**Supporting Information**

# In-Use Vehicle Emissions Test Results in Beijing Analyzed Using Logistic Regression

Cheng Chang and Leonard Ortolano*

Department of Civil and Environmental Engineering, Stanford University
Jerry Yang & Akiko Yamazaki Environment & Energy Building
Room 249, Stanford, California 94305-4020

Number of pages: 17

Number of tables: 8

Number of figures: 0

## Variables in Data Set

Table S1 shows the abbreviations and definitions used to represent the predictor variables after the raw data were coded. Information in the table reflects various characteristics of tested vehicles, such as vehicle model and model year, ownership, and inspection station. Descriptive statistics were computed for DOMESTIC and CLASS, but these variables were not used in the regression studies because of collinearity problems; those two variable are determined from the vehicle model.

**TABLE S1. Summary of the Predictor (or “Independent”) Variables**

| Variable abbreviation | Range of value | Definition |
|---|---|---|
| MODEL | 1-52 | Vehicle model |
| DOMESTIC | 1 | Vehicle produced by a domestic manufacturer |
| | 0 | Vehicle produced by a foreign company or a joint venture involving a foreign investor |
| OWNERSHIP (Type) | 1 | Government |
| | 2 | Business |
| | 3 | Individual |
| MY (Model Year) | 1999-2002 | Vehicle model year |
| SITE | 1-16 | Inspection station number |
| CLASS | 1 | Luxury sedan |
| | 2 | Economic sedan |
| | 3 | Truck, van, and minivan* |
| SUBURBAN (Registration Location) | 1 | Vehicle registered at a suburban office |
| | 0 | Vehicle registered at a urban office |

* Hereinafter, when “truck” is used to describe a category of the CLASS variable, it refers to a category that includes trucks, vans and minivans.

## Vehicle Emissions Testing Procedures

Details of the certification and emissions testing procedures are given below. To clarify distinctions between certification and emissions testing, consider an example involving two vehicles: a 1,500kg (3,307lb) sedan and an 1,800kg (3,968lb) light-duty truck. Furthermore, assume that both were subject to Euro I emissions standards for *new* vehicle certification when they were sold in the Beijing market in 2000. Table S2 illustrates the cutpoints for the two models for type approval under the Chinese Euro I emissions standards; these are in units of g/km. The sedan model has much more stringent limits than the light-duty truck model.

**Table S2. Cutpoints in Chinese Euro I Standards for Type Approval**

| | Reference weight (kg) | Euro I Emissions Standards for New Vehicle Type Approval | |
|---|---|---|---|
| | | CO (g/km) | HC+NOx (g/km) |
| Vehicle 1 (Sedan) | 1,500 | 2.72 | 0.97 |
| Vehicle 2 (Truck) | 1,800 | 6.90 | 1.70 |

In the Beijng ASM emissions standards for in-use vehicles, there is a category of emission concentration cutpoints corresponding to vehicle models subject to Euro I emissions standards. Table S3 lists the cutpoints for the above two illustrative models. As shown in the table, as vehicle weight increases, the cut points for all three pollutants decrease.

**Table S3. Cutpoints in Beijing ASM Standards**

| | Reference weight (**kg**) | Cutpoints in **Beijing** ASM Emissions Standards for In-use Vehicles | | | | | |
|---|---|---|---|---|---|---|---|
| | | 5024 | | | 2540 | | |
| | | CO (%) | HC ($10^{-6}$) | NO ($10^{-6}$) | CO (%) | HC ($10^{-6}$) | NO ($10^{-6}$) |
| Vehicle 1 (Sedan) | 1,500 | 0.6 | 80 | 750 | 0.5 | 75 | 700 |
| Vehicle 2 (Truck) | 1,800 | 0.5 | 75 | 700 | 0.4 | 70 | 690 |

Note: (a) The first two digits in "5024" and "2540" refer to loads exerted by a dynamometer, and the last two digits represent the testing speeds. In the Beijing ASM standards, the high (50%) and low (25%) loads are specified using the following equations: $PA_{5024}$=Rm/150 and $PA_{2540}$=Rm/180, where: $PA_{5024}$ is the high load (kW) on the tested vehicle; $PA_{2540}$ is the low load (kW); and Rm is the vehicle's reference mass, defined as the mass of the vehicle with all equipment plus 100kg. The "24" in "5024" refers to 24 km/hour, equivalent to 15 miles/hour; similarly, the "40" in "2540" refers to 40 km/hour, equivalent to 25 miles/hour. Source: (*1*).

