APPENDIX J

METHODS TO ESTIMATE ANNUAL MILES DRIVER PER VEHICLE

J-1. INTRODUCTION

In the 2001 NHTS, the amount of driving (VMT) driven by an NHTS household vehicle could be estimated in three different ways. First, one can annualize the odometer readings recorded approximately two to three months apart. Second, a designated household member was asked to report the total number of miles driven in each of the household vehicles (hereafter referred to as “self-reported VMT”). Finally, the amount of annual driving can be estimated based on the amount a vehicle is driven during the designated sample day (i.e., the travel day). Ideally, annualizing the odometer readings would probably generate the most reliable VMT estimate compared to estimates based on the other two approaches. Unfortunately, not all vehicles had their odometer readings recorded. Furthermore, of those that had their odometer reading recorded, the quality of some of the odometer readings is less than desirable. As such, ORNL was asked to estimate the number of miles driven by each of the NHTS vehicles based on the best available data. This estimate is hereafter referred to as the *BESTMILE*. *BESTMILE*s are furnished only for automobiles, pickup trucks, vans, sport utility vehicles, and motorcycles. The value of the *BESTMILE* for motorcycles equals the value of the self-reported VMT. There are no *BESTMILE* estimates for other trucks or recreational vehicles (RV). The *BESTMILE* estimates were computed only for vehicles in the National Sample households.

The *BESTMILE* estimates were developed using the National sample of the NHTS data (Version 1, dated January 2003). Since then, some of the data used in the development of the *BESTMILE* estimates were modified during the editing process for the January 2004 version of the NHTS. These modifications affected data on 187 sampled vehicles including data on vehicle type (*VEHTYPE*), self-estimated annual miles driven (*ANNMILES*), household vehicle count (*HHVEHCNT*), and other variables that were used to estimate the annual miles driven per vehicle. Due to the resource constraints in this study, the *BESTMILE* estimates for these vehicles were not updated to reflect January 2004 data. For these vehicles, *BESTMILE* was set to “Not Ascertained,” and the associated *BEST_FLG* (i.e., How the *BESTMILE* was computed) was set to “No Best_Estimate, underlying values changed in editing, and all other associated variables were changed accordingly.”

J-2. QUALITY OF ODOMETER READINGS

Odometer readings were collected for each household vehicle at two points in time. The first was at or around the time of the person interviews. The second was at least two months later. The dates of each reading were recorded to facilitate the estimation of annual mileage. Of the 31,939 vehicles with two valid recording dates, 3% of them recorded odometer readings less than 2 months apart while less than 1% recorded the readings more than one year apart (Table J-1).

Table J-1. Lag Time between Two Odometer Readings 2001 NHTS

Lag Time in Days	No. of NHTS Vehicles with two valid dates	% of NHTS Vehicles
0 - 30	100	0.3%
31 - 60	979	3.1%
61 - 90	5,367	16.8%
91 - 120	12,121	38.0%
121 - 150	1,871	5.9%
151 - 180	2,157	6.8%
181 - 210	2,889	9.1%
211 - 365	6,235	19.5%
> 365	220	0.7%
Total	31,939	100.0%

* Applies to 31,939 vehicles that have two valid recording dates.

To determine whether odometer reading data are usable for estimating annual mileage, they were checked with respect to:

(1) the *completeness* of data – Both the beginning and the ending odometer readings and the corresponding recording dates are reported.

(2) the *reasonableness* of the readings – Example of such reasonableness checks are that the second reading must be greater than the first reading, and the second date chronologically follows the first date.

(3) the *consistency* to the self-reported VMT – If the odometer readings of a vehicle are not reliable for estimation purposes, its self-reported VMT is used. Therefore, the relationship between the difference between two odometer readings and the self-reported VMT needs to be reasonable and consistent. Arbitrary boundaries were set to determine the “consistency” between the self-reported VMT and the difference between two odometer readings. If the ratio of the odometer-based daily driving to the self-reported daily driving was

greater than 4 or less than 0.25, and the difference between the two VMT estimates is greater than 10,000 miles per year, then the odometer readings were considered *unusable*.

