Understanding Changes in Youth Mobility

TECHNICAL APPENDICES TO THE FINAL DELIVERABLE

Prepared for
NCHRP 08-36

Task 132
Transportation Research Board
of
The National Academies

Table of Contents: The Four Technical Appendices

TECHNICAL APPENDIX 1: ADDITIONAL DOCUMENTATION ABOUT VARIATION ASSOCIATED WITH COHORT GROUPS.

Examining the Trends in VMT per Driver

Appendix 1 presents an analysis of the role of two factors in the explanation of change in VMT over the past decades. Those factors which reflect the age of a given group are expected to change as that group matures are called "time-effect" factors. Those factors which reflect the uniqueness of the behavior of a given group, and are expected to remain influencing behavior as that group matures -- are called "cohort effects." The content presented in this Technical Appendix was originally included in the Interim Report, but deleted from the final report because of its length and complexity.

Separating Out the Cohort-effect from the Time-effect in Variation in VMT

A key question for many observers of transportation (and other public behaviors) is whether the group we tend to call the "Millennial Generation" is characterized by preferences, values, and beliefs that will continue to influence their choices as they proceed through their life cycles and aging process. If so, this would be labeled a "cohort-effect." As they age, they will be subject to social and economic realities which will change over time; understanding these non-cohort based influences will help us interpret the "time-effect."

In order to provide an exploration of these two separate demographic phenomena (cohort-effect vs time-effect), we created a test case where the two patterns could be seen together. Remarkably enough, the U.S. Census Bureau does not have published definitions of the major generations referred to so often in the media, with the exception of the Baby Boomer Generation. In general, however, the generations defined in the public realm tend to be about 15 years in duration, except the Baby Boomer generation which is somewhat longer. For this analysis, we have focused on two National Household Travel Survey (NHTS) years which are, indeed roughly 15 years apart: the 2009 survey was chosen because it is the main source of information about the decline in VMT, and 1995 was roughly 15 years before that.

The Cohort-effect on VMT

Figure 1 and Figure 2 have been created to explore the two separate demographic patterns. In both diagrams, the reader is recommended to read *down* within any given column to reveal the logic of the effect being presented. Figure 1 is a diagram of the cohort-effect in which four generational groups are defined. For each column, a range of birth years are defined, which together comprise the cohort group, defined in groups of 15 years.

Figure 1 provides an example of the effect of a given cohort group ('generation') proceeding through the aging process, with their change in VMT noted in the lower two rows. In 1995, the youngest age category was populated (largely) by the Generation X'ers. Fourteen years later, in 2009, the youngest age category was populated (largely) by the Millennial Generation. The bottom rows report that when the Millennials populated this age category, they generated about 14% fewer vehicle miles per driver than did the Generation X'ers when they were in this age category.

When the Generation X'ers took the 2009 NHTS survey, they then populated the "Early Middle Age" category. From the bottom row of those columns, we see that for those between 31 and 45 years of age, the Generation X cohort had a driving rate about 3% lower than the Baby Boomers had reported in their 1995 survey, at the same age of their lives. While the next two generations are not

the primary focus of this NCRHP study, it should be noted that both Baby Boomers and the Silent Generation reported higher rates of driving than the previous generation when they were the same age.[1] The Silent Generation reported higher VMT than did the oldest generation when they were both between 61 and 75 years of age—an example of the cohort-effect.

The Time-effect on VMT

Applying the same definitions for the groups, Figure 2 is presented to show the effect of one cohort moving from one time in life to a later time in life, which is part of the time-effect. Looking at the Generation X-ers, *here* defined as those born between 1964 and 1979, we find that when this group left the life-phase of being under 31 years of age, and entered the life phase of being between 31 and 45, the older phase in life was associated with a driving rate 17% higher than the younger life phase.

When the Baby Boomers moved up one life-phase, (to being between 46 and 60 years of age) their driving rate went down by about 9%. When the Silent Generation took the 2009 NHTS survey, their driving rate was 39% lower than they had reported when they were 14 years younger. Although these older generations are not the primary focus of this NCHRP study of youth mobility, these patterns are shown here to illustrate the logic of the time-effect on VMT per driver. By contrast, the reader will note that there *are no* time effects reported here for the Millennials, as we only have data from them from one life-phase, that of being between 16 and 30 years of age.

While the two terms used in the diagrams of Figure 1 and Figure 2 might seem obvious, it is important to note that the *two concepts often get confused in the public media*. In many cases, an implicit assumption is made that if Millennials, for example, are consuming more transit services than (say) Generation X-ers, then it follows that Millennials will continue to manifest those same behaviors as they proceed into an older age category, or life phase. Figure 1 and Figure 2 are presented here to show as clearly as possible that there are two very different demographic patterns occurring at once. The concept of two patterns at play simultaneously is also revealed in **Error! Reference source not found.** (shown earlier) where the differences associated with progressive age categories are shown on the same chart as the differences associated with three separate cohorts being shown for each age category.

[1] The "Silent Generation" includes those born before the arrival of the Baby Boomers after the conclusion of World War II.

	Figure 1: Cohort-Effect on VMT per Driver for Each Age Category												
	Youngest **Age Category**			***Early Middle*** **Age Category**			***Late Middle*** **Age Category**			***Senior*** **Age Category**		
Age in both surveys	*16-20*	*21-25*	*26-30*	*31-35*	*36-40*	*41-45*	*46-50*	*51-55*	*56-60*	*61-65*	*66-70*	*71-75*
Cohort	*Born between 1965 and 1979*			*Born between 1950 and 1964*			*Born between 1935 and 1949*			*Born between 1920 and 1934*		
1995 VMT	17,190	28,363	31,194	30,851	31,138	30,318	30,928	26,868	24514	20,840	17,103	7,474
Cohort	*Born between 1979 and 1993*			*Born between 1964 and 1978*			*Born between 1963 and 1949*			*Born between 1934 and 1948*		
2009 VMT	13,827	24,891	28,237	30,083	29,681	30,109	30,317	28,316	25,249	22,443	19,934	7,964
2009 VMT as Percent of 1995 VMT	***80%***	***88%***	***91%***	***98%***	***95%***	***99%***	***98%***	***105%***	***103%***	***108%***	***117%***	***107%***
	Millennials had about 14% *less* VMT than **Gen X'ers** when they were in this *Youngest* age category			**Gen X'ers** had about 3% *less* VMT than **Boomers** when they were this *Early Middle* age category			**Boomers** had about 3% *more* VMT than **Silents** when they were in this *Late Middle* age category			**Silents** had about 11% *more* VMT than **Oldest** cohort when they were in this *Senior* age category		

Figure 1. Diagram showing the effect of the change of cohort group within each age category. Source: data from NHTS, ORNL website.

Figure 2: Time-Effect on VMT per Driver for Each Cohort Group												
	Cohort Born 1979 - 1994			**Cohort Born 1964-1979**			**Cohort Born 1949 - 1964**			**Cohort Born 1934 -1949**		
Age Category in 1995	(Too young for survey)			*Youngest* Age Category			*Early Middle* Age Category			*Late Middle* Age Category		
Age in 1995	*1 to 5*	*6 to 10*	*11 to 15*	*16-20*	*21-25*	*26-30*	*31-35*	*36-40*	*41-45*	*46-50*	*51-55*	*56-60*
1995 VMT	(Not collected)			17,190	28,363	31194	30,851	31,138	30,318	30,928	26,868	24,514
Age Category in 2009	*Youngest* Age Category			*Early Middle* Age Category			*Late Middle* Age Category			*Senior* Age Category		
Age in 2009	*16-20*	*21-25*	*26-30*	*31-35*	*36-40*	*41-45*	*46-50*	*51-55*	*56-60*	*61-65*	*66-70*	*71-75*
2009 VMT	13,827	24,891	28,237	30,083	29,681	30,109	30,317	28,316	25,249	22,443	19,934	7,964
Older Age Category VMT as Percent of Younger	n.a.	n.a.	n.a.	***175%***	***105%***	***97%***	***98%***	***91%***	***83%***	***73%***	***74%***	***32%***
Summary of Age Effect on Cohort	**Millennials** entered the *Youngest* age Category			**Gen X'er**s increased their VMT by about 17% as they aged into the *Early Middle* age category, 14 years later. The increase is strongest for younger age sub-categories.			**Boomers** decreased their VMT by about 9% as they aged into the *Late Middle* age category, 14 years later. Decrease strongest for older age sub-categories.			**Silents** decreased their VMT by about 39% as they aged into the *Senior* age category, 14 years later. Decrease strongest for older age sub-categories.		