(b) Units of cutpoints in Table S3 (i.e., % and ppm or parts per million, which is noted as $10^{-6}$) refer to volume/volume. In other words, 1% would mean one cubic meter of a pollutant out of a total of 100 cubic meters of exhaust gases.

Because the Beijing ASM standards are similar to the U.S. ASM standards, we use Table S4 to demonstrate what the cutpoints for the above two vehicle models would be if they were tested in the

U.S. (The unit for reference weight in Table S4 is lb, rather than kg in Table S2 and S3). The category of cutpoints that the two vehicle models are subject to would be 1994+ Tier 1 (see Table S4) (*2*).

The general relationships between cutpoints and vehicle's reference weights in both Table S3 and S4 are similar: cutpoints decrease with increasing weights. The U.S. ASM standards are more stringent than the Beijing ASM standards on the whole, except for the HC cutpoints for Vehicle 1 (sedan).

**Table S4. Cutpoints in U.S. ASM Standards**

| | Reference weight (**lb**) | Cutpoints in **US** ASM Emissions Standards for In-use Vehicles | | | | | |
|---|---|---|---|---|---|---|---|
| | | 5015 | | | 2525 | | |
| | | CO (%) | HC ($10^{-6}$) | NO ($10^{-6}$) | CO (%) | HC ($10^{-6}$) | NO ($10^{-6}$) |
| Vehicle 1 (Sedan) | 3,307 | 0.45 | 81 | 647 | 0.43 | 78 | 585 |
| Vehicle 2 (Truck) | 3,968 | 0.39 | 71 | 553 | 0.38 | 68 | 501 |

Note: the "15" in "5015" refers to 15 miles/hour; similarly, the "25" in "2525" refers to 25 miles/hour. Source: (*2*)

In addition to the differences in cut points, the Beijing ASM standards are a little different from the U.S. ones regarding fast pass. In both modes (5015 and 2525) of the U.S. ASM, a fast pass is given if at any point after the initial 25 seconds of testing, the 10-second average emissions are below the cut points for the vehicle. In the Beijing ASM, a fast pass in the 5024 mode is given if at any point after the initial 25 seconds of testing, the 10-second moving average emissions are below the cut points for the vehicle. In contrast, a fast pass in the 2540 mode is given if at any point after the initial 5 to 15 seconds of testing, at least 15 of the 10-second moving average emissions are below the cut points for the vehicle. In addition, the Beijing ASM test allows a vehicle to pass the inspection after passing the initial 5024 mode, while the US ASM test requires a vehicle to go through 2525 tests even if it fast passes the 5015 test.

## Data Quality Control Methods

Our efforts to control data quality are detailed below. For vehicles that passed inspections on the first try, much of the database was produced by the testing equipment in electronic form, and our review of this data revealed no obvious inconsistencies. Vehicle information (e.g., odometer reading, make and model) was manually entered into computers by inspectors. Our inspection revealed no obvious problems, except in the case of vehicle mileage; thus we eliminated odometer readings in our database.

Test results for vehicles failing their first emissions tests during the summer of 2003 were generated by measuring instruments but recorded on paper sheets instead of in computers. Each paper sheet correspondes to one vehicle test failure.

To put the information for vehicles failing their first tests in electronic form, we rented a computer lab at a college in Beijing and hired about 60 college students to manually enter the data into computers. Each of the sheets contained vehicle and owner information as well as test results.

We trained the students to understand the dataset, and we then used various measures to ensure the quality of data entry. For example, we defined and constrained the type and range of inputs for certain spreadsheet cells so that an error message appeared if the wrong type of entry or range was entered.

In addition, we employed the following quality assurance approach. Cheng Chang (the first author) organized the students engaged in data input into 10 groups. Each group had a leader helping Chang check the quality of the inputs. During the first three days of work, Chang and the group leaders randomly checked about 10 percent of the data input daily by the students. Most students had consistently high accuracy (i.e., greater than 99 percent), but a few had accuracy below 99 percent. Chang and the group leaders worked with these few students to improve their performance. Common data entry problems were rapidly identified and resolved.

After the first several days of data input work, Chang reduced the fraction of sheets checked randomly to about 5 percent, but only for those students that demonstrated consistency in attaining accuracy rates greater than 99 percent.

