SFB 823

A focused information criterion for quantile regression: Evidence for the rebound effect

Peter Behl, Holger Dette, Manuel Frondel, Colin Vance

Nr. 39/2016

Discussion Paper

A Focused Information Criterion for Quantile Regression: Evidence for the Rebound Effect.

Peter Behl, **Holger Dette**, Ruhr University Bochum (RUB)

Manuel Frondel, RWI – Leibniz Institut für Wirtschaftsforschung and Ruhr University Bochum (RUB)

Colin Vance, RWI – Leibniz Institut für Wirtschaftsforschung and Jacobs University Bremen

Abstract. In contrast to conventional model selection criteria, the Focused Information Criterion (FIC) allows for the purpose-specific choice of model specifications. This accommodates the idea that one kind of model might be highly appropriate for inferences on a particular focus parameter, but not for another. Using the FIC concept that is developed by Behl, Claeskens and Dette (2014) for quantile regression analysis, and the estimation of the rebound effect in individual mobility behavior as an example, this paper provides for an empirical application of the FIC in the selection of quantile regression models.

JEL classification: C3, D2.

Key words: Information Criteria, Fuel Efficiency, Price Elasticities.

Correspondence: Manuel Frondel, RWI – Leibniz Institut für Wirtschaftsforschung, Hohenzollernstr. 1-3, D-45128 Essen. frondel@rwi-essen.de.

Acknowledgements: This work has been supported by the Collaborative Research Center "Statistical Modelling of Nonlinear Dynamic Processes" (SFB 823) of the German Research Foundation (DFG), within the framework of project A3, "Dynamic Technology Modelling", and project C1, "Model Choice and Dynamic Dependence Structures". The work of Peter Behl has been supported by a doctoral scholarship of the Hanns-Seidel-Foundation.

1 Introduction

Common model selection methods, such as the AKAIKE (1974) criterion (AIC) and the SCHWARZ (1978) criterion (SIC), do not require the specification of any purpose of inference, that is, a focus parameter. This also holds true for alternative model selection methods, such as goodness-of-fit tests, which are proposed by, among many others, DETTE (1999), DETTE, PODOLSKIJ and VETTER (2006), and PODOLSKIJ and DETTE (2008). However, conditional on the underlying purpose, some specifications might be better suited than others in terms of estimation efficiency with respect to the focus parameter. Recognizing this argument, CLAESKENS and HJORT (2003) designed the Focused Information Criterion (FIC) for the targeted search of parametric regression models that are estimated using maximum-likelihood methods, thereby explicitly taking the purpose of inference into account (BEHL, CLAESKENS, DETTE, 2014:602).

This is of high relevance in many fields of applied research, such as estimating the well-known direct rebound effect, which captures the behaviorally induced offset in the reduction of energy consumption following efficiency improvements (e. g. SORRELL, DIMITROUPOULOS, 2008; FRONDEL, PETERS, VANCE, 2008). To this end, alternative focus parameters are estimated in the context of individual transportation: First, the efficiency elasticity of mobility demand s:

$$\eta_{\mu}(s) := \frac{\partial \ln s}{\partial \ln \mu}\,, \tag{1}$$

reflecting the relative change in mobility demand s due to a percentage increase in efficiency μ (see e. g. BERKHOUT *et al.*, 2000),[1] and, second, the negative of the

[1] In line with the economic literature (e. g. FRONDEL, VANCE, 2013), energy efficiency is defined here by

$$\mu = \frac{s}{e} > 0,$$

where the efficiency parameter μ characterizes the technology with which a service demand s is satisfied and e denotes the energy input employed for a service such as mobility. For the specific example of individual conveyance, parameter μ designates fuel efficiency, which can be measured in terms of vehicle kilometers per liter of fuel input. The efficiency definition reflects the fact that fuel price elasticity of mobility demand, $\eta_{p_e}(s)$:

$$\eta_{p_e}(s) := \frac{\partial \ln s}{\partial \ln p_e} \, . \tag{2}$$

While $\eta_\mu(s)$ is the most natural definition of the direct rebound effect, the negative of the fuel price elasticity $\eta_{p_e}(s)$ is frequently the preferred measure for various reasons (FRONDEL, RITTER, VANCE, 2012), most notably because of the likely endogeneity of efficiency variable μ. For instance, if a more efficient car is purchased in response to a job change that results in a longer commute, fuel efficiency would not be exogenous (see e. g. SORRELL, DIMITROUPOULOS, SOMMERVILLE, 2009:1361). To avoid endogeneity bias, it would be wise to refrain from including this variable in any model specification aiming at estimating the response to fuel price effects, as fuel efficiency may be a bad control (ANGRIST and PISCHKE, 2009:63).