Figure 2. Diagram showing the impacts of all other factors on the cohort over time. Source: data from NHTS, ORNL website.

TECHNIAL APPENDIX 2: ADDITIONAL DOCUMENTATION CONCERING DRIVER'S LICENSE RATES

Driver's License Rates and Decline in VMT

Appendix 2 now provides a more detailed analysis of driver licensing and VMT data for the youth market in the United States, as contained in the Interim Report, some of which was deleted from the final report for the sake of brevity and simplicity of presentation. [2]

Driver licensing data were obtained from the US Federal Highway Administration (FHWA), and population data were obtained from the US Census Bureau. The available data regarding the number of licenses held by members of particular age groups are disaggregated by year for persons aged 16-24. Beyond the age of 24, driver licensing data is only available in 5-year aggregations beginning with the 25-29 year-old group and continuing onwards. In addition, the analysis relies on travel data, in particular VMT provided by the 1995, 2001, and 2009 National Household Travel Surveys (NHTS).

Trends in Drivers Licensing among American Youth

The percentage of licensed drivers in the United States between the ages of 16 and 30 relative to the total population of each particular age group has decreased over the last two decades (Figure 1). Despite the decrease, a majority of young Americans are still obtaining their licenses, but many are obtaining them at later ages. For example, 43% of persons born in 1978 held licenses by the time they were 16 years old (1994) but only 29% of persons born in 1990 held licenses by the time they were 16 years old (2006). This represents a decrease of 43%. By the time the cohorts reached their 24th birthdays the gap was considerably smaller with 88% of those born in 1978 holding driver's licenses and 81% of those born in 1990 holding licenses.

About 85% of Americans between the ages of 25-30 were license holders in 2014. However, this number has decreased as well, from just over 90% between 1994 and 2004, suggesting that there are some cohort effects in decreasing rates of license holding but a focus on the very young (16 to 18) may lead to spurious conclusions.

[2] Figure numbers are from the Interim Report.

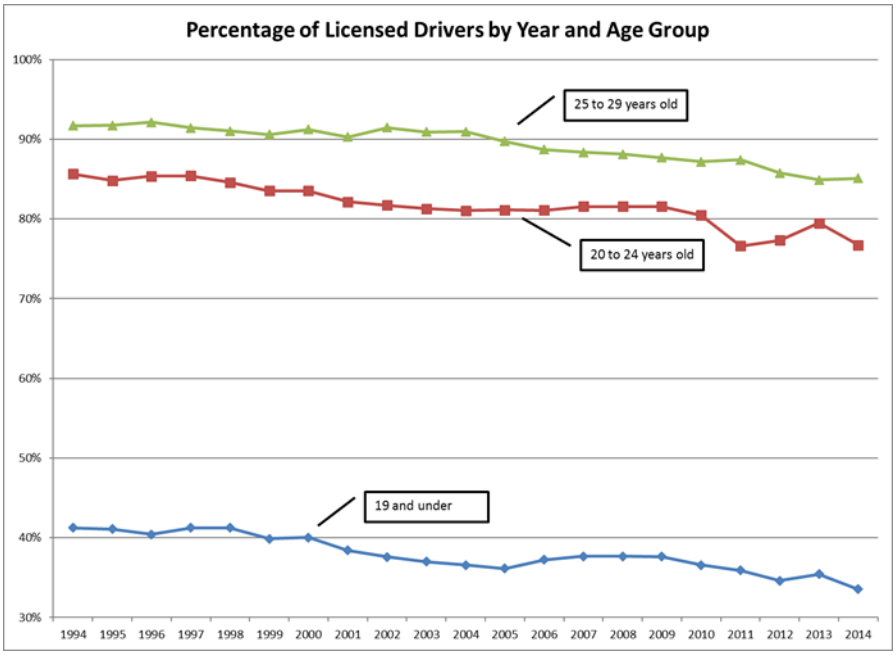

Figure 1. Percentage of licensed drivers in relation to population.

The decrease has been the most notable for those in the 16-18 year old age groups whose percentage of license carriers relative to their population appears to have fallen by about sixteen percentage points for both 16 and 17 year olds, and thirteen percentage points for 18 year olds, respectively (Figure 2). However, for each age group between 19 and 24, the decreases were also significant, with each individual age group seeing a percentage of licensed drivers in relation to population reduced by between about eight and ten percentage points over the same time period (Figure 3).

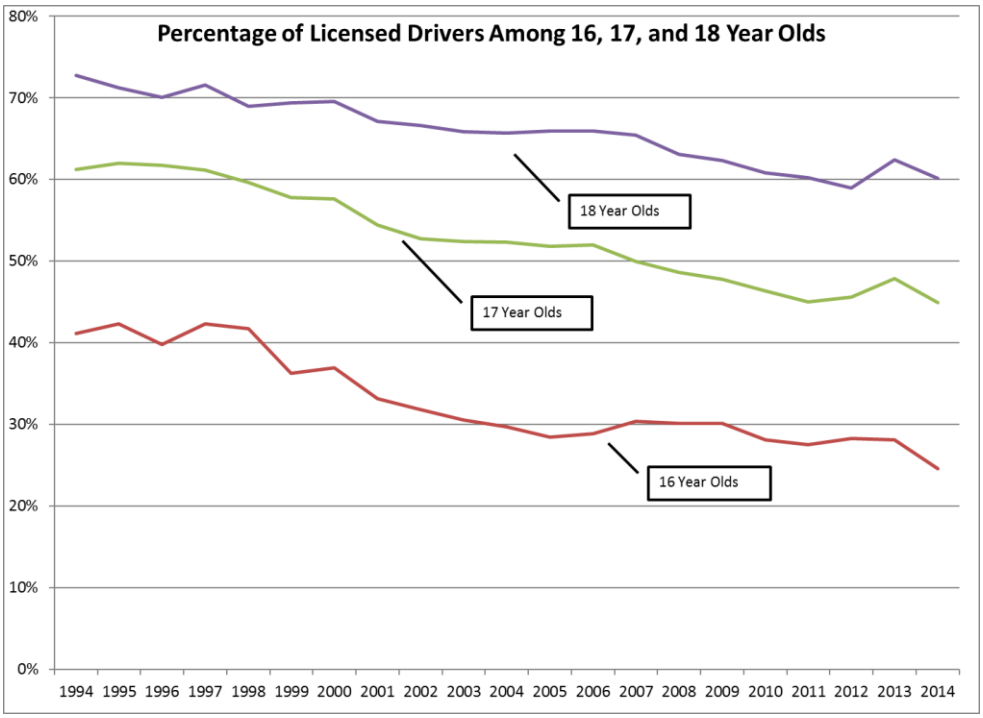

Figure 2. Percentage of license holding 16-18 year olds in the US

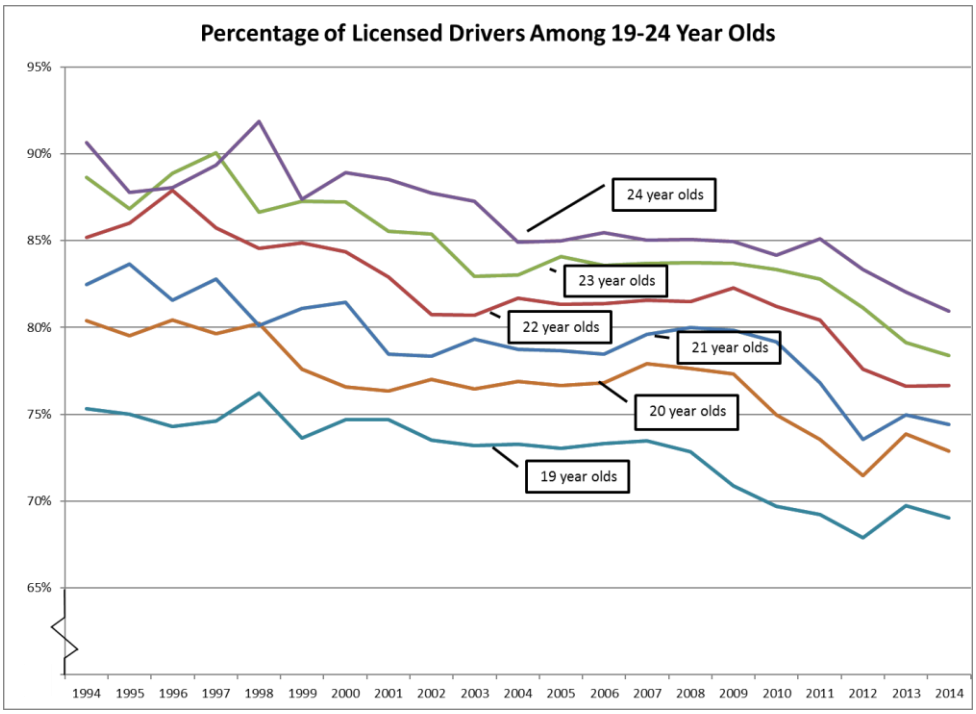

Figure 3. Percentage of license holding 19-24 year olds in the US

Overall, these reductions equated to a drop of about eight total percentage points of Americans under the age of 19, including 14 and 15 year olds, from just over 41% in 1994 to a third in 2014. For those aged 20-24 the drop proved even more significant, with about a drop of about nine percentage point, from about 86% to about 77% of the total population (Figure 1).