After accounting for the vehicle types that are excluded for the purpose of estimating the *BESTMILE* and the vehicles that failed the data quality checks, 25,292 of the 53,278 vehicles in the national sample had reasonable odometer readings and other key data elements that could be used to estimate *BESTMILE* based on odometer readings (Table J-2).

Table J-2. 2001 NHTS Sample Vehicles by Data Required to Estimate Odometer-based *BESTMILE*

Data Quality Checks	Number of Sample Vehicles*	%
Total 2001 NHTS Vehicles (National Sample)	**53,278**	**100.0**
- Incomplete odometer readings and/or date data	21,357	40.09
- Negative differences between 2 odometer readings	2	0.00
- Differences between 2 odometer readings too large (more than 550 miles per day)	64	0.12
- The second reading chronologically proceeded the first one.	7	0.01
Vehicles with usable odometer reading data (Total vehicles less all of the above)	**31,848**	**59.77**
- No primary driver associated with the vehicle	3,492	6.55
- Out-of-scope vehicles (such as “Other truck”, RV, motorcycle, “other”, “don’t know” vehicle types)	943	1.77
- The ratio of odometer-based daily driving to self-reported daily driving is outside the range of 0.25 - 4.0; and the absolute difference between the two driving estimates is greater than 10,000 miles per year	2,121	3.98
Vehicles with their *BESTMILE* estimated based on odometer readings	**25,292**	**47.47**

* Data quality checks were done **sequentially**. Therefore, the number of vehicles reported is those that have passed all of the preceding data quality check criteria. For example, 64 vehicles that report too large a difference between 2 readings had passed preceding data quality checks on missing readings or dates, and on negative difference between 2 readings.

J-3. EXTENT OF DATA AVAILABILITY

Given that only 25,292 vehicles could have their *BESTMILE* estimated based on odometer readings, the next question was how to estimate *BESTMILE* for the remaining 25,558 vehicles. Even with usable odometer readings, it was felt that the *BESTMILE* can be more accurately estimated if other driving-related data were taken into account, such as the self-reported VMT and the characteristics of the primary driver.

Therefore, the method to estimate the *BESTMILE* is largely dictated by the extent to which information is available for a vehicle. The method becomes less sophisticated as less data are available for an individual vehicle. The NHTS sample vehicle population was grouped into six categories based on the availability and usefulness of the key data elements that are necessary to estimate the number of miles driven by a vehicle in a year. The key data elements are: (1) usable odometer readings at two time points, (2) the self-reported VMT, (3) the number of miles driven on the travel day, and (4) the characteristics of the vehicle's primary driver. Because characteristics of the primary driver are used in the estimation of *BESTMILE*, the *BESTMILE*s are not estimated for vehicles that are regularly driven by more than one person or vehicles where a primary driver was not indicated. The vehicle population was further categorized by the relationship between the number of vehicles and the number of drivers in the household. The rationale for this additional classification is the hypothesis that vehicles in a household where there are more drivers than vehicles are likely to be driven more than those in a household where there are fewer drivers than vehicles, everything else being equal. Table J-3 summarizes the distribution of the NHTS sample vehicles based on the availability of these key data elements. This distribution helps outline how BESTMILE is estimated for each subgroup of vehicles (Section 5).