Chang and the group leaders continued to focus intensive checking efforts on the few students with less than 99 percent accuracy. They checked almost all the inputs by these relatively low-accuracy students. Before these students were allowed to continue working, the causes of input errors were identified and all but two of the students improved their accuracy rapidly. Those two students were not allowed to continue their data input. During the last few days of data input work, Chang and the group leaders randomly selected about two percent of the input sheets processed daily for thorough follow up checks.

Throughout the process, whenever shortfalls in accuracy were identified in random checks, the rate of checking was increased on the items causing problems; this allowed for rapid resolution of particular problems.

## Vehicle Model and Main Effects Model

Before presenting results from the main effects model it is necessary to comment on the selection of reference categories. For binary predictor variables, one of the two values (i.e., categories) of the variable is selected as a reference value. For predictor variables that have more than two categories,[1] statistics software packages like *Stata* would randomly drop one of the categories to prevent collinearity in regression models. However, if one of the categories is in some sense a normal or baseline condition, it is a good practice to select the category to delete before the program is run. Among the seven predictor variables, only SUBURBAN and DOMESTIC are binary predictor variables. The other five predictor variables (MODEL, MY, SITE, OWNERSHIP and CLASS) have more than two categories. Thus, we selected one of the several categories within each of these five predictor variables for deletion from the regression models before the models are built. The deleted category is the "reference category." When used with a specific predictor variable, it is often convenient to use the name of the variable in place of the word "category." For example, the reference category for the variable called "MODEL" is also called the "reference model," and the reference category for the inspection station variable is also called "reference station," or "reference site." Modeling results indicate whether some

[1] For predictor variables that have more than two groups, I will use "category" to denote these groups in each of the predictor variables hereafter.

categories of a specific predictor variable are associated with high or low failure rates, compared with the failure rate for the reference category of the same predictor variable.

For example, inspection station 10 has an average failure rate of 6.7%, which is closest, among the sixteen stations, to the overall average failure rate (6.2%) for the entire data set. Ignoring other predictors for the moment, site 10 represents, in some sense, an average station and can be approximately regarded as normal or typical site.[2] This normal site (site 10) was therefore chosen as a reference category, or control, and deleted from the regression model. Each of the other fifteen stations was included in the regression model. Model results allow an examination of whether the influence of each of these other fifteen stations on the dependent variable (probability of failure in an emissions test) are significantly different from the effect of site 10. The same rationale was used in selecting reference categories for other predictor variables: MODEL, OWNERSHIP. Because MODEL is strongly correlated with CLASS and DOMESTIC, the latter two variables were not used in the multivariable regression models.

We treated the variable MY in a different way. For MY, we were interested to see how earlier model year vehicles behaved in emissions tests compared with the 2002 model year vehicles. Therefore, we picked MY=2002 as the reference category rather than using the value of MY with the closest to the overall average failure rate. This reference category happens to have the largest sample size.[3]

Based on the above reasoning, MODEL=2 (M2), MY=2002, SITE=10 (site 10), OWNERSHIP=3 (individual ownership) were chosen as reference categories and were deleted from the regression models. Table S5 shows the information for the selected reference categories. Except for the variable SITE, all selected reference categories for the predictor variables have sample sizes larger than average for the other categories for each predictor variable.

---

[2] “Approximately” is used here because the average failure rate for the category is obtained without controlling for the influences of other predictor variables.

[3] If there is not a category that seems to have a normal condition, then the category with largest number of observations is usually picked as the one to eliminate.

**TABLE S5. Reference categories of predictor variables**

| Predictor variable* | Total number of levels within predictor variable | Selected reference category | Average failure rate for the reference category | Sample size of the reference category (percentage of the total) |
|---|---|---|---|---|
| **MODEL** | **52** | **MODEL=2** | **6.35%** | **13.4%** |
| **SITE** | **16** | **SITE=10** | **6.67%** | **3.3%** |
| **OWNERSHIP** | **3** | **OWNERSHIP=3** | **4.8%** | **74.3%** |
| **MY** | **4** | **MY=2002** | **2.3%** | **47.9%** |

Note: * — reference values for the binary predictor variable (i.e., SUBURBAN) are used directly in the model and are not listed in the table.

Results from the main effects model highlighted the significance of the variable called "MODEL." The fourteen worst performing vehicle models, defined as the fourteen models with the highest failure probabilities, were identified by inspection of results for the main effects model (see Table S6).