The drop in driver's license acquisitions among 16, 17, and 18 year olds appears to have had minimal effect, though, on a particular cohort's overall ratio of acquisitions over time, particularly once the cohorts are over the age of 19. Instead, it appears that the rate of acquisitions of most cohorts is simply delayed by about one year (Figure 4). As shown in Figure 2 and Figure 4, about 41% of 16 year olds had obtained a license in 1994 compared to about 29% of 16 year olds who had obtained a license in 2006. However, by the age of 19 both of these cohorts had licensing rates between 70% and 75%, and have remained within 8 percentage points, and frequently closer, of each other each year since (Figure 4). A similar pattern can be found when comparing most of the cohorts. By the age of 19 all cohorts had a license holding rate within about 5% of each other as their ratios increased (generally) year after year. This pattern continued for each cohort through at least the age of 24.

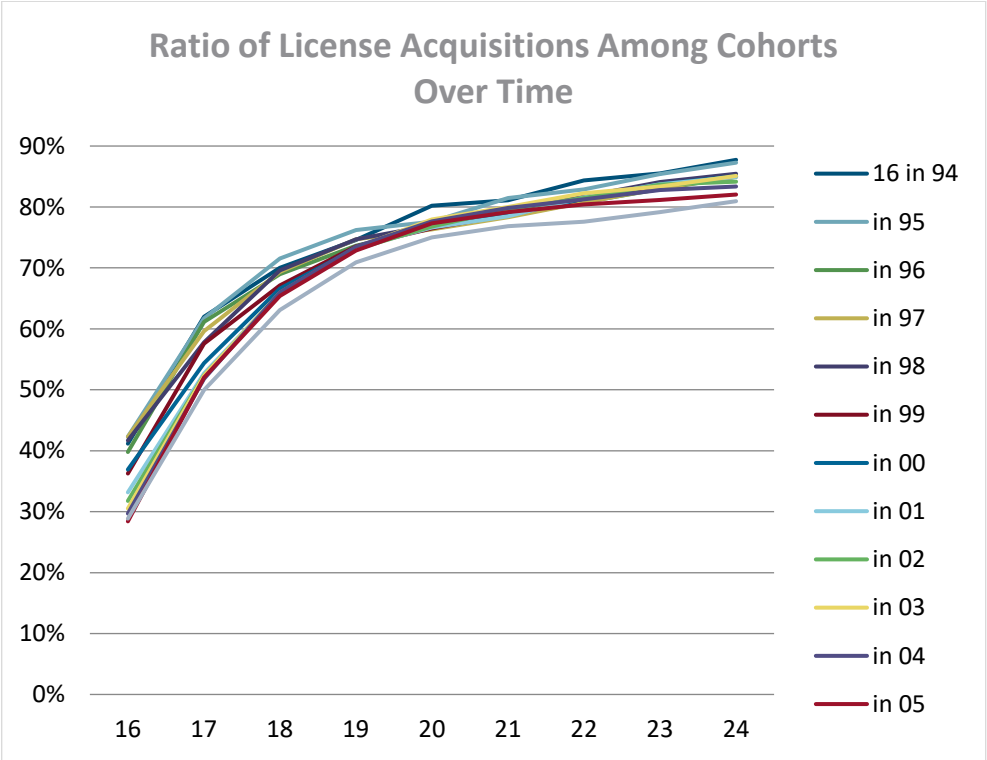

Figure 4. Ratio of license acquisitions among cohorts over time

While American youths do, for whatever reason, appear to be getting licenses somewhat later, *their reasons for doing so are not immediately clear*. There are a number of factors which may contribute to this phenomenon, including social factors, access to transit, or lifestyle choices. Economic factors are also possible contributors, as evidenced by the fact that all ages, and particularly those between the ages of 19 and 24, dropped *slightly faster between 2009 and 2012*, the worst period of the global recession. The most likely major contributing factor, though, is the change in rules and laws regarding the age young people are allowed to get a driver's license, and how they can use it once they do.

The Role of Graduated Licensing

The age of 16 had long been a standard average age that an American could expect to be able to get a license to drive. However, beginning in the 1990's concerns over safety and reasonable expectations of responsibility led many states to adopt stricter rules and procedures for minors

wishing to obtain a license. Now, most states operate under Graduated Driver Licensing systems, which require young drivers to advance through a two-stage learning permit process with built-in probationary periods before obtaining a standard driver's license. Therefore, where many young people used to be able to obtain a "full" or unrestricted driver's license at the age of 16, all but a handful of states now only allow young people to begin the second of the two-stage learning process at age 16 upon successful completion of the initial supervised learning stage. Only 15 states now offer a path for a 16-year-old to obtain a fully unrestricted license, compared to 23 states whose rules would not allow a person to obtain a license until either 17 or 18 years of age. The remaining 12 states offer licenses at varying ages, but with varying restrictions.[3]

Most states have also adopted a number of restrictions on drivers under the age of 18, even those who have successfully completed the learning stages and have obtained a standard license. Several of these restrictions could disincentivize driving, and therefore obtaining a license. Most notably, these restrictions include varying restrictions on the number, age, and/or familial status of allowable passengers, and restrictions on driving at certain parts of the day. As of 2015, at least 46 states limit the number of passengers that can be in a vehicle with a driver who is at an intermediate stage and/or who is below a certain age, most of which limit the number to a single passenger. Similarly, all but one state restricts when intermediate-stage drivers are allowed to drive at night, with most banning night driving altogether at that stage. Some of these states continue to impose night driving restrictions after an otherwise "full" license has been obtained, until the driver reaches a certain predetermined age.[4]

Implications of Delayed Licenses for VMT

The VMT implications of delayed driver license acquisitions among American youth are likely to be trivial, particularly among the youngest cohorts. The average VMT of 16-20 year olds over this time period has steadily reduced, but only retirement-age Americans drive less on average. Cohorts aged 21-25 and 26-30, however, have remained among the heaviest drivers on average *despite* having also reduced their driving over this time period. Also, as discussed previously, most cohorts have been shown to "catch up" to license rates of previous cohorts after only a year delay upon reaching the age of 19. Therefore, delayed licensing is only likely to produce meaningful effects to the VMT of 16-18 year olds. However, as this group is among the smallest in terms of licenses held, and overall VMT of any age group, the implications of delayed licensing to VMT are small.

Thus, our analysis of rates of licensure concludes that, yes, the exact year in which the license is obtained is being delayed compared to historical patterns, but, the groups seem to be returning to historical levels in their twenties.

Age of License Acquisition by Age Group

In the TransitCenter survey of 2014, Millennials and Gen X-ers both have *lower* than the average propensity to have obtained their driver's license while they were 16 years age, with the Baby Boomers and Silent Generations both reporting higher than average rates of obtaining it while 16 years old (Table 1).

---

[3] Governor's Highway Safety Association (2016). Graduated Driver Licensing (GDL) Laws. URL: http://www.ghsa.org/html/stateinfo/laws/license_laws.html

[4] Governor's Highway Safety Association (2016). Graduated Driver Licensing (GDL) Laws. URL: http://www.ghsa.org/html/stateinfo/laws/license_laws.html

Table 1. Age of obtaining first driver's license. Source: TransitCenter, 2014.

