1 Structural estimation of fuel consumption

Consider the following parameterization of structural model of fuel consumption. For ship i in hour j, let

$$\frac{W_{ME,ij}}{W_{ME,ref}} = C\tilde{t}_{ij}^{0.66}\tilde{v}_{ij}^{3} \tag{1}$$

with $\tilde{t}_{ij} := \frac{t_{ij}}{t_{ref}}$ and $\tilde{v}_{ij} := \frac{v_{ij}}{v_{ref}}$ and

$$SFC_{ME_{ij}} = SFC_{base} \cdot \left(\beta_1 - 1.6\beta_2 \frac{W_{ME,ij}}{W_{ME,ref}} + \beta_2 \left(\frac{W_{ME,ij}}{W_{ME,ref}}\right)^2\right). \tag{2}$$

where the choice of $-1.6\beta_2$ is motivated by the assumption that $SFC_{ME_{ij}}$ is minimized at fuel efficient point of $\frac{W_{ME,ij}}{W_{ME,ref}} = 0.8$.[1] Here, the unkonwn parameters are C, $\tilde{\theta}_1$, and $\tilde{\theta}_3$.

Then, from (1)-(2),

$$FC_{ME,ij} = W_{ME,ij}SFC_{ME_{ij}} = \theta_1\tilde{t}_{ij}^{0.66}\tilde{v}_{ij}^{3} + \theta_2\left(\tilde{t}_{ij}^{0.66}\tilde{v}_{ij}^{3}\right)^2 + \theta_3\left(\tilde{t}_{ij}^{0.66}\tilde{v}_{ij}^{3}\right)^3$$

where $\theta_1 := \beta_1 C$, $\theta_2 := -1.6\beta_2 C^2$, and $\theta_3 := \beta_2 C^3$.

Because we observe aggregate fuel consumption over the period of one year, we sum up over j's to obtain a specification for annual fuel consumption as

$$\frac{FC_{ME,i}}{W_{ref}SFC_{base}} = \theta_1 \sum_j \tilde{t}_{ij}^{0.66}\tilde{v}_{ij}^{3} + \theta_2 \sum_j \left(\tilde{t}_{ij}^{0.66}\tilde{v}_{ij}^{3}\right)^2 + \theta_3 \sum_j \left(\tilde{t}_{ij}^{0.66}\tilde{v}_{ij}^{3}\right)^3. \tag{3}$$

Let $y_i = \frac{FC_{ME,i}}{W_{ref}SFC_{base}}$, $\boldsymbol{\theta} = (\theta_0, \theta_1, \theta_2)^\top$, and $\mathbf{x}_i = (x_{1i}, x_{2i}, x_{3i})^\top$ with $x_{1i} := \sum_j \tilde{t}_{ij}^{0.66}\tilde{v}_{ij}^{3}$, $x_{2i} := \sum_j \left(\tilde{t}_{ij}^{0.66}\tilde{v}_{ij}^{3}\right)^2$, and $x_{3i} := \sum_j \left(\tilde{t}_{ij}^{0.66}\tilde{v}_{ij}^{3}\right)^3$. Then, we may estimate $\boldsymbol{\theta}$ by

$$\hat{\boldsymbol{\theta}} = \min_{\boldsymbol{\theta}\in\Theta} \sum_{i=1}^{n} (y_i - \mathbf{x}_i^\top \boldsymbol{\theta})^2.$$

We may also extend the above model by specifying so that θ_0, θ_1, and θ_2 depend on $\mathbf{1}\left\{\frac{W_{ME,ij}}{W_{ME,ref}} > 0.8\right\}$ to allow for the asymmetry around the optimal fuel efficiency point. Specifically, we may allow $\tilde{\theta}_1$ and $\tilde{\theta}_3$ to depend on $\mathbf{1}\left\{\frac{W_{ME,ij}}{W_{ME,ref}} > 0.8\right\}$ in Equation (2) so that $\tilde{\theta}_1 = \tilde{\theta}_{10} + \tilde{\theta}_{11}\mathbf{1}\left\{\frac{W_{ME,ij}}{W_{ME,ref}} > 0.8\right\}$ and $\tilde{\theta}_3 = \tilde{\theta}_{30} + \tilde{\theta}_{31}\mathbf{1}\left\{\frac{W_{ME,ij}}{W_{ME,ref}} > 0.8\right\}$. Then, repeating the above argument and

[1] The first order condition of minimizing $f(x) = \theta_1 + \tilde{\theta}_2 x + \tilde{\theta}_3 x^2$ gives $x = -\tilde{\theta}_2/\tilde{\theta}_3$. When $x = 0.8$, $\tilde{\theta}_2 = -1.6\tilde{\theta}_3$.

aggregating over j's gives

$$\frac{FC_{ME,i}}{W_{ref}SFC_{base}} = \theta_1 \cdot \sum_j \frac{W_{ME,ij}}{W_{ME,ref}} + \theta_2 \sum_j \left(\frac{W_{ME,ij}}{W_{ME,ref}}\right)^2 + \theta_3 \cdot \sum_j \left(\frac{W_{ME,ij}}{W_{ME,ref}}\right)^3$$
$$+ \theta_4 \cdot \sum_j \mathbf{1}\left\{\frac{W_{ME,ij}}{W_{ME,ref}} > 0.8\right\} \frac{W_{ME,ij}}{W_{ME,ref}} + \theta_5 \sum_j \mathbf{1}\left\{\frac{W_{ME,ij}}{W_{ME,ref}} > 0.8\right\} \left(\frac{W_{ME,ij}}{W_{ME,ref}}\right)^2$$
$$+ \theta_6 \cdot \sum_j \mathbf{1}\left\{\frac{W_{ME,ij}}{W_{ME,ref}} > 0.8\right\} \left(\frac{W_{ME,ij}}{W_{ME,ref}}\right)^3.$$

We may test the assumption of symmetry by testing $\theta_4 = \theta_5 = \theta_6 = 0$.