Using the FIC developed by BEHL, CLAESKENS and DETTE (2014) for quantile regression analysis and building on FRONDEL, RITTER and VANCE (2012), who investigate the heterogeneity of the rebound effect in individual mobility behavior on the basis of quantile regressions, this paper provides for an empirical application of the FIC in the selection of quantile regression models.[2] It will become evident from our empirical illustration that model selection may depend on the percentiles of the dependent variable under scrutiny.

Because of its usefulness in balancing modeling bias against estimation variability, the FIC has been increasingly applied in the realm of statistics (see e. g. CLAESKENS, CROUX, VAN KERCKHOVEN, 2007, CLAESKENS, HJORT, 2008, and HJORT, CLAESKENS, 2006), but this concept appears to be virtually unknown

the higher the efficiency μ of a given technology, the less energy $e = s/\mu$ is required for the provision of a service. The above efficiency definition assumes proportionality between service level and energy input regardless of the level – a simplifying assumption that may not be true in general, but provides for a convenient first-order approximation of the relationship of s with respect to e.

[2]R code is available from the authors upon request.

in the economics literature, particularly in transport economics. The contributions of BEHL et al. (2012, 2013) represent the sole exceptions for the literature on economic modeling, while the analysis of BROWNLEES and GALLO (2008) is a rare example originating from financial economics.

The general idea underlying the FIC, which ultimately results from estimating the mean squared error of the modeling bias (CLAESKENS, HJORT, 2003:902), is to study perturbations of a parametric model, with the known parameter vector $\gamma^0 := (\gamma_1^0, ..., \gamma_q^0)^T$ as the point of departure. A variety of models may then be considered that depart from γ^0 in some or all of q directions: $\gamma \neq \gamma^0$. On the basis of parameter estimates of the altogether 2^q (sub-)models that candidate model will be selected for which the FIC is minimal for a given focus parameter $\Lambda = \Lambda(\gamma)$.

By minimizing the FIC, one captures the trade-off between modeling bias, which, by definition, is zero for the most general model for which $\gamma_i \neq \gamma_i^0$ for $i = 1, ..., q$, and relative estimation variability, which, by definition, is zero for the most restricted model for which $\gamma_i = \gamma_i^0$ for $i = 1, ..., q$. For the sake of simplicity, in our empirical example on how to estimate the direct rebound effect, we will confine ourselves to $q = 1$. That is, we choose between just $2^q = 2$ model specifications, where the unrestricted specification includes the critical variable fuel efficiency μ, while the restricted specification does not.

The following Section 2 provides for a concise introduction into the concept of the FIC. Section 3 presents the regression method, followed by the presentation of the empirical example in Section 4. The last section summarizes and concludes.

2 The Example of the Rebound Effect

To illustrate the concept of the FIC with the empirical example of the heterogeneity in individual mobility behavior, we follow FRONDEL, RITTER, and VANCE

(2012) and choose the negative of the fuel price elasticity of transport demand, $-\eta_{p_e}(s)$, for the identification of the rebound effect, although theory would suggest estimating the efficiency elasticity $\eta_\mu(s)$ to directly capture the rebound. In line with FRONDEL, RITTER, and VANCE (2012), however, we argue that the indirect way to elicit the rebound effect via estimating fuel price elasticities is empirically advantageous, as fuel prices typically exhibit sufficient variation and, in contrast to fuel efficiency, can be regarded as parameters that are largely exogenous to individual households. In short, our preferred focus parameter Λ for the empirical identification of the direct rebound effect is given by $\Lambda := -\eta_{p_e}(s)$.