**TABLE S6. Coefficients in the Main Effects Model[a]**

| Variable | Coef.[b] | Variable | Coef. | Variable | Coef. |
|---|---|---|---|---|---|
| MY=2002 Ref. | | Model 19 | 0.43* | Model 45 | 2.29* |
| MY=2001 | 0.80* | Model 20 | 2.36* | Model 46 | -0.26 |
| MY=2000 | 2.38* | Model 21 | 0.16 | Model 47 | 0.54* |
| MY=1999 | 2.29* | Model 22 | 1.35* | Model 48 | 0.39 |
| Owner=1 | 0.05 | Model 23 | 1.36* | Model 49 | 1.74* |
| Owner=2 | 0.41* | Model 24 | 0.96* | Model 50 | 1.05* |
| Owner=3 Ref. | | Model 25 | 1.36* | Model 51 | 1.16* |
| Suburban | -0.11* | Model 26 | 0.96* | Model 52 | 0.72* |
| Model 1 | -0.36* | Model 27 | 3.32* | Site 1 | -0.35* |
| Model 2 Ref. | | Model 28 | -0.37 | Site 2 | 0.33* |
| Model 3 | 0.48* | Model 29 | 0.07 | Site 3 | 0.61* |
| Model 4 | -0.11 | Model 30 | 3.16* | Site 4 | -0.20* |
| Model 5 | -0.54* | Model 31 | 3.48* | Site 5 | 0.63* |
| Model 6 | 0.69* | Model 32 | 1.84* | Site 6 | -0.45* |
| Model 7 | 1.27* | Model 33 | 0.51* | Site 7 | -0.71* |
| Model 8 | 0.46* | Model 34 | 1.25* | Site 8 | 0.11 |
| Model 9 | 0.75* | Model 35 | 0.23 | Site 9 | -0.31* |
| Model 10 | -0.11 | Model 36 | 2.03* | Site 10 Ref. | |
| Model 11 | 1.39* | Model 37 | 1.36* | Site 11 | -0.21* |
| Model 12 | 0.55* | Model 38 | 1.00* | Site 12 | 0.26* |
| Model 13 | 0.54* | Model 39 | -0.49 | Site 13 | -0.03 |
| Model 14 | 1.27* | Model 40 | 2.36* | Site 14 | 0.16 |
| Model 15 | -0.52* | Model 41 | 2.66* | Site 15 | 0.1 |
| Model 16 | 0.68* | Model 42 | 1.72* | Site 16 | -0.73* |
| Model 17 | 0.90* | Model 43 | 1.80* | | |
| Model 18 | -0.38 | Model 44 | 0.84* | | |

[a]. Based on 129,604 observations; log likelihood = -28586; Likelihood ratio=13108.

[b]. Significance level: *--0.05 level.

Having vehicles manufactured by multinational corporations that perform relatively poorly on emissions tests is not unique to China. Consider examples provided by Stewart et al. (*3*). They found Japanese and European vehicle models (with model years from 1998 to 2002) tested in British Columbia that had very high failure rates: the 1998 Acura 1.6 EL had the highest average failure rate (41%), and the 2000 Audi TT had the second highest average failure rate (37%) based on emissions tests conducted in British Columbia in 2004 (*3*). In comparison, the comparable failure rates overall (in

2004) ranged from only 0.24% for model year 2002 to 3.65% for model year 1998. Although these results are striking, they are not definitive because other factors, such as inspection station location and vehicle odometer readings, may explain part of the high failure rates.

Stewart's presentation shows that the 1998 Acura 1.6 EL had the highest average failure rate (41%), about 11 times the average failure rate for all vehicles with model year 1998, and the 2000 Audi TT had the second highest average failure rate (37%), about 29 times the average failure rate for all vehicles with model year 2000, based on emissions tests conducted in British Columbia in 2004.[4]

In a study of I/M test data for the US, Washburn et al. (*4*) shows that GM vehicles had the highest average failure rate, approximately 2.5%, about double the overall average failure rate for all vehicles (~1.3-1.4%). The paper does not give the numbers directly. We estimated the total numbers of vehicles tested using Figure 1 in the Washburn paper and estimated the failure rates using Figure 2. Then we calculated a vehicle number weighted failure rate, which is about 1.37%. GM's failure rate is about double the overall average.

The main reason that the magnitude of the differences in failure rates are so different in the US, Canada and China is that the testing methods for all the major results are different. The Chinese data were based on the ASM testing method, which has two parts: high load with low speed and low load with high speed, both under steady state (no acceleration and deceleration). The Washburn paper analyzes data from a two-part idle and loaded dynamometer testing method. The loaded dynamometer part is somewhat similar to the ASM testing method. The British Columbia of Canada uses IM240, a more advanced and sophisticated transient emissions testing method. It includes accelerations, decelerations, and cruise conditions.