Age obtained driver's license	**18-24**	**25-34**	**35-49**	**50-64**	**65+**
Age 16 or before	38%	48%	52%	60%	59%
Age 17 to 18	45%	31%	30%	27%	26%
Age 19 to 24	16%	13%	13%	10%	11%
Age 25 to 35	0%	7%	4%	2%	3%
After age 35	0%	0%	2%	1%	1%

The propensity to have a driver's license did not change substantially between 2001 and 2009, as self-reported by the participants in the two NHTS surveys (Table 2). For males, those under 35 had about a 3% lower rate of license holding in 2009 than they did in 2001. Males above 35 years of age dropped by about 1% over the same period. For all age groups together, males went down by 1% and females went up by 1%, confirming that this is not a major explanatory factor in change in VMT between 2001 and 2009, at least as self-reported in the two surveys.

Table 2. Change in self-reported driver status in NHTS, 2009 as a percent of 2001. Source: data from NHTS, ORNL website.

Change in self-reported driver status	**18-24**	**25-34**	**35-49**	**50-64**	**65+**
Male (*year 2009 rate as percent of year 2001 rate)*	97%	97%	99%	99%	99%
Female (*year 2009 rate as percent of year 2001 rate)*	104%	97%	99%	100%	101%

Technical Appendix 3. Additional Analysis in New Model Building for Change in VMT

Technical Appendix 3 now presents a fuller and more technically complete description of model building activity undertaken by the Research Team and reported in Chapter 4. While the Final Report includes our most up do date description of the analysis, some content was shortened from the content of the Interim Report, for the sake of brevity. This Technical Appendix includes the original version of that text, as included in the Interim Report which is somewhat longer, and more detailed than included in the Final Report. In addition, this Appendix presents new description of the models, and key technical back-up information about them.

The following section describes the further analyses that were carried out to extend the approach used by McDonald (2015), in terms of both methodology and results.

A Further Exploration of NHTS Analyses Indicating a Relative Decline in Auto Usage among Millennials

Extensions of the research approach: description of methods and results

The research team carried out a variety of analyses, starting from the approach used by McDonald. These included combinations of including new explanatory variables, including additional interactions between variables, changing the dependent variable used in the regression, and changing the function form of the regression model itself. In this section we give a qualitative description of the analysis methods and results, without providing the quantitative model estimation results in detail. (A more detailed paper with full model estimation results will be prepared and available upon request.)

Replicating the original analysis

Our first analysis was to use the description of the survey data processing and model specification provided in the McDonald paper and attempt to replicate the original model and results. This test was successful, as we were able to produce sample statistics and regression model estimation results that were very close to those reported in the paper, with only very minor differences. This test provided strong evidence that there were no major data processing or analysis errors made in either McDonald's analysis or our own, and that our initial models were a reliable basis for further exploration of the data.

Adding additional explanatory variables

One hypothesis is that there may be additional variables in the NHTS data that could help explain why the younger age groups in particular showed the largest drop-off in auto use in 2009 compared to the earlier years. Below is a list of additional explanatory variables that were tested, and the model estimation results:

- **Parents with children in different age categories**: One societal change that has happened over the past 20 years is that many young adults are now having their first child at a later age, so the percent of people in the different age groups with children in different age categories will be shifting. The model results showed that parents with children in the K-12 non-driving age range (5 to 16) travel significantly more auto miles per day, presumably to

take their children to school and other activities. Similar but smaller effects were found for parents of children age 0-4 or parents of children at home age 17-21. Interestingly, when these variables were interacted with the gender of the parent, they were found to have a stronger positive effect for males compared to females, going against the stereotype that females do most of the chauffeuring of children.

- **People age 19+ living with their parents**: "Children" in the analysis still living at home with parents show significantly less auto use than others, although the magnitude is only about 1 mile less per day (about 3% of the average daily distance). When this variable was interacted with gender it was found that only the male children living at home tend to have less auto miles, while the effect is in the opposite direction for females.
- **Saturdays versus Sundays**: While the original analysis only used a single variable for weekend travel, the analysis actually showed about 3 miles *more* auto use on Saturdays and 3 miles *less* auto use on Sundays as compared to weekdays, all else equal.
- **Households with no cars**: Even though the original model included a variable for autos per licensed driver in the household, there is an additional significant negative effect of owning 0 vehicles, above and beyond the effect of vehicles per license. (In other words, the effect of auto ownership is non-linear, with a higher negative effect at 0.)
- **More detailed land use categorization**: While the original model included an "urban area" variable, the NHTS data also includes the "Claritas" land use categorization into four different area types based on Census block group level data. This categorization provides significant estimates for the added variables, even when added along with the simpler two-way classification. The population density variables still remain significant as well. The more urban the area type, and the higher the population density, the lower the daily auto mileage, as one would expect.
- **People looking for work**: One employment category that was not included in the original models is those that are listed as "looking for work." Those who are unemployed and looking for work travel significantly fewer auto miles than others, all else equal.
- **Education level and student status**: Compared to those with no college education, those with at least some college education and/or a college degree travel more auto miles, even after accounting for other related differences such as income level and employment status. Those who were currently college students also traveled more auto miles, all else equal.
- **Non-linear effects of income**: While the original model included a single linear income variable, variables for income over $100,000 (in 2009 dollars), and for income below the poverty level were added to test for non-linear effects. None of the added variables was significant, indicating that a linear relationship with income is appropriate in this model.

Our main reason for testing these additional variables was to see if they would account somewhat for the lower auto miles traveled found for the youngest age groups in 2009. If that were the case, then we would expect the interactions between age and survey year (the purple bars in **Error! Reference source not found.**) to become smaller and less statistically significant. That was not the case, however. While adding all of the additional variables listed above improved the statistical fit and explanatory power of the models, the interaction effects between the age groups and survey years remained at a similar magnitude and significance. At first pass, this finding suggests that the reason for the Millennials' reduced auto use in 2009 (relative to other age groups and time periods) was at least partly due to variables that were not measured in the travel survey data, such as differences in attitudes towards travel and/or differences in lifestyle, possibly related to the use of mobile technology and social media. However, we hypothesized other possible statistical reasons for the estimation results that were due to the way that the regression model was specified. The tests of these further hypotheses are reported below.

Adding additional interactions between explanatory variables

One possible drawback of the way that the original regression model was specified is that only the age group variable was interacted with the survey year. Thus, the implicit assumption is that all of the other demographic variables, such as employment status, income, etc. had the same marginal effects across the different survey years. If, for example, being unemployed had a more significant effect in reducing auto usage in 2009 than it did in 2001, and if more people in the youngest age categories were unemployed in 2009 relative to the older age groups, then that could lead to estimation of a (at least partially) spurious relationship between age and survey year. We tested this hypothesis by including a number of different interactions between various key explanatory variables and the survey year:

- **Gender interacted with survey year**: Compared to females, males' auto distance declined much more in 2009 relative to the earlier survey years, even after accounting for differences in employment status, income and other variables.
- **Employment status interacted with survey year**: Those who were not working had significantly less auto miles traveled in 2009 as compared to those not working in the earlier survey years. This did not apply to those who were "looking for work," however, as that effect does not appear any stronger in 2009 than in the other years.
- **Education level and student status interacted with survey year:** The positive relationship between having at least some college education and traveling more auto miles was significantly stronger in 2009 than in the other survey years, particular among those who hold college degrees and those who were currently students at the time of the survey.
- **Income below poverty level, Hispanic ethnicity, and child age 19+ living with parents, all interacted with survey year**: These three variables were also tested for different effects across the survey years, but none showed significant interactions.
- **Households with no cars interacted with age group**: While not owning a car is strongly related to less auto miles traveled, this negative relationship is not as strong for the youngest (19-24) age group. It may be the case that people in this age group more often live in situations in which they can borrow a vehicle from family or friends/housemates when needed.

We also tested three-way interactions between age group, survey year and other key variables, essentially replicating the year x age interaction variables shown in **Error! Reference source not found.** for different demographic groups:

- **Gender interacted with age group and survey year**: Although males' auto distance was found to have declined more than females' in 2009 compared to the earlier survey years (see above), there was no additional significant interaction found across the age groups. In other words, this difference between males and females in 2009 appears to apply to all of the age groups tested, and not just Millennials.
- **Employment status interacted with age group and survey year:** Similar to the finding for gender, while not having a job was shown to have a stronger relationship to lower auto miles in 2009 than in the earlier years, this effect was not found to be significantly stronger for the youngest ("Millennial") age group than it was for the other age groups.
- **Households with no cars interacted with age group and survey year**: It was stated above that the effect of not owning a car was generally found to be less of a negative influence on auto miles traveled for the 19-24 age group compared to other age groups. However, when also interacted with the survey year, the opposite effect is found, with zero-car ownership showing more of a negative relationship with car use in 2009 than in the earlier years.