To capture heterogeneity in the rebound response, we estimate the conditional quantile function (CQF) of the logged monthly vehicle kilometers traveled, $\log(s)$, for a given percentile $\tau \in (0,1)$, using quantile regression methods developed by KOENKER and BASSETT (1978):

$$
\begin{aligned}
Q_\tau(\log(s_i)|p_{e_i}, \mathbf{z}_i) &= \alpha(\tau) + \alpha_{p_e}(\tau)\log(p_{e_i}) + \mathbf{z}_i^T \boldsymbol{\alpha}_{\mathbf{z}}(\tau) \\
&= (\boldsymbol{\alpha}^0(\tau))^T \mathbf{x}_i \,, \qquad (3)
\end{aligned}
$$

where $\log(p_e)$ designates logged fuel prices, $\log(\mu)$ denotes logged efficiency and $\mathbf{x}_i := (1, \log(p_{e_i}), \mathbf{z}_i, \log(\mu_i))^T$, with T indicating the transposition of a vector. $\mathbf{z}$ is a vector of control variables, such as household income, employment status of adult household members and number of children, and $\boldsymbol{\alpha}^0(\tau)$ is defined by $\boldsymbol{\alpha}^0(\tau) := (\boldsymbol{\xi}(\tau), \gamma^0)^T$ with $\boldsymbol{\xi}(\tau) := (\alpha(\tau), \alpha_{p_e}(\tau), \boldsymbol{\alpha}_{\mathbf{z}}(\tau))^T$. As efficiency μ is not included as a regressor in model (3), $\gamma^0 = \alpha_\mu = 0$.

Instead from specification (3), where efficiency μ is omitted, the rebound effect is frequently estimated from a wider model that includes the likely endogenous efficiency variable μ:

$$
\begin{aligned}
Q_\tau(\log(s_i)|p_{e_i}, \mu_i, \mathbf{z}_i) &= \alpha(\tau) + \alpha_{p_e}(\tau)\log(p_{e_i}) + \boldsymbol{\alpha}_{\mathbf{z}}^T(\tau)\mathbf{z}_i + \alpha_\mu(\tau)\log(\mu_i) \\
&= (\boldsymbol{\alpha}^{full})^T \mathbf{x}_i \,, \qquad (4)
\end{aligned}
$$

where $\boldsymbol{\alpha}^{full} := (\boldsymbol{\xi}(\tau), \gamma(\tau))^T$ and $\gamma(\tau) := \alpha_\mu(\tau)$.

Adopting the terminology of CLAESKENS and HJORT (2003), specification (3) is called the narrow or null model, as efficiency variable μ is lacking, whereas it is included in the full model (4). Using the terminology introduced in the previous section, by estimating the full model, we depart from $\alpha_\mu = \gamma^0 = 0$ in just $q = 1$ direction: $\gamma(\tau) = \alpha_\mu(\tau) \neq 0$.

In our example, in which model (3) is nested in specification (4) and we choose between only these two models, the quantile regression formulae for the FIC adopt a straightforward shape that strongly resembles those for specifications that are estimated using maximum-likelihood methods (see e. g. BEHL et al., 2013). Following BEHL, CLAESKENS, and DETTE (2014), the FIC for the null model is given by

$$\mathrm{FIC}^0 := \boldsymbol{\omega}_I^T \mathbf{B}\mathbf{B}^T \boldsymbol{\omega}_I, \tag{5}$$

where for $q = 1$, as in our case, bias vector $\mathbf{B} := \sqrt{n}(\gamma - \gamma^0)$ degenerates to a scalar: $B = \sqrt{n}\gamma(\tau)$.

The FIC depends on focus parameter Λ via vector $\boldsymbol{\omega}_I$, which is defined by

$$\boldsymbol{\omega}_I := I_{10} I_{00}^{-1} \frac{\partial \Lambda}{\partial \boldsymbol{\xi}} - \frac{\partial \Lambda}{\partial \gamma} = I_{10} I_{00}^{-1} \frac{\partial \Lambda}{\partial \boldsymbol{\xi}} - \frac{\partial \Lambda}{\partial \alpha_\mu},$$

and also simplifies to a scalar, as our focus parameter Λ is given by

$$\Lambda(\boldsymbol{\alpha}^{full}(\tau)) = -\alpha_{p_e}(\tau) = -\eta_{p_e}(s). \tag{6}$$

Hence, $\frac{\partial \Lambda}{\partial \alpha_\mu} = 0$ and

$$\frac{\partial \Lambda}{\partial \boldsymbol{\xi}} = (0, -1, \mathbf{0})^T. \tag{7}$$