Table J-3. NHTS Vehicles[1] by Data Required for BESTMILE Estimation

	Usable Data to Estimate Odometer-based *BESTMILE*?					
	Yes		*No*			
	Usage self-reported VMT?		**Usage self-reported VMT?**			
	Yes	*No*	*Yes*		*No*	
	Information on Primary Driver?		**Information on Primary Driver?**		**Information on Primary Driver?**	
	Yes		*Yes*	*No*	*Yes*	*No*
One driver/One vehicle HHs	3,137	129	1,835	8	176	22
Two drivers/two vehicles HHs	8,729	1,143	5,427	968	946	394
Other 'Drivers=Vehicles' HHs	1,996	393	1,718	238	469	159
'Drivers > Vehicles' HHs	1,341	208	1,189	444	230	182
'Drivers < Vehicles' HHs	7,357	859	6,284	2,492	1,100	1,277
Subtotal	22,560	2,732	16,453	4,150	2,921	2,034
Subtotal by Usable Data	***25,292***		***25,558***			

[1] There were 53,278 vehicles included in the NHTS national sample. However, 2,428 of these vehicles were out of scope for the *BESTMILE* estimate. The out-of-scope vehicle types include "other trucks," "recreational vehicles," and vehicles with missing vehicle type information.

J-4. ADJUSTING ODOMETER READINGS TO A FIXED TWELVE-MONTH TIME FRAME

The recordings of odometer readings began on January 1, 2001 and ended on December 30, 2002, spanning a period of twenty-four months. In order to facilitate the estimation of fuel economy (MPG), fuel consumption, and fuel costs (Appendix K), the odometer readings were standardized to a 12-month period, from May 1, 2001 to April 30, 2002. This time frame was selected because it contained the largest proportion of odometer readings compared to all other possible time spans beginning on the first day of a given month. The goal, therefore, was to adjust odometer readings that were recorded outside the designated 12-month period to within the designated 12-month period. This adjustment was not done in the 1995 NPTS.

Aggregate VMTs reported in Federal Highway Administration's (FHWA) *Highway Statistics* were used to account for seasonal and yearly differences in driving (Table J-4). For example, the seasonal difference in driving between June and July of 2001 was that the amount of driving in July was about 2.6% greater than that in June. And, the difference in driving between June and July of 2002 was 3.2%. It was assumed that the seasonal difference and the difference from one year to the next apply to the amount of driving by individual vehicles.

Therefore, if the odometer readings of a vehicle were recorded on July 1 and July 31, **2002** (outside the designated 12-month period), then the adjustment needs to be made as if the readings were recorded on July 1 and 31, **2001** (within the designated period), respectively. Since the dependent variable for estimating the *BESTMILE* was the miles driven per day, the adjustment to this vehicle was

$$\frac{Odometer2 - Odometer1}{Date2 - Date1} \times \frac{VMT_{July2001}}{VMT_{July2002}} \left(= \frac{248.8}{254.2} \right)$$

where Odometer 2 and Date 2 are the reading and the recording date of the second odometer reading while Odometer 1 and Date 1 are the first reading and the recording date. According to Table J-4, driving on July 2001 is 97.8% (=248.8/254.2) of that on July 2002.

Table J-4. Total Miles Traveled (VMT) in the United States for All Systems (in Billions)

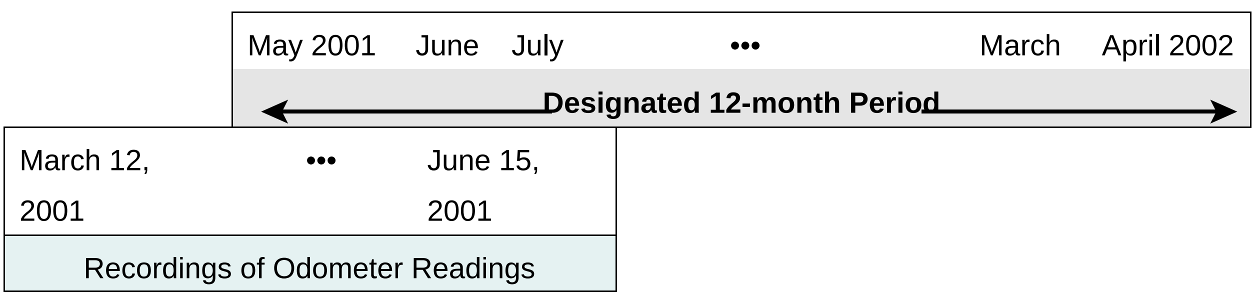
Month	Year 2001	Year 2002
January	209.3	213.8
February	199.9	206.5
March	231.5	234.0
April	231.4	234.8
May	244.3	249.9
June	242.6	246.4
July	248.8	254.2
August	251.7	256.2
September	224.6	230.6
October	240.0	243.2
November	229.5	228.4
December	228.1	231.3