In summary, the tests of interaction effects indicate that lower auto use in 2009 compared to the earlier years is greatest among males, non-employed, and those without a college education, all else equal, but none of those effects appear to be much different for the younger (Millennial) age group than for the other age groups. In fact, even when all of these additional interaction effects were added to the model specification, the original interaction effects between age group and survey year remain very close in size and significance to those shown in **Error! Reference source not found.**.

Testing other dependent variables

"Auto usage" from a travel survey can be defined as person-miles traveled (PMT) or vehicle-miles traveled (VMT). PMT includes miles for all trips by auto, whether as a driver or a passenger, while VMT generally includes only miles for car driver trips. (Alternatively, one could include all trips and divide the miles traveled by auto occupancy, but it is questionable as to whether VMT should be allocated to child passengers for any behavioral analysis. A third option would be to include all trips and divide by the number of *adult* auto occupants to obtain an estimate of VMT, but that is not a straightforward calculation using NHTS data that does not always identify all passengers.)

Although it is not explicitly stated in the McDonald paper, it appears that PMT was used as the dependent variable, as it also was in our comparative analyses. If, however, car occupancy rates shifted across the different age groups over time, then using VMT as the dependent variable may give different results. This was tested empirically by using the same NHTS data observations and independent variables, but using only auto driver trips rather than all auto trips to calculate one-day VMT as opposed to PMT. The VMT estimation results are quite similar to the analogous model regressed on PMT, with all of the same variables being significant with the same signs and comparable magnitudes. However, the interaction effects between age group and survey year, while remaining statistically significant, decrease in magnitude by about 50% compared to those shown in **Error! Reference source not found.**. This finding indicates that the "unexplained" drop in auto miles for Millennials in 2009 (as compared to the same age group in earlier survey years) is less pronounced when one only considers car driver trips rather than all auto trips. Another way of interpreting this finding is that the "unexplained" drop in car miles traveled was greater among Millennial car passengers than among Millennial car drivers.

Testing alternative model specifications

A linear ordinary least-squares (OLS) regression model assumes that the unexplained variation (residual error) in the model has an approximately normal, bell-shaped distribution. This in turn implies that the dependent variable itself should have a distribution that is more or less bell-shaped, or at least not strongly skewed or truncated in one direction or the other. Auto miles traveled per day in the NHTS data, however, has a distribution resembling a log-normal distribution, with a large proportion of cases at 0, and a distribution skewed to the left, with a long tail towards the right. This is true when using either PMT or VMT, with PMT during the travel day equal to 0 for about 13% of the NHTS travel days in the estimation sample, and VMT equal to 0 for about 18% of the travel days. Since this distribution violates the assumptions underlying OLS regression, it raises the question of whether the data would yield different results if using a more appropriate model specification. To address this question, we tested three alternative functional forms:

1. **Using a logarithmic transformation of the dependent variable**: Instead of using PMT or VMT directly as the dependent variable, this model form uses Log(PMT + 1) or Log(VMT + 1). Adding 1 before taking the logarithm is necessary to keep the cases of 0 in the model (since Log(0) is undefined). This transformation does provide more of a bell-shaped normal

distribution for the non-zero PMT or VMT values, and as a result the model fit of the regression models improves considerably. In most cases, the variables that are significant with the linear specification become even more significant with this log-linear specification, including the key variables in the analysis—the interactions between the age group effects and survey year effects. However, the fact that the dependent variable remains truncated at zero for almost 20% of the observations remains an issue.

2. **Using a TOBIT model form**: The TOBIT model is a generalization of linear regression to deal with truncated dependent variables. The estimation method essentially extends the regression line past the truncation value. In this case, it would assume that when 0 vehicle-miles are observed, the "true" latent value would be a negative number of miles traveled in the day. Behaviorally, one could also interpret this as a desire to avoid traveling by auto. Empirically, the difference in estimation results between the OLS regression and the TOBIT regression are fairly minor, both when using the linear form of the dependent variable and when using the logarithmic transformation. The most pronounced differences are for the variables that are most strongly related to observations of 0 auto miles, such as not owning a car and not having a driving license. For the key variables in our analysis, the interactions between age effects and year effects, the estimates become somewhat larger and more statistically significant.

3. **Using a two-stage model**: This approach deals with the truncation issue in way that is less statistically sophisticated than the TOBIT model, but is stronger in terms of explaining behavior in an interpretable manner. The approach is to estimate two different models: (a) a binary logit model to explain which people do not use auto at all during the travel day (0 auto miles versus a positive number of auto miles), and (b) a log-linear regression model which only includes the cases with one or more auto miles. When estimating the models in this form, the same set of explanatory variables generally gave significant estimates in both stages of the model. The interaction effects between age group and survey year retained their negative, significant values in both stages.

Description of models

In all models, the key "Millennial interaction terms" are highlighted in yellow – with the "age 19-24 in 2009" variable most representatives of Millennials. Significantly negative estimates for these interactions means that daily car use decreased more for these age groups than for the base (37-42) age group in 2009 relative to previous years.

Tests of dependent variable specification and model functional form

Model A1: The most similar to the model in the McDonald paper (linear ordinary least-squares (OLS) regression on auto miles--driver or passenger) , but includes a few additional explanatory variables and re-specification of the base category for certain variables, such as race. The results are similar to McDonald's for all key variables.

Model A2: The same explanatory variables and sample as A1, but with the explanatory variable changed to the log of (auto miles+1)—log-linear OLS regression. The model fit improves substantially, and the statistical significance of the interaction terms is higher. The implication is that the distribution of the log of auto miles more closely follows a normal distribution in the sample. (The coefficient values are smaller than in A1 because of the logarithmic transformation.)

Model A3: The same as model A2, but excluding cases where auto miles is 0 (the respondent made no auto trips in the day. The model fit and t-statistics are lower than model A2, but similar to Model A1.

Model A4: The selection model corresponding to A3—a binary discrete choice model for 0 auto miles versus positive auto miles. The interaction coefficients are negative and statistically significant. Taken together with A3, this means the "Millennial effect" on auto use applies for both the propensity to make any auto trips in the day and for the distance traveled by those who did make auto trips.

(**Note**: Another functional form that was tested, but results not shown here, was a "Tobit" model, which is a form of regression model designed for censored data—in this case the fact that the data is truncated at 0 because it is not possible to travel less than 0 miles. The Tobit estimation results with linear and logarithmic dependent variables are not substantially different from Models A1 and A2.)

Models B1 to B4 are the same as Models A1 to A4**,** but using auto driver miles as the dependent variable rather than total auto miles. (Only auto driver trips are counted.)

Model B1: The overall model fit is somewhat higher than for A1, but the interaction terms become smaller and less significant. Only the "age 19-24 in 2009" interaction remains significant, but with a smaller magnitude (-2.55 miles per day vs. -4.11 miles per day in Model A1).

Model B2: The overall model fit is higher than for A2, and, unlike B1, all of the the age-year interactions for 2009 are significant.

Model B3: When selecting only those who made driver trips, the overall model fit as about the as for A3, and the "age 19-24 in 2009" interaction is only marginally significant.

Model B4: For the binary logit model, the model fit is higher than for A4, but that is due to other variables (particularly the license holding and car ownership variables), as the age-year interaction effects are less significant than in A4.

Models C2 to C4 are the same as models A2 to A4, but going in the other direction to be more inclusive, using miles traveled by all modes as the dependent variable.

Model C2 has a somewhat lower fit than A2, and the age-year interaction effects are somewhat smaller than in A2, but still quite significant. This result indicates that the drop in auto distance is not compensated by an increase in distance by other modes, as overall travel distance also goes down.

Model C3 has a better overall fit than A3, and the age-year interaction terms remain about the same.

Conversely**, Model C4** has a worse overall fit than A4 and the age-year interaction terms are less significant. When expanding the dependent variable to include all modes, more of the explanatory power goes to the distance traveled by those who make any trips (Model C3).

Models A2M and A2F are the same as Model A2 (which is repeated in the table alongside them), but with separate models estimated for Male and Female respondents. The results for the three models are strikingly similar, particularly for the highlighted interaction terms. This indicates that the "Millennial effect" in 2009 does not appear to be gender-specific.