---

[4] Canada does not have a mandatory vehicle emission recall, but, as noted at the Environment Canada website, http://www.ec.gc.ca/CEPARegistry/documents/part/LEV_MOU.cfm: "The motor vehicle manufacturers intend to continue their voluntary practice of implementing emission recalls on vehicles sold in Canada, where a U.S. emission recall is initiated on the equivalent model, and where the reasons

## Final Effects Model Results

Due to the large number of interaction terms between MODEL and MY, the results for these two variables are shown separately from all other modeling results. Table S7 shows model results for all variables other than MODEL and MY. MODEL and MY were used to form a series of interaction terms, and the regression results are presented in Table S8.

for that recall are valid in Canada." This assertion is confirmed at a website for Ford Canada: http://www.ford/ca/english/Maintaining/Recalls/Default.asp.

**TABLE S7. Estimated Probabilities of Failure for Variables other than MODEL and MY**

| Variable | Probability of failure[a] | P>\|z\| | Prob. [95% Conf. Interval] | |
|---|---|---|---|---|
| Gov't | 0.05 | 0.367 | 0.05 | 0.06 |
| Business | 0.07* | 0 | 0.07 | 0.07 |
| Individual | 0.05 | Reference ownership (excluded from the regression) | | |
| Suburban | 0.06* | 0.022 | 0.05 | 0.06 |
| Urban | 0.06 | Reference (excluded from the regression) | | |
| Site 1 | 0.05* | 0 | 0.04 | 0.06 |
| Site 2 | 0.09* | 0 | 0.08 | 0.11 |
| Site 3 | 0.12* | 0 | 0.10 | 0.13 |
| Site 4 | 0.06* | 0.007 | 0.05 | 0.06 |
| Site 5 | 0.12* | 0 | 0.10 | 0.14 |
| Site 6 | 0.04* | 0 | 0.04 | 0.05 |
| Site 7 | 0.03* | 0 | 0.03 | 0.04 |
| Site 8 | 0.07 | 0.204 | 0.06 | 0.09 |
| Site 9 | 0.05* | 0 | 0.04 | 0.06 |
| Site 10 | 0.07 | Reference site (excluded from the regression) | | |
| Site 11 | 0.05* | 0.023 | 0.05 | 0.06 |
| Site 12 | 0.09* | 0.002 | 0.07 | 0.10 |
| Site 13 | 0.06 | 0.744 | 0.05 | 0.08 |
| Site 14 | 0.08 | 0.105 | 0.06 | 0.09 |
| Site 15 | 0.07 | 0.242 | 0.06 | 0.09 |
| Site 16 | 0.03* | 0 | 0.03 | 0.04 |

[a]. The probabilities of failure are the relative probabilities compared with the reference level of each variable except those for the reference categories.

[b]. Significance level: *--0.05 level.

In addition to the reference term (M2 at MY=2002) that was dropped before the modeling, there are three other types of terms that were not included in the regression modeling processes. These are variables with a sample size of zero (e.g., M52 at MY=1999), variables that are collinear with other variables (e.g., M37, M51, etc., at MY=1999), and variables in which all vehicles either passed or failed the tests, i.e., "perfect prediction for pass" (e.g., M13 at MY=1999) and "perfect prediction for failure" (e.g., M7 at MY=1999). Each of these three types of terms is marked as 0. Furthermore, the probabilities of failure without confidence intervals for the third type of variables are given in the table although they are indicated as perfect predictions.

In sum, there are 22 terms with no vehicles, and six other terms have collinearity problems; there are 17 terms that perfectly predict passes and seven terms that perfectly predict failures. Each of these 52 terms was dropped from the regression model.