Let's assume that (Odometer 2 - Odometer 1) = 300 miles. Without any adjustments to the designated 12-month period, the number of miles driven per day is 10 miles (=300 miles/30 days). However, if the odometer readings were recorded on July 2001 instead of

July 2002, then the number of miles driven per day for the days on July 2001 would have been 9.8 miles (=$\frac{300 miles}{30 days} \times \frac{248.8}{254.2}$).

The adjustment procedure is complicated when there were multiple months involved and/or odometer readings were recorded over two years. The adjustment procedure became even more complicated if the recordings of the odometer readings straddled dates *inside* and *outside* the designated time period, such as if the odometer readings were recorded from March 12, 2001 to June 15, 2001 (Figure J-1). Rather than shifting the entire recording period to "inside" the designated time period, the procedure adjusted only the miles driven <u>outside</u> the designated period, and left the miles driven <u>inside</u> the designated period intact. In this example, the task was to adjust the number of miles driven during the period from March 12, **2001** through April 30, **2001**, as if the driving had taken place from March 12, **2002** through April 30, **2002**.

Figure J-1. Example of Odometer Readings Straddling Dates *Inside* and *Outside* the Designated 12-Month Period



Rather than adjusting the number of miles driven between the first and the second odometer readings, the adjustment was made to the shares of driving attributable to each of the forty-eight different temporal categories. These forty-eight categories are the unique combinations of 2 years, 12 months, and weekend vs. weekday. A set of forty-eight percentages was created for each NHTS vehicle, each percentage being the proportion of days in a given category over the total number of days between two readings. Each of these 48 percentages was used as a proxy of the share of driving that was attributable to the days in that year/month/weekday (vs. weekend) category. The sum of these 48 percentages is equal to 1.

The rationale for adjusting the proportion of days in each of the 48 categories, rather than the total miles driven, was the modeling specifications of the *BESTMILE* estimation approaches. First, the dependent variable of the *BESTMILE* estimation approaches was the miles driven per day. Second, terms were included in the models to account for the temporal differences in driving (i.e., the seasonal, the yearly, and weekend/weekend differences in driving). The adjustments to the percentages of days, rather than the total miles driven between two readings, facilitated the modeling process.

Table J-5 illustrates the rationale and steps taken to adjust the odometer readings recorded from March 12 through June 15, 2001. Reading from March 12 through April 30 was outside the designated period while the remaining dates were inside the period. The odometer readings were recorded over a period of 96 days, from March 12 through June 15. Therefore, the contribution of *weekday* driving in March 2001 toward the total driving during the 96-day period was approximated to be 15.6%, while the corresponding *weekend* contribution was 5.2%. May and June percentages were not adjusted because they were within the designated time period. Since *Highway Statistics* does not categorize VMT by weekday vs. weekend, the adjustment factors for weekend-days and weekdays are identical.

Based on *Highway Statistics*, the driving in March 2001 was 98.9% (=231.5/234.0) that in March 2002 (Table J-4). To adjust the March 2001 driving as if it took place in March 2002, the share of March 2001 days was adjusted by multiplying it by 0.989. Similarly, the share of April 2001 days was adjusted by multiplying it by 0.986 (=231.4/234.8). The percentages were re-calculated after the adjustments [Column (7)].