Tests of adding new variables and interaction terms

Models D1 and D2 are the same as Model A2 (which is again repeated in the table alongside them), but model D1 adds a number of new variables:

- New interaction variables for males, Hispanics, those living with parents, those not employed and those not owning cars, for years and/or year/age group combinations.
- New variables related to those looking for work, student occupation, education level, and poverty status, along with some interactions with survey year.
- The effects for child-rearing and children living at home with parents are split by gender.

Compared to model A2, the model fit in D1 is improved only slightly (r-squared of 0.179 vs. 0.174), even after the addition of several variables, some of which are significant. A number of the significant new variables are highlighted. Auto use in 2009 dropped most for males, for those not employed, and for those age 19-24 with no car in the household. Even with the addition of these effects, the original age-year interaction terms for 2009 remain significant in D1, with a magnitude very similar to model A1.

Model D2 is the same as model D1, but uses weighting on the data (all of the preceding models used no weights on the cases in estimation). The weight is the expansion weight for each observation divided by the average expansion weight in the sample, so the sum of weights remains the same. Using weighting increases the model fit somewhat, and the significance of some of the variables. With weighted data, the extra interaction effects by age group and year for males are now negative and marginally significant, suggesting that the drop in auto use by younger groups in 2009 was somewhat stronger for males. However, the main age-year interaction effects also remain significant with a similar magnitude to D1 and A2.

In summary, none of the changes in model specification that were attempted "explained away" the significant effect that auto travel in 2009 relative to previous years decreased more for the youngest age groups—particularly the 19-24 age group that was comprised of "Millennials"--than for older age groups.

Model	**A1**		**A2**		**A3**		**A4**	
Model form	**OLS regression**		**OLS regression**		**OLS regression**		**Binary logit**	
Dependent variable	**auto miles**		**log(auto miles+1)**		**log(auto miles+1)**		**IF(auto miles > 0)**	
Sample selection	**all cases**		**all cases**		**exclude auto miles=0**		**all cases**	
Adjusted r-squared	**0.070**		**0.174**		**0.061**		**0.228 (rho-squared)**	
Variable	**Coefficient**	**T-value**	**Coefficient**	**T-value**	**Coefficient**	**T-value**	**Coefficient**	**T-value**
Intercept	10.765	9.8	.834	21.2	-.788	-19.1	-1.392	-18.0
age 19 to 24	3.489	5.0	.177	7.2	.073	2.8	.329	5.6
age 25 to 30	1.873	3.1	.072	3.3	-.001	.0	.056	1.1
age 31 to 36	1.043	1.9	.042	2.1	.001	.0	.063	1.3
year 2001	-1.858	-3.7	-.051	-2.8	-.033	-1.8	.007	.2
year 2009	-3.495	-7.2	-.094	-5.4	-.020	-1.1	-.191	-4.5
age 19-24 in 2001	-1.447	-1.7	-.099	-3.2	-.092	-2.8	-.284	-3.9
age 25-30 in 2001	-.729	-.9	-.045	-1.6	-.045	-1.5	-.093	-1.4
age 31-36 in 2001	-.869	-1.2	-.055	-2.1	-.040	-1.5	-.143	-2.2
age 19-24 in 2009	-4.107	-4.9	-.218	-7.3	-.135	-4.3	-.370	-5.4
age 25-30 in 2009	-2.187	-2.8	-.140	-5.1	-.081	-2.8	-.245	-3.9
age 31-36 in 2009	-1.989	-2.8	-.106	-4.2	-.070	-2.7	-.214	-3.5
male	-1.954	-2.1	-.114	-3.4	-.107	-3.1	-.148	-2.6
driver's license	12.231	18.4	1.028	43.4	1.732	69.6	1.464	36.4
male w/ license	5.476	5.7	.167	4.9	.414	11.6	.119	2.0
Hispanic	1.748	4.3	.078	5.3	.048	3.1	.158	4.9
race=Black	1.283	2.6	.022	1.3	.081	4.4	-.041	-1.1
race=Asian	-2.931	-4.8	-.029	-1.3	-.050	-2.2	.036	.8
race=other non-white	.948	1.4	.035	1.5	.030	1.2	.027	.5

full time worker	7.324	23.7	.514	46.7	.600	51.9	.943	42.1
part time worker	4.336	11.2	.376	27.1	.447	30.7	.741	25.3
one person household	-.703	-1.3	.009	.5	.164	8.1	.116	2.6
parent w/child age 0-4	.149	.5	.076	7.0	.128	11.2	.243	9.8
parent w/child age 5-16	2.001	5.7	.146	11.6	.206	15.7	.382	12.8
parent w/child age 17-21	.811	1.1	.075	2.8	.129	4.7	.187	3.1
live at home w/parents	-.988	-2.5	.046	3.3	.120	8.1	.189	6.2
day=Saturday	3.454	10.4	.039	3.3	-.326	-26.0	.039	1.4
day=Sunday	-2.265	-7.0	-.211	-18.3	-.612	-50.5	-.296	-11.8
income data missing	1.978	3.8	.072	3.9	.085	4.3	-.018	-.4
income is over $100k	.283	.7	-.024	-1.7	-.031	-2.1	-.061	-1.8
income (in 2009$)	.077	16.5	.003	20.6	.003	16.5	.004	10.1
household has no cars	-3.253	-3.3	-.483	-13.7	.282	7.6	-.780	-12.0
cars/license in household	9.754	12.7	.706	25.8	1.209	42.1	1.214	22.6
population density	-10.299	-18.0	-.235	-11.5	-.210	-9.8	.120	2.6
population density squared	2.132	10.8	-.007	-1.0	.001	.1	-.177	-12.0
"urban" block group	-5.484	-10.8	-.153	-8.5	-.128	-6.8	-.046	-1.2
"second city" block group	-4.984	-14.4	-.139	-11.2	-.105	-8.1	.096	3.3
"suburban" block group	-4.341	-14.7	-.085	-8.0	-.057	-5.2	.076	3.0
data given by proxy	-2.938	-10.8	-.164	-16.8	-.241	-23.5	-.342	-15.8

Model	B1		B2		B3		B4	
Model form	OLS regression		OLS regression		OLS regression		Binary logit	
Dependent variable	driver miles		log(driver miles+1)		log(driver miles+1)		IF(driver miles > 0)	
Sample selection	all cases		all cases		exclude driver miles=0		all cases	
Adjusted r-squared	0.100		0.242		0.059		0.356 (rho-squared)	
Variable	**Coefficient**	**T-value**	**Coefficient**	**T-value**	**Coefficient**	**T-value**	**Coefficient**	**T-value**
Intercept	-6.824	-6.9	-.788	-19.1	2.517	16.0	-6.516	-36.8
age 19 to 24	1.383	2.2	.073	2.8	.038	1.9	.120	2.3
age 25 to 30	.112	.2	-.001	.0	.031	1.8	-.046	-1.0
age 31 to 36	-.007	.0	.001	.0	.008	.5	-.003	-.1
year 2001	-1.214	-2.7	-.033	-1.8	-.056	-4.0	.006	.2
year 2009	-1.382	-3.1	-.020	-1.1	.003	.2	-.076	-2.0
age 19-24 in 2001	-1.141	-1.5	-.092	-2.8	.011	.4	-.243	-3.8
age 25-30 in 2001	-.227	-.3	-.045	-1.5	.005	.2	-.115	-1.9
age 31-36 in 2001	-.378	-.6	-.040	-1.5	-.004	-.2	-.099	-1.8
age 19-24 in 2009	-2.547	-3.4	-.135	-4.3	-.060	-2.4	-.200	-3.2
age 25-30 in 2009	-.745	-1.1	-.081	-2.8	-.026	-1.2	-.130	-2.3
age 31-36 in 2009	-1.090	-1.7	-.070	-2.7	-.029	-1.5	-.113	-2.1
Male	-1.466	-1.8	-.107	-3.1	-.354	-1.6	.283	1.2
driver's license	16.328	27.3	1.732	69.6	.293	1.9	5.409	32.5
male w/ license	9.421	11.0	.414	11.6	.518	2.4	.121	.5
Hispanic	1.592	4.3	.048	3.1	.030	2.5	.049	1.6
race=Black	2.505	5.7	.081	4.4	.086	5.7	.034	.9
race=Asian	-2.422	-4.4	-.050	-2.2	-.041	-2.3	-.036	-.8
race=other non-white	1.437	2.4	.030	1.2	.076	3.8	-.027	-.5
full time worker	8.351	30.0	.600	51.9	.213	22.1	.917	44.7
part time worker	5.150	14.7	.447	30.7	.117	9.9	.744	27.9
one person household	1.780	3.7	.164	8.1	.022	1.4	.383	8.8