**TABLE S8. Logistic regression results for MODEL and MY**

| | MY=2002 | | | | MY=2001 | | | | MY=2000 | | | | MY=1999 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Vehicle Model | Prob. of Failure | P>\|z\| | 95% C.I. | | Prob. of Failure | P>\|z\| | 95% C.I. | | Prob. of Failure | P>\|z\| | 95% C.I. | | Prob. of Failure | P>\|z\| | 95% C.I. | |
| Model 1 | 0.01 | 0.00 | 0.01 | 0.02 | 0.02 | 0.86 | 0.02 | 0.03 | 0.07 | 0.00 | 0.05 | 0.08 | 0.07 | 0.00 | 0.06 | 0.09 |
| Model 2 | 0.02 | Reference (not included in the regression model) | | | 0.03 | 0.02 | 0.02 | 0.04 | 0.11 | 0.00 | 0.09 | 0.14 | 0.09 | 0.00 | 0.07 | 0.11 |
| Model 3 | 0.02 | 0.33 | 0.01 | 0.02 | 0.07 | 0.00 | 0.05 | 0.09 | 0.13 | 0.00 | 0.10 | 0.16 | 0.20 | 0.00 | 0.15 | 0.25 |
| Model 4 | 0.01 | 0.00 | 0.01 | 0.01 | 0.03 | 0.19 | 0.02 | 0.04 | 0.08 | 0.00 | 0.06 | 0.11 | 0.11 | 0.00 | 0.08 | 0.16 |
| Model 5 | 0.01 | 0.00 | 0.01 | 0.01 | 0.01 | 0.05 | 0.01 | 0.02 | 0.08 | 0.00 | 0.06 | 0.11 | 0.05 | 0.00 | 0.04 | 0.07 |
| Model 6 | 0.02 | 0.18 | 0.01 | 0.02 | 0.05 | 0.00 | 0.03 | 0.06 | 0.21 | 0.00 | 0.18 | 0.26 | | No vehicles | | |
| Model 7 | 0.03 | 0.00 | 0.03 | 0.05 | 0.10 | 0.00 | 0.08 | 0.12 | 0.25 | 0.00 | 0.20 | 0.31 | 1.00 | Perfectly predict failure (not included in the regression model) | | |
| Model 8 | 0.01 | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.02 | 0.04 | 0.20 | 0.00 | 0.16 | 0.24 | 0.58 | 0.00 | 0.34 | 0.78 |
| Model 9 | 0.02 | 0.31 | 0.01 | 0.03 | 0.04 | 0.00 | 0.03 | 0.06 | 0.24 | 0.00 | 0.20 | 0.29 | 0.20 | 0.00 | 0.15 | 0.26 |
| Model 10 | 0.01 | 0.01 | 0.00 | 0.02 | 0.02 | 0.36 | 0.01 | 0.03 | 0.09 | 0.00 | 0.07 | 0.12 | 0.17 | 0.00 | 0.12 | 0.23 |
| Model 11 | 0.05 | 0.00 | 0.04 | 0.06 | 0.08 | 0.00 | 0.05 | 0.11 | 0.30 | 0.00 | 0.19 | 0.44 | 0.62 | 0.00 | 0.25 | 0.89 |
| Model 12 | 0.02 | 0.86 | 0.01 | 0.03 | 0.00 | Perfectly predict pass (not included in the regression model) | | | 0.00 | Perfectly predict pass (not included in the regression model) | | | 0.61 | 0.00 | 0.08 | 0.96 |
| Model 13 | 0.03 | 0.16 | 0.02 | 0.04 | 0.02 | 0.49 | 0.01 | 0.03 | 0.00 | Perfectly predict pass (not included in the regression model) | | | 0.00 | Perfectly predict pass (not included in the regression model) | | |
| Model 14 | 0.04 | 0.00 | 0.03 | 0.06 | 0.10 | 0.00 | 0.07 | 0.13 | 0.29 | 0.00 | 0.23 | 0.36 | 0.29 | 0.00 | 0.23 | 0.36 |
| Model 15 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | 0.02 | 0.00 | 0.01 | 0.28 | 0.00 | 0.17 | 0.44 | 1.00 | Perfectly predict failure (not included in the regression model) | | |
| Model 16 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | Perfectly predict pass (not included in the regression model) | | | 0.57 | 0.00 | 0.44 | 0.70 | 0.46 | 0.00 | 0.21 | 0.73 |
| Model 17 | 0.03 | 0.08 | 0.02 | 0.05 | 0.05 | 0.00 | 0.03 | 0.09 | 0.23 | 0.00 | 0.16 | 0.33 | | No vehicles | | |
| Model 18 | 0.01 | 0.01 | 0.00 | 0.02 | 0.00 | Perfectly predict pass (not included in the regression model) | | | 0.00 | Perfectly predict pass (not included in the regression model) | | | | No vehicles | | |