Table J-5. Adjustments to Miles Driven during March 12 through June 15, 2001

Month (1)	(2)	No. of Days (3)	% of Days (4)	Adjustment factor[1] (5)	Adjusted No. of Days (6)	% of Adjusted Days (7)
March	Weekday	15	15.6%	0.989=231.5/234.0	14.8=15×0.989	15.6%
	Weekend	5	5.2%	0.989=231.5/234.0	4.95=5×0.989	5.2%
April	Weekday	21	21.9%	0.986=231.4/234.8	20.7=21×0.986	21.7%
	Weekend	9	9.4%	0.986=231.4/234.8	8.87=9×0.986	9.3%
May	Weekday	23	24.0%	*	23	24.1%
	Weekend	8	8.3%	*	8	8.4%
June	Weekday	11	11.5%	*	11	11.5%
	Weekend	4	4.2%	*	4	4.2%
Total		96	100.0%		95.36	100.0%

[1] Based on the total monthly VMT as reported in the *Highway Statistics*.

* No adjustments were made because May and June were inside the designated time period.

These 48 percentages were then consolidated into 24 terms for modeling purposes by adding together the proportions of days in the same month (e.g., March) but in different years. For example, if odometer readings were recorded from March 2001 through March 2002. Then, the proportion of weekdays in March 2001 (which was outside the designated period) was adjusted and added to the unadjusted proportion of weekdays in March 2002 (which was inside the designated period). A single term was then created to represent the contribution of *weekday* driving in March 2002. Similarly, a single term was created to represent the contribution of *weekend* driving in March 2002. This consolidation was necessary so that the temporal differences in driving are represented by 23 terms (*n1* - *n23*) in the *BESTMILE* estimation approaches (Section 5). Driving attributable to the weekends in December 2001 was the baseline in the estimation models; thus there were only 23 terms.

J-5. ESTIMATION METHODS

Six different estimation approaches were developed.

Approach 1. *For vehicles with usable odometer readings, self-reported VMT, and information on the primary driver.*

There were 22,560 vehicles in this category (Table J-3). This approach assumes that the daily driving of a vehicle, which is calculated based on two odometer readings, is a function of:

- the daily driving based on self-reported VMT,
- the characteristics of the primary driver, and
- other household and geographical attributes.

Since the number of drivers and vehicles in a household affects the amount of driving per vehicle, models were estimated separately for three different types of households: (1) households with one vehicle and one driver, (2) multi-driver households with an equal number of vehicles and drivers, and (3) households with unequal numbers of vehicles and drivers. The models are represented in Equation (1),

$$Y \quad X\beta \quad R, \qquad (1)$$

where *Y* is the vector of observed average daily mileages (based on odometer readings), *X* is the vector of independent variables, β is the matrix of model parameter estimates, and *R* is the vector of residuals. The vector of independent variables, *X*, includes the month-weekend/weekday terms (*n1* - *n23*), daily self-reported VMT (*ANNMILES/365*), education level (*EDUC*), age of the respondent (*R_AGE*), vehicle age class (*VEHAGEC*), vehicle type (*VEHTYPE*), area size (*MSASIZE*), census division (*CENSUS_D*), life cycle of the household (*LIF_CYC*), worker status and gender of the primary driver (*WORKER* and *R_SEX*, respectively), and size of the household (*HHSIZE*). The model for the case with an unequal number of drivers and vehicles also used a categorical variable for the driver to vehicle ratio (*DRVEH*).

The seasonally-adjusted daily driving estimates are computed using the residual for each vehicle, plus the predicted value of the model, and adjusted to values of *n1* through *n23* to reflect the twelve months from May 2001 to April 2002. The residual was retained since the goal was to create seasonally-adjusted annualized estimates, as opposed to predictions completely free from random noise. The best estimate on the annual driving per vehicle was computed by multiplying this seasonally-adjusted daily driving by 365.

Approach 2. *For vehicles with self-reported VMT, and information on the primary driver, but <u>without</u> usable odometer readings.*

Equation (1) was used to estimate their annual driving for the 16,453 vehicles with usable self-reported VMT and information on the primary driver, but without usable odometer readings. However, since information from these vehicles was not used to estimate the model, there were no residuals. As such, the estimates for this group of vehicles have a significantly small random noise term than the 22,560 vehicles used to estimate the model. To overcome this problem, residuals from the original 22,560 vehicles were randomly assigned, *without* replacement, to the 16,453 vehicles.