parent w/child age 0-4	1.115	4.1	.128	11.2	-.003	-.3	.282	12.9
parent w/child age 5-16	2.708	8.6	.206	15.7	.025	2.5	.425	16.3
parent w/child age 17-21	1.239	1.9	.129	4.7	-.001	-.1	.275	5.1
live at home w/parents	.526	1.5	.120	8.1	.012	1.0	.264	9.2
day=Saturday	-3.416	-11.4	-.326	-26.0	-.101	-10.4	-.568	-24.7
day=Sunday	-8.303	-28.5	-.612	-50.5	-.246	-25.0	-.924	-43.6
income data missing	1.922	4.1	.085	4.3	.099	6.3	.018	.5
income is over $100k	.105	.3	-.031	-2.1	.000	.0	-.080	-2.7
income (in 2009$)	.062	14.8	.003	16.5	.002	17.4	.002	7.0
household has no cars	5.248	5.9	.282	7.6	.125	2.4	-.668	-9.1
cars/license in household	15.036	21.8	1.209	42.1	.291	12.1	1.905	39.7
population density	-8.293	-16.0	-.210	-9.8	-.377	-22.0	.171	4.1
population density squared	1.763	9.9	.001	.1	.102	15.5	-.167	-11.7
"urban" block group	-4.617	-10.1	-.128	-6.8	-.145	-9.8	-.014	-.4
"second city" block group	-3.821	-12.2	-.105	-8.1	-.177	-18.0	.098	3.8
"suburban" block group	-3.384	-12.7	-.057	-5.2	-.113	-13.7	.082	3.7
data given by proxy	-3.642	-14.8	-.241	-23.5	-.058	-7.2	-.453	-23.2

Model	C2		C3		C4	
Model form	OLS regression		OLS regression		Binary logit	
Dependent variable	log(all miles + 1)		log(all miles + 1)		IF(all miles > 0)	
Sample selection	all cases		exclude all miles=0		all cases	
Adjusted r-squared	0.106		0.094		0.098 (rho-squared)	
Variable	**Coefficient**	**T-value**	**Coefficient**	**T-value**	**Coefficient**	**T-value**
Intercept	1.373	36.2	2.430	77.1	-.164	-1.9
age 19 to 24	.147	6.2	.076	4.0	.274	4.3
age 25 to 30	.076	3.6	.074	4.5	.036	.6
age 31 to 36	.035	1.8	.038	2.5	.005	.1
year 2001	-.034	-2.0	-.070	-5.1	.123	2.5
year 2009	-.046	-2.7	-.059	-4.5	.031	.7
age 19-24 in 2001	-.076	-2.6	-.035	-1.5	-.193	-2.4
age 25-30 in 2001	-.051	-1.9	-.042	-2.0	-.064	-.9
age 31-36 in 2001	-.048	-1.9	-.036	-1.8	-.070	-1.0
age 19-24 in 2009	-.169	-5.9	-.110	-4.8	-.244	-3.2
age 25-30 in 2009	-.127	-4.8	-.090	-4.3	-.157	-2.2
age 31-36 in 2009	-.090	-3.7	-.063	-3.3	-.121	-1.8
male	.077	2.4	.009	.3	.168	2.9
driver's license	.747	32.8	.457	22.9	.890	20.5
male w/ license	.001	.0	.058	2.0	-.094	-1.5
Hispanic	.050	3.5	.027	2.4	.085	2.3
race=black	.064	3.8	.071	5.3	.011	.3
race=asian	-.042	-2.0	-.033	-2.0	-.041	-.8
race=other non-white	.036	1.5	.056	3.0	-.066	-1.1
full time worker	.529	50.0	.210	24.2	1.049	42.4
part time worker	.409	30.7	.109	10.1	.923	27.7
one person household	.022	1.2	-.028	-1.9	.213	3.9

parent w/child age 0-4	.037	3.5	.002	.2	.113	4.0
parent w/child age 5-16	.105	8.7	.036	3.8	.274	8.1
parent w/child age 17-21	.025	1.0	.013	.7	.040	.6
live at home w/parents	.019	1.4	.015	1.4	.032	.9
day=saturday	.003	.3	.023	2.6	-.058	-1.9
day=Sunday	-.259	-23.3	-.148	-16.6	-.423	-15.5
income data missing	.088	4.9	.118	8.1	-.055	-1.3
income is over $100k	-.019	-1.4	-.018	-1.6	-.003	-.1
income (in 2009$)	.004	23.9	.003	22.8	.005	10.8
household has no cars	-.004	-.1	-.151	-5.4	.289	3.7
cars/icense in household	.438	16.6	.314	14.9	.519	8.1
population density	-.261	-13.2	-.336	-21.5	.152	2.8
population densit squared	.052	7.7	.073	13.4	-.050	-2.8
"urban" block group	-.088	-5.1	-.159	-11.6	.214	4.5
"second city" block group	-.127	-10.7	-.200	-21.3	.223	6.7
"suburban" block group	-.075	-7.4	-.118	-14.8	.122	4.4
data given by proxy	-.174	-18.5	-.045	-6.0	-.476	-20.0

Model	A2 (repeated)		A2M		A2F	
Model form	OLS regression		OLS regression		OLS regression	
Dependent variable	log(auto miles+1)		log(auto miles+1)		log(auto miles+1)	
Sample selection	all cases		males only		females only	
Adjusted r-squared	0.174		0.171		0.178	
Variable	**Coefficient**	**T-value**	**Coefficient**	**T-value**	**Coefficient**	**T-value**
Intercept	.834	21.2	.943	18.8	.596	11.0
age 19 to 24	.177	7.2	.180	5.3	.183	5.1
age 25 to 30	.072	3.3	.049	1.7	.112	3.5
age 31 to 36	.042	2.1	.048	1.7	.039	1.3
year 2001	-.051	-2.8	-.059	-2.4	-.041	-1.6
year 2009	-.094	-5.4	-.064	-2.7	-.128	-5.0
age 19-24 in 2001	-.099	-3.2	-.067	-1.6	-.135	-3.0
age 25-30 in 2001	-.045	-1.6	.005	.1	-.104	-2.5
age 31-36 in 2001	-.055	-2.1	-.045	-1.3	-.067	-1.8
age 19-24 in 2009	-.218	-7.3	-.221	-5.3	-.204	-4.7
age 25-30 in 2009	-.140	-5.1	-.138	-3.7	-.146	-3.6
age 31-36 in 2009	-.106	-4.2	-.108	-3.1	-.108	-2.9
male	-.114	-3.4				
driver's license	1.028	43.4	1.059	42.2	1.131	39.3
male w/ license	.167	4.9				
Hispanic	.078	5.3	.044	2.2	.116	5.4
race=black	.022	1.3	.036	1.6	.009	.3
race=asian	-.029	-1.3	-.047	-1.5	-.004	-.1
race=other non-white	.035	1.5	.018	.5	.056	1.6
full time worker	.514	46.7	.451	33.4	.609	29.9
part time worker	.376	27.1	.358	22.0	.410	15.3
one person household	.009	.5	.010	.4	.007	.3

parent w/child age 0-4	.076	7.0	.031	2.0	.107	6.7
parent w/child age 5-16	.146	11.6	.123	7.3	.164	8.7
parent w/child age 17-21	.075	2.8	.047	1.4	.118	2.8
live at home w/parents	.046	3.3	.120	5.8	-.002	-.1
day=saturday	.039	3.3	.097	5.9	-.027	-1.6
day=Sunday	-.211	-18.3	-.132	-8.3	-.303	-17.9
income data missing	.072	3.9	.053	2.1	.089	3.2
income is over $100k	-.024	-1.7	-.004	-.2	-.047	-2.2
income (in 2009$)	.003	20.6	.004	15.6	.003	13.2
household has no cars	-.483	-13.7	-.538	-11.2	-.429	-8.2
cars/icense in household	.706	25.8	.579	15.3	.849	21.4
population density	-.235	-11.5	-.259	-9.2	-.206	-6.9
population densit squared	-.007	-1.0	-.005	-.5	-.010	-1.0
"urban" block group	-.153	-8.5	-.136	-5.5	-.170	-6.5
"second city" block group	-.139	-11.2	-.133	-7.9	-.144	-8.0
"suburban" block group	-.085	-8.0	-.071	-4.9	-.101	-6.6
data given by proxy	-.164	-16.8	-.186	-12.6	-.154	-11.8