| Model 19 | 0.03 | 0.62 | 0.01 | 0.05 | 0.01 | 0.29 | 0.00 | 0.04 | 0.27 | 0.00 | 0.21 | 0.34 | 0.06 | 0.00 | 0.04 | 0.10 |
| Model 20 | 0.11 | 0.00 | 0.08 | 0.15 | 0.19 | 0.00 | 0.14 | 0.24 | 0.66 | 0.00 | 0.56 | 0.74 | | No vehicles | | |
| Model 21 | 0.01 | 0.22 | 0.01 | 0.03 | 0.00 | Perfectly predict pass (not included in the regression model) | | | 0.00 | Perfectly predict failure (not included in the regression model) | | | | No vehicles | | |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model 22 | 0.05 | 0.00 | 0.03 | 0.08 | 0.12 | 0.00 | 0.08 | 0.18 | 0.30 | 0.00 | 0.23 | 0.37 | | No vehicles | | |
| Model 23 | 0.01 | 0.03 | 0.00 | 0.02 | 0.03 | 0.44 | 0.01 | 0.08 | 0.57 | 0.00 | 0.45 | 0.68 | 0.45 | 0.00 | 0.30 | 0.61 |
| Model 24 | 0.02 | 0.95 | 0.01 | 0.05 | 0.02 | 0.50 | 0.01 | 0.04 | 0.33 | 0.00 | 0.26 | 0.42 | 0.23 | 0.00 | 0.13 | 0.36 |
| Model 25 | 0.17 | 0.00 | 0.13 | 0.22 | 0.51 | 0.00 | 0.42 | 0.60 | 0.96 | 0.00 | 0.89 | 0.99 | | No vehicles | | |
| Model 26 | 0.01 | 0.05 | 0.00 | 0.02 | 0.03 | 0.61 | 0.01 | 0.06 | 0.00 | Perfectly predict pass (not included in the regression model) | | | | No vehicles | | |
| Model 27 | 0.02 | 0.87 | 0.00 | 0.12 | 0.03 | 0.39 | 0.02 | 0.05 | 0.00 | Perfectly predict pass (not included in the regression model) | | | | No vehicles | | |
| Model 28 | 0.42 | 0.00 | 0.34 | 0.51 | 0.35 | 0.00 | 0.28 | 0.44 | 0.53 | 0.00 | 0.44 | 0.62 | | No vehicles | | |
| Model 29 | 0.46 | 0.00 | 0.37 | 0.55 | 0.41 | 0.00 | 0.33 | 0.51 | 0.66 | 0.00 | 0.58 | 0.74 | | No vehicles | | |
| Model 30 | 0.03 | 0.07 | 0.02 | 0.06 | 0.15 | 0.00 | 0.06 | 0.35 | 1.00 | Perfectly predict failure (not included in the regression model) | | | 0.74 | 0.00 | 0.23 | 0.97 |
| Model 31 | 0.02 | 0.80 | 0.01 | 0.05 | 0.01 | 0.38 | 0.00 | 0.06 | 0.23 | 0.00 | 0.14 | 0.34 | 0.14 | 0.00 | 0.06 | 0.29 |
| Model 32 | 0.00 | 0.08 | 0.00 | 0.03 | 0.00 | Perfectly predict pass (not included in the regression model) | | | 0.34 | 0.00 | 0.24 | 0.44 | 0.50 | 0.00 | 0.36 | 0.65 |
| Model 33 | 0.03 | 0.55 | 0.01 | 0.05 | 0.04 | 0.30 | 0.01 | 0.16 | 0.15 | 0.00 | 0.10 | 0.23 | | Collinearity (not included in the regression model) | | |
| Model 34 | 0.08 | 0.00 | 0.05 | 0.12 | 0.00 | Perfectly predict pass (not included in the regression model) | | | 1.00 | Perfectly predict failure (not included in the regression model) | | | | No vehicles | | |
| Model 35 | 0.01 | 0.38 | 0.00 | 0.06 | | Collinearity (not included in the regression model) | | | 0.40 | 0.00 | 0.31 | 0.50 | 0.32 | 0.00 | 0.17 | 0.51 |
| Model 36 | 0.02 | 0.96 | 0.01 | 0.05 | 0.10 | 0.00 | 0.06 | 0.16 | 0.10 | 0.01 | 0.03 | 0.27 | | No vehicles | | |
| Model 37 | 0.01 | 0.21 | 0.00 | 0.03 | 0.00 | Perfectly predict pass (not included in the regression model) | | | 0.13 | 0.00 | 0.06 | 0.26 | | Collinearity (not included in the regression model) | | |
| Model 38 | 0.20 | 0.00 | 0.11 | 0.35 | 0.15 | 0.00 | 0.07 | 0.31 | 0.56 | 0.00 | 0.47 | 0.65 | 0.61 | 0.00 | 0.22 | 0.89 |