If an estimated $\hat{y}$ is less than 0 or greater than 200,000 miles per year, then a second randomly assigned residual was used. This process was repeated one more time if the estimated $\hat{y}$ was still outside the range of 0 and 200,000 miles per year. However, after this point, if $\hat{y}$ was still outside the range, then *BESTMILE* was set at 0 or 200,000 miles.

Approach 3. *For vehicles with self-reported VMT, but <u>without</u> odometer readings and usable information on the primary driver.*

There were 4,150 vehicles in this category (Table J-3). Although odometer readings were missing for these vehicles, the strong relationship between self-reported VMT and odometer readings suggested the following estimation approach:

$$BESTMILE_i = \hat{\alpha} + \hat{\beta} \times ANNMILES_i + R_i \qquad (2)$$

where $\hat{\alpha}$ and $\hat{\beta}$ are from Equation (1). The pseudo-residuals were assigned similar to Approach 2, *without* replacement.

Approach 4. *For vehicles with usable odometer readings and information on the primary driver, but <u>without</u> self-reported VMT.*

There were 2,732 vehicles in this category (Table J-3). The estimation model was similar to Equation (1), except for the exclusion of the self-reported VMT. Furthermore, the small sample sizes precluded separate estimation models from being developed for households with different ratios of vehicles to drivers. Instead, the *DRVEH* variable was included in the model. The *BESTMILE* estimate for this group of vehicles was $\hat{y}$ multiplied by 365.

Approach 5. *For vehicles with usable information on the primary driver, but <u>without</u> odometer readings and self-reported VMT.*

There were 2,921 vehicles in this group (Table J-3). The estimation model was that developed based on data from the vehicles described in Approach 4. $\hat{y}'s$ for this group of vehicles had the issue of a small random noise term, similar to the vehicles included in Approach 2. Therefore, this group of 2,921 vehicles required the assignment of pseudo-residuals. However, because the estimation model was developed based on information from 2,732 vehicles and there were 2,921 vehicles requiring pseudo-residuals, the residuals were randomly assigned *with* replacement.

Approach 6. *For vehicles with no driving information except that collected on the travel day.*

There were 2,034 vehicles that had no usable odometer readings, no self-reported VMT, and no information on the primary driver. Of these, driving was recorded for only 958 vehicles during the travel day while the remaining 1,076 vehicles were assumed not to be used on that day. For the 958 vehicles, the total number of miles driven on the travel day was annualized by multiplying the number of miles driven in the travel day by 365. Since a vehicle was unlikely to be driven every day of the year, the annualized miles were adjusted by the probability that a vehicle was driven on a typical weekday, or weekend, depending on whether the travel day was a weekday or a weekend. The equation below indicates how the *BESTMILE* was estimated for vehicles in which their travel day fell on a weekday:

$$\begin{aligned} \mathit{BESTMILE} = {} & 365 \times (\text{Miles driven on the travel day}) \\ & \times \mathit{Prob}(\text{vehicle was driven on a weekday}) \\ & \times \mathit{Mean}(\text{miles driven on a weekday}) \end{aligned} \tag{3}$$

where *Prob*(vehicle was driven on a weekday) is the weighted proportion of vehicles driven on a *weekday* travel day to all vehicles; and *Mean*(daily miles driven on a weekday) is the weighted average of miles driven for vehicles driven on a *weekday* travel day. A similar approach was used for vehicles that were driven on a travel day that was a weekend.

BESTMILES was missing for vehicles that were not used on the designated travel day and that had no information on their primary drivers or driving patterns (e.g., odometer readings, self-reported VMT). There were 2,080 such vehicles.