Model	A2 (repeated)		D1		D2	
Model form	OLS regression		OLS regression		OLS regression	
Dependent variable	log(auto miles+1)		log(auto miles+1)		log(auto miles+1)	
Sample selection	all cases		all cases		all cases- weighted	
Adjusted r-squared	0.174		0.179		0.212	
Variable	**Coefficient**	**T-value**	**Coefficient**	**T-value**	**Coefficient**	**T-value**
Intercept	.834	21.2	.993	21.7	1.084	25.9
age 19 to 24	.177	7.2	.161	5.5	.170	6.9
age 25 to 30	.072	3.3	.075	3.4	.116	6.0
age 31 to 36	.042	2.1	.047	2.4	.051	2.8
year 2001	-.051	-2.8	-.044	-1.6	-.022	-.8
year 2009	-.094	-5.4	-.047	-1.7	-.060	-2.1
age 19-24 in 2001	-.099	-3.2	-.074	-1.7	-.069	-1.8
age 25-30 in 2001	-.045	-1.6	-.042	-1.3	-.089	-2.7
age 31-36 in 2001	-.055	-2.1	-.058	-1.9	-.015	-.5
age 19-24 in 2009	-.218	-7.3	-.206	-4.8	-.204	-5.0
age 25-30 in 2009	-.140	-5.1	-.170	-5.3	-.150	-4.4
age 31-36 in 2009	-.106	-4.2	-.114	-3.9	-.072	-2.3
male	-.114	-3.4	-.015	-.4	-.071	-2.2
driver's license	1.028	43.4	.991	41.3	.985	46.1
male w/ license	.167	4.9	.119	3.5	.129	4.3
Hispanic	.078	5.3	.086	2.4	.158	7.1
race=black	.022	1.3	.036	2.1	.015	1.1
race=asian	-.029	-1.3	-.047	-2.1	-.013	-.6
race=other non-white	.035	1.5	.037	1.5	.054	2.1
full time worker	.514	46.7	.426	17.6	.465	21.6
part time worker	.376	27.1	.278	10.8	.321	13.7
one person household	.009	.5	-.007	-.3	-.008	-.5

female w/child age 0-4	.076	7.0	.044	3.0	.041	2.7
female w/child age 5-16	.146	11.6	.140	8.6	.109	6.4
female w/child age 17-21	.075	2.8	.071	2.1	.048	1.4
male w/child age 0-4	.076	7.0	.108	7.0	.121	8.1
male w/child age 5-16	.146	11.6	.171	9.4	.167	9.2
male w/child age 17-21	.075	2.8	.124	3.0	.150	3.8
male at home w/parents	.046	3.3	-.027	-.9	-.025	-1.0
female at home w/parents	.046	3.3	.102	3.5	.058	2.2
day=saturday	.039	3.3	.040	3.4	.102	8.8
day=Sunday	-.211	-18.3	-.210	-18.2	-.136	-12.0
income data missing	.072	3.9	.020	1.0	-.028	-1.4
income is over $100k	-.024	-1.7	-.025	-1.7	-.075	-4.7
income (in 2009$)	.003	20.6	.003	13.7	.003	14.7
household has no cars	-.483	-13.7	-.554	-14.8	-.628	-19.2
cars/icense in household	.706	25.8	.668	24.4	.565	22.7
population density	-.235	-11.5	-.104	-4.6	-.162	-7.6
population densit squared	-.007	-1.0	-.047	-6.2	-.036	-5.3
urban designation			-.158	-14.6	-.136	-11.7
"urban" block group	-.153	-8.5	-.155	-8.6	-.148	-9.3
"second city" block group	-.139	-11.2	-.117	-9.3	-.105	-8.4
"suburban" block group	-.085	-8.0	-.072	-6.7	-.088	-8.2
data given by proxy	-.164	-16.8	-.161	-16.4	-.155	-15.9
New interaction terms			**Coefficient**	**T-value**	**Coefficient**	**T-value**
male - 2001			.015	.5	.139	4.7
male - 2009			-.118	-4.5	-.055	-2.0
male age 19-24 in 2001			-.091	-2.2	-.136	-3.3
male age 25-30 in 2001			-.043	-1.1	-.136	-3.6
male age 31-36 in 2001			-.022	-.6	-.096	-2.6

male age 19-24 in 2009		.007	.2	-.092	-2.3
male age 25-30 in 2009		.060	1.7	-.093	-2.3
male age 31-36 in 2009		-.005	-.2	-.111	-3.0
Hispanic - 2001		.008	.2	.000	.0
Hispanic - 2009		.039	1.0	-.032	-1.1
live w/parents - 2001		.064	1.8	.048	1.5
live w/parents - 2009		.006	.2	.018	.6
not employed - 2001		.028	.9	.074	2.4
not employed - 2009		-.159	-5.4	-.119	-4.0
not employed- age 19-24		-.035	-.7	.025	.6
not employed- 19-24 in 2001		-.062	-.9	-.199	-3.3
not employed- 19-24 in 2009		-.029	-.5	-.010	-.2
no cars - age 19-24		.334	4.0	.485	6.8
no cars - 19-24 in 2001		-.088	-.8	-.282	-3.0
no cars - 19-24 in 2009		-.330	-2.6	-.537	-5.2
New variables		**Coefficient**	**T-value**	**Coefficient**	**T-value**
looking for work		-.128	-2.9	-.036	-.9
looking for work in 2009		.079	1.4	-.060	-1.1
occupation=student		.076	2.3	.130	4.1
occ=student in 2009		.243	5.4	.123	2.8
education=some college		.117	6.2	.113	6.9
education=bach.degree		.129	6.1	.136	7.2
education=graduate degree		.056	1.9	.101	3.5
some college- 2001		.022	.9	.011	.5
bach. degree - 2001		.049	1.8	.017	.7
graduate degree - 2001		.026	.7	-.028	-.7
some college- 2009		.011	.4	.005	.2
bach. degree - 2009		.071	2.6	.038	1.4

graduate degree - 2009		.114	3.1	.037	1.0
poverty status		-.030	-1.0	-.111	-4.3
poverty status - 2001		-.038	-1.0	.064	1.9
poverty status - 2009		-.037	-1.0	.002	.1

Our Re-examination: Summary and Recommendations

The analysis reported by McDonald has been the most complete multivariate disaggregate analysis of national household travel survey (NHTS) data to date. Through the use of interactions between age group effects and survey year effects, McDonald was able to show strong evidence that Millennials travelled less by auto than in 2009 than what can be explained by a wide variety of other variables that were controlled for in the analysis.

In the work reported here, we were able to replicate McDonald's analysis, and then test a number of different ways of extending the analysis.

First, a number of new explanatory variables and interaction effects were added to the model that could potentially help explain why Millennials' auto use declined in 2009 relative to what would be indicated by the other variables controlled for in the analysis. Although several of these added variables were significant and helped explain some trends in travel behavior, the interaction effects between age group and survey year reported by McDonald remained significant over and above the new variables added.

We also tested a number of different model specifications, including log-linear regression models, TOBIT regression models, and a two-stage binary logit and regression model approach. Again, although these changes gave better improved model fit and some further explanatory power compared to the simpler OLS regression method, the interaction effects related to Millennials' reduced auto use remained at about the same values, and become even more statistically significant.

The one change tested that had the largest effect on the model results was to use vehicle-miles travelled (VMT) rather than person-miles traveled (PMT) as the dependent variable. With the model based on VMT, the interaction effects between age group and survey year remained statistically significant, but only about half as large as the same variables in a model based on PMT. This finding indicates that the "unexplained" relative decline in auto miles traveled among Millennials has been somewhat more pronounced for auto passenger trips than for auto driver trips.

Given that the evidence remains strong that Millennials in 2009 reduced their auto travel more than the older age groups did, all else equal, we recommend further research in a few different directions. One is to repeat this analysis when the 2017 NHTS data becomes available. If, as one hopes, the data from the next survey year is compatible with the data from the past three survey years for analysis purposes, having another year of data can establish whether these same "unexplained" differences have persisted since 2009 as the economy has improved and the entire Millennial generation has reached driving age.