| Model 39 | 0.11 | 0.00 | 0.06 | 0.19 | 0.37 | 0.00 | 0.27 | 0.48 | 0.57 | 0.00 | 0.45 | 0.69 | | No vehicles | | |
| Model 40 | 0.01 | 0.35 | 0.00 | 0.06 | 0.00 | Perfectly predict pass (not included in the regression model) | | | 0.48 | 0.00 | 0.29 | 0.68 | 0.56 | 0.00 | 0.42 | 0.70 |
| Model 41 | 0.02 | 0.98 | 0.00 | 0.13 | 0.06 | 0.06 | 0.02 | 0.18 | 0.52 | 0.00 | 0.41 | 0.63 | 0.31 | 0.00 | 0.17 | 0.49 |
| Model 42 | 0.19 | 0.00 | 0.09 | 0.37 | 0.00 | Perfectly predict pass (not included in the regression model) | | | | Collinearity (not included in the regression model) | | | 0.00 | Perfectly predict pass (not included in the regression model) | | |
| Model 43 | 0.14 | 0.00 | 0.09 | 0.21 | 0.05 | 0.27 | 0.01 | 0.17 | 1.00 | Perfectly predict failure (not included in the regression model) | | | | No vehicles | | |
| Model 44 | 0.01 | 0.15 | 0.00 | 0.04 | | No vehicles | | | | No vehicles | | | 0.34 | 0.01 | 0.04 | 0.86 |
| Model 45 | 0.12 | 0.00 | 0.04 | 0.28 | 0.04 | 0.31 | 0.01 | 0.16 | 0.18 | 0.00 | 0.11 | 0.29 | 0.07 | 0.04 | 0.02 | 0.20 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model 46 | 0.02 | 0.69 | 0.00 | 0.06 | 0.02 | 0.76 | 0.00 | 0.10 | 0.50 | 0.01 | 0.06 | 0.94 | 1.00 | Perfectly predict failure (not included in the regression model) | | |
| Model 47 | 0.30 | 0.00 | 0.15 | 0.51 | 0.00 | Perfectly predict pass (not included in the regression model) | | | 0.34 | 0.00 | 0.22 | 0.50 | 0.61 | 0.00 | 0.35 | 0.82 |
| Model 48 | 0.03 | 0.42 | 0.01 | 0.08 | | No vehicles | | | | No vehicles | | | | No vehicles | | |
| Model 49 | 0.05 | 0.38 | 0.01 | 0.29 | | Collinearity (not included in the regression model) | | | 0.48 | 0.00 | 0.25 | 0.72 | | No vehicles | | |
| Model 50 | 0.04 | 0.54 | 0.01 | 0.23 | 0.03 | 0.65 | 0.00 | 0.20 | 0.25 | 0.00 | 0.13 | 0.43 | 0.14 | 0.00 | 0.05 | 0.33 |
| Model 51 | 0.11 | 0.00 | 0.04 | 0.27 | 0.04 | 0.48 | 0.01 | 0.25 | 0.65 | 0.00 | 0.44 | 0.81 | | Collinearity (not included in the regression model) | | |
| Model 52 | 0.64 | 0.00 | 0.47 | 0.78 | 0.97 | 0.00 | 0.80 | 1.00 | 0.95 | 0.00 | 0.74 | 0.99 | | No vehicles | | |

Note: * — The probability of failure for the reference model (M2) at MY=2002 is 2.12%. The probabilities of failure of the other models are the relative probabilities compared with the reference model (M2) at MY=2002.

## References

(1) Beijing Environmental Protection Bureau. "Emission Standard for Exhaust Pollutants from Gasoline Vehicles under Steady-State Loaded Mode, DB 11/122-2003," February 2003.

(2) US EPA. "Technical Guidance for Acceleration Simulation Mode Test Procedures, Emission Standards, Quality Control Requirements, and Equipment Specifications, EPA-AA-RSPD-IM-96-2," July 1996.

(3) Stewart, S.; Wong, J. In *Proceedings of the 15th CRC On-road Vehicle Emissions Workshop*: San Diego, California, 2005.

(4) Washburn, S.; Seet, J.; Mannering, F. Statistical modeling of vehicle emissions from I/M testing data: an exploratory analysis; *Transportation Research Part D* **2001**, *11*, 21-36.