J-6. SCREENING OF BESTMILE ESTIMATES

Once the *BESTMILE* estimates were derived using one of the aforementioned approaches, they were evaluated for *reasonableness*. With two odometer readings recorded less than one year apart, if the value of the *BESTMILE* is less than the difference between two odometer readings, then the *BESTMILE* is deemed unreasonable. Furthermore, since the estimation approaches do not constrain the values, some *BESTMILE* estimates could be negative. As in 1995, adjustments were made to the estimated annual mileage if it had a negative value, or if it was less than the difference between the two odometer readings. Similar to the approach used in the 1995 NPTS, unreasonable *BESTMILE* estimates were adjusted in one of two ways. First, if odometer readings were present and valid, then the *BESTMILE* is

$$\frac{(Odometer2 - Odometer1)}{(Date2 - Date1)} \times 365 \qquad (4)$$

where *Odometer2* and *Date2* are the reading and the recording date of the second odometer reading while *Odometer1* and *Date1* are for the first reading. The annual milage calculated using Equation (4) is hereafter referred to as the *Crude Estimate*. However, If no odometer readings were present or valid, then the *BESTMILE* was set at 0.

Finally, in keeping with the cap of 200,000 miles on the self-estimated VMT, if the value of *BESTMILE* exceeds 200,000 miles per year, then the estimate was set at 200,000 miles per year. Table J-6 summarizes various adjustments to the *BESTMILE* estimates.

Table J-6. Adjustments to *BESTMILES*

Adjustment Code	Frequency	Percent	Criteria	Adjustment
	52,361	98.28%	No adjustment	None
1	798	1.50%	No. of days between two odometer readings is less than 365 and the *BESTMILE* is less than the difference between two readings	*BESTMILE* = difference between two readings
2	5	.0.1	No. of days between two odometer readings is greater than 365 and the *BESTMILE* is greater than the difference between two readings	*BESTMILE* = difference between two readings
3	8	0.02%	No. of days between two odometer readings is greater than 365 and the value of *BESTMILE* is negative	*BESTMILE = Crude Estimate*
4	3	0.01%	*BESTMILE* > 200,000 miles	*BESTMILE* = 200,000 miles
6	103	0.19%	No usable odometer readings and *BESTMILE* is negative	*BESTMILE* = 0
Total	**53,278**	**100.00%**		

Consistently to the data quality checks used in the 1995 NPTS, each *BESTMIL*E estimate was compared to its *Crude Estimate* (Equation (4)) and the corresponding self-reported VMT. Outlier codes were then assigned on the basis of these comparisons and the subjectively determined thresholds (Table J-7). Because the self-reported VMTs were considered less reliable than the *Crude Estimates*, the thresholds are tighter for the *Crude*-vs-*BESTMILE* comparisons. Codes based on comparisons of the *BESTMILE* estimate and the *Crude Estimate* were only assigned if the difference exceeded 5,000 miles. Codes based on comparisons of the *BESTMILE* estimate and the self-reported estimate were only assigned if the difference exceeded 10,000 miles. The outlier codes were recorded as numeric codes (*BEST-OUT*) as indicated in Table J-7.

Table J-7. Outline Codes of *BESTMILES*

BEST_OUT	Frequency	Percent	Criteria
No code	52,155	97.89%	
1	10	0.02%	$BESTMILE < \frac{CruedEstimate}{2}$ and $\lvert BESTMILE - CrudeEstimate \rvert > 5{,}000 miles$
2	173	0.32%	$BESTMILE < \frac{Self - reportedVMT}{4}$ and $\lvert BESTMILE - Self - reportedVMT \rvert > 10{,}000 miles$
4	9	0.02%	$BESTMILE > CrudeEstimate \times 2$ and $\lvert BESTMILE - CrudeEstimate \rvert > 5{,}000 miles$
5	931	1.75%	$BESTMILE. > Self - reportedVMT \times 4$ and $\lvert BESTMILE - Self - reportedVMT \rvert > 10{,}000 miles$
Total	53,278	100.00%	