ➲ Seasonal data: capture more bike and walk trips; capture different times leaving for work
Disadvantages

➲ DISCLOSURE—getting enough data for transportation uses
    ♦ Huge concern that individuals aren't identified from released data
    ♦ We may want mode by travel time by race by people who go from A to B but because some modes are infrequently used and some races are infrequent, this may lead to specific individuals
    ♦ In general, census will release *statewide* data with 75 observations per tabulated "cell"
    ♦ Census Transportation Planning Package (CTPP), closer to 30, some arrangements of five observations representing 30 people
    ♦ Race by mode by place of work…getting even five observations is difficult with 17% sample; now 3% sample
    ♦ Regression requires similar "cells"
        • Travel time = a + b1*income +b2*place of work requires (75) in each income/place combination
➲ Aggregation of five-year data for some areas, shorter periods for some sub areas—within the same metropolitan statistical area
➲ Error may propagate over five years
➲ One planning agency uses 2003-2007 data; another using 2005-2009 data may lead to different baselines
➲ Worries about continuous congressional funding
➲ Unexpected costs per completed interview may lead to longer data accumulation periods or only voluntary data. Currently the Census Bureau is "experimenting" with collecting only voluntary data (mail sample only, no follow-up phone or in-person interviews) (Quesinberry, 2003).

TABLE 2 U.S. Census Variables Used in Transportation Planning and Programming Processes (Table is modified from (Christopher, Murakami, Riklin, and Srinivasan, 2003))

Journey to Work and Mobility Questions	Demographic Variables
Place of Work	Sex
Travel Mode to Work	Age
Vehicle Occupancy	Race
Travel Time to Work	National origin
Time Left for Work (or Time Arrived)	Citizenship
Number of Vehicles in Household	Education
Disability status affecting employment	Building Type
Usual Hours Worked Per Week	Employment Status
Actual Hours Worked Last Week	Employer Industry
Vehicles Available	Employee Occupation
	Worker Earnings
	Household Income
	Household Size
	Household Type
	Geographic Mobility
	Language

Appendix A

RDC Timeline for Emily Parkany's Seasonality of Transportation Data Project (Winter '03-Fall '03)
October 8, 2003

December 14, 2002 In a visit to use data in the Boston RDC, the administrator explained I may be able to get access in the Maryland RDC and that there were advantages to being near Washington DC and closer to demographic/JTW and ACS census staff and the Disclosure Review Board.

January 2003 The Boston RDC administrator said that I would have to transfer the data from Boston to DC as I did not have a satisfactory reason for needing access to multiple centers.

Feb-March 2003 A series of emails were exchanged about transferring the project and how long I would have access to the data in Boston. It was decided that the desired third year of data would not be available in Boston (but that I would have permission to use it in Suitland).

April 10 2003 I was instructed to revise my RDC proposal (AGAIN) to describe why I needed the third year of data. Revisions were sent immediately.

April 29, 2003 I received a draft agreement from Arnie Reznek, MD RDC administrator, to review that needed the name of a Villanova signatory (Director of the Office of Sponsored Research was designated). This included a start date of June 4.

May 1, 2003 I received revised draft agreement including signatories and revised start date of May 16.

May 21, 2003 After emails inquiring about status, I learned that the agreement had to be revised by them with some "minor changes" which would change the start date back to June 4.

June 17, 2003 I received agreement signed by Dr. Ron Jarmin on June 5. I obtained the necessary Villanova signatures and returned on June 19, 2003.

July 3, 2003 I received account and data request form. (It was first sent to my University of Massachusetts account on June 30.) I printed out 20 pages and sent it overnight. (The form said that if it was not sent overnight it could take weeks to process.)

July 10, 2003 The overnight mail was misdirected as the USPS address was different from the FedEx address. The Center received it on this date and I was told that July 14 would be the only day I could start in the next two weeks.

July 14, 2003 I started. By the end of the day, the three years of ACS data was transferred into STATA. I agreed to come back Friday and get my badge processed at the Suitland location (2 ½ miles away from the RDC in Upper Marlboro, MD). (Badge processing was not done on Mondays.)

July 18, 2003 Returned for my badge. The process took about two hours because security observed that they had never processed my fingerprints and they were over one year old. So I "was lucky" that they could do my fingerprints that day. Additionally, initially they made a badge with only 6:30-6:30 M-F access. Another trip to the security office and some phone calls later, I was able to get a 24-hour access badge. The badge allows access to the census buildings and has a chip for access to the Washington Plaza building after hours, the RDC/CES office at all times, the WP parking garage, the canteen, etc. Access to the RDC rooms requires a numerical code.) It turns out that the delay was OK because Washington Plaza had an electrical outage and I came back to dark computers.

Mid-August Started conversation about Disclosure procedure. Discovered that two forms: Project Clearance Record and Request for Clearance were necessary.

August 29, 2003 Met with Arnie Reznek, MD RDC administrator, and Phil Salopek about how to prepare the data for disclosure. Filled out the forms and left the data for them to access.

September 23, 2003 Told that Arnie and Phil had looked at the data and that they had questions that they wanted to discuss via a conference call. They gave me two preferred times on Wednesday, September 24.

September 24, 2003 Conference call. Lasted 90 minutes. After some logistical delays, it was determined that I would go to the RDC and label the desired tables for the disclosure review board—both the desired tables and the unweighted supplemental tables that the board would need.

September 29, 2003 Visited the RDC but could not get past the login screens because my file storage was too large. Without my knowledge, I had been generating a "core" file that prevented my login.

October 5, 2003 I visited the RDC again and left edited files for Arnie and Phil to review for disclosure.

[The desired tables were released in November 2003 after review by the Disclosure Board and values representing two or fewer respondents were deleted.]