Furthermore, we may specify θ_0, θ_1, and θ_2 in terms of observed ship characteristics.

2 Extended structural model

As an extension, we treat the exponents of $\tilde{t}_{ij}$ and $\tilde{v}_{ij}$ as unknown parameters as:

$$\frac{W_{ME,ij}}{W_{ME,ref}} = C\tilde{t}_{ij}^{\alpha_1}\tilde{v}_{ij}^{\alpha_2}. \tag{4}$$

Then, because the annual fuel consumption is determined as $FC_{ME,i} = \sum_j SFC_{ME_{ij}} W_{ME,ij}$, we have

$$\frac{FC_{ME,i}}{W_{ref}SFC_{base}} = \theta_1 \sum_j \tilde{t}_{ij}^{\alpha_1}\tilde{v}_{ij}^{\alpha_2} + \theta_2 \sum_j \left(\tilde{t}_{ij}^{\alpha_1}\tilde{v}_{ij}^{\alpha_2}\right)^2 + \theta_3 \sum_j \left(\tilde{t}_{ij}^{\alpha_1}\tilde{v}_{ij}^{\alpha_2}\right)^3, \tag{5}$$

where $\theta_1 = C\beta_1$, $\theta_2 = -1.6C^2\beta_2$, and $\theta_3 = C^3\beta_2$.

We also consider a more flexible specification of $\frac{W_{ME,ij}}{W_{ME,ref}}$ given by

$$\frac{W_{ME,ij}}{W_{ME,ref}} = C\psi_1(\tilde{t}_{ij};\boldsymbol{\gamma}_1)\psi_2(v_{ij};\boldsymbol{\gamma}_2)$$

with

$$\psi_1(\tilde{t}_{ij};\boldsymbol{\gamma}_1) = \sum_{k=0}^{K} \gamma_{1k} B_k\left(\tilde{t}_{ij}\right) \quad \text{and}$$
$$\psi_2(\tilde{v}_{ij};\boldsymbol{\gamma}_2) = \sum_{k=0}^{K} \gamma_{2k} B_k\left(\tilde{v}_{ij}\right).$$

In this case, we have

$$\frac{FC_{ME,i}}{W_{ref}SFC_{base}} = \beta_1 \underbrace{\sum_j \psi_1(\tilde{t}_{ij};\boldsymbol{\gamma}_1)\psi_2(v_{ij};\boldsymbol{\gamma}_2)}_{:=\Psi_{1i}(\boldsymbol{\gamma})}$$
$$+ \beta_2 \underbrace{\sum_j \left(-1.6\left(\psi_1(\tilde{t}_{ij};\boldsymbol{\gamma}_1)\psi_2(v_{ij};\boldsymbol{\gamma}_2)\right)^2 + \left(\psi_1(\tilde{t}_{ij};\boldsymbol{\gamma}_1)\psi_2(v_{ij};\boldsymbol{\gamma}_2)\right)^3\right)}_{:=\Psi_{2i}(\boldsymbol{\gamma})}. \tag{6}$$

3 Framework

Consider a random sample of n observations $S = \{(X_i, Y_i)\}_{i=1}^n$, where $(X_i, Y_i) \overset{iid}{\sim} F(x, y)$. We assume that the data is generated as

$$Y_i = m(X_i) + \epsilon_i, \quad \epsilon_i | X_i \overset{iid}{\sim} F_\epsilon$$

form some $m(x)$.

We

4 Structural model vs. non-parametric model

We are interested in evaluating the predictive performance of different models when we use the test set in which the support of covariates does not overlap with that in the training set.

Let $\{Y_i, X_i\}$ be a sample of size n independently drawn from $F(y, x)$. We are interested in predicting the mean value of Y when $X = x_0$, where x_0 is located outside of the support of $F(y, x)$. Let $f(x)$ be the probability density function X.

The predictive ability of estimated models depends on how far away the location of x_0 is from the distribution of X_i in the effective sample used in the estimation of predictive models.

Let $\rho : \mathcal{P} \times \mathcal{X} \to \mathbb{R}$ be a distance measure between a distribution function and a point. For example, when we use the Euclidean distance between the mean of the distribution and the point as a measure, $\rho(f(x), x_0) = || \int x f(x) dx - x_0 ||$.

Example 1 (Local linear regression) *Given x_0, we construct a predictive model based on local polynomial regression at x_d given a bandwidth h as:*

$$\hat{\theta}(x_d, h) = \arg\min_\theta \sum_{i=1}^n w_h(|X_i - x_d|)(Y_i - X_i^\top \theta)^2.$$

In this case, the "effective" sample distribution that is used for estimating this local polynomial regression is given by

$$g_{h,x_d}(x) = \frac{f(x) \times w_h(|x - x_d|)}{\int f(x) \times w_h(|x - x_d|) dx}.$$

Then, we define the distance between the effective sampling distribution g_{h,x_d} and x_0 by $\rho(g_{h,x_d}, x_0)$.

Example 2 (Structural models) *The structural models (3)-(6) are estimated using the sample distribution $f(x)$. Therefore, the distance between the sample distribution and the evaluation point x_0 is $\rho(f(x), x_0)$.*

We expect that the bias is an increasing function of the distance between the effective sampling distribution and the evaluation point x_0. On the other hand, the variance is a decreasing function of the effective sample size.

For each predictive model, we will illustrate how the bias and the variance depends on the distance between the effective sampling distribution and the evaluation point as well as the effective sample size.

Suppose we want to evaluate the predictive performance at x_0.