From the definition of $\boldsymbol{\omega}_I$ and derivative (7), it follows that $\boldsymbol{\omega}_I$ equals the negative of the second element of matrix $I_{10} I_{00}^{-1}$, with I_{10} and I_{00} belonging to the information matrix

$$\mathbf{I} := \begin{pmatrix} I_{00} & I_{01} \\ I_{10} & I_{11} \end{pmatrix}, \tag{8}$$

whose components are defined as follows:

$$\begin{aligned} I_{00} &:= \lim_{n\to\infty} \tfrac{1}{n} \textstyle\sum_{i=1}^{n} f_i(\varepsilon_i|\mathbf{x}_i)\, \mathbf{x}_i^0 (\mathbf{x}_i^0)^T, & I_{01} &:= \lim_{n\to\infty} \tfrac{1}{n} \textstyle\sum_{i=1}^{n} f_i(\varepsilon_i|\mathbf{x}_i)\, \mathbf{x}_i^0 \log(\mu_i), \\ I_{10} &:= \lim_{n\to\infty} \tfrac{1}{n} \textstyle\sum_{i=1}^{n} f_i(\varepsilon_i|\mathbf{x}_i)\, (\mathbf{x}_i^0)^T \log(\mu_i), & I_{11} &:= \lim_{n\to\infty} \tfrac{1}{n} \textstyle\sum_{i=1}^{n} f_i(\varepsilon_i|\mathbf{x}_i)\, \log(\mu_i)^2, \end{aligned}$$

where $\mathbf{x}_i^0 := (1, \log(p_{e_i}), \mathbf{z}_i)^T$ and $f_i(\varepsilon_i|\mathbf{x}_i)$ denotes the unknown conditional density of the error term $\varepsilon_i := \log(s_i) - (\boldsymbol{\alpha}^{full})^T \mathbf{x}_i$, which has to be estimated by smoothing techniques that are explained in the subsequent section.

As becomes evident from the formula for $\boldsymbol{\omega}_I$, information matrix $\mathbf{I}$ is a key element for the calculation of the FIC in quantile regression analysis, whereas for specifications that are estimated by maximum-likelihood methods, the well-known Fisher information measure represents such a key element (BEHL et al., 2012). Information matrix $\mathbf{I}$ also plays an important role for the asymptotic covariance matrix $\mathbf{V}$ defined by

$$\mathbf{V} := \tau(1-\tau)\mathbf{I}^{-1}\mathbf{V}_{\mathbf{x}}\mathbf{I}^{-1}, \tag{9}$$

where $\mathbf{V}_{\mathbf{x}}$ is a covariance matrix that is based on the vector $\mathbf{x}_i^{full}$ of the explanatory variables of the full model:

$$\mathbf{V}_{\mathbf{x}} := \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i (\mathbf{x}_i)^T.$$

For the FIC formula for the full model, we need the inverse of the asymptotic covariance matrix $\mathbf{V}$:

$$\mathbf{V}^{-1} = \frac{1}{\tau(1-\tau)} \mathbf{I}\mathbf{V}_{\mathbf{x}}{}^{-1}\mathbf{I} = \begin{pmatrix} J_{00} & J_{01} \\ J_{10} & J_{11} \end{pmatrix},$$

with the dimensions of the block matrices J_{00}, J_{01}, J_{10}, and J_{11} equaling those of matrices I_{00}, I_{01}, I_{10}, and I_{11}, respectively. On this basis, the FIC formula for the full model reads:

$$\mathrm{FIC}^{full} = \boldsymbol{\omega}_J^T \mathbf{V} \boldsymbol{\omega}_J, \tag{10}$$

where $\mathbf{V}$ captures the relative estimation variability and $\boldsymbol{\omega}_J$ is defined similar to $\boldsymbol{\omega}_I$:

$$\boldsymbol{\omega}_J := J_{10} J_{00}^{-1} \frac{\partial \Lambda}{\partial \boldsymbol{\xi}} - \frac{\partial \Lambda}{\partial \alpha_\mu}.$$

In the one-dimensional case $q = 1$ investigated here, due to $\frac{\partial \Lambda}{\partial \alpha_\mu} = 0$ and $\frac{\partial \Lambda}{\partial \boldsymbol{\xi}} = (0, -1, \mathbf{0})^T$, $\boldsymbol{\omega}_J$ degenerates to the negative of the second element of the vector $J_{10} J_{00}^{-1}$.

The FIC formula (10) reflects the fact that for the full model, there is no modeling bias by definition: $\mathbf{B} = \mathbf{0}$, whereas relative estimation variability $\mathbf{V}$ vanishes by definition for the null model and, hence, does not emerge from formula (5). Instead, modeling bias $\mathbf{B}$ becomes the pivotal factor in FIC formula (5) for the null model. In short, the FIC formulae for the null and full models reveal the trade-off between modeling bias and estimation variability.

3 Estimation Method

For obtaining estimates of FIC^0 and FIC^{full}, linear minimization problems have to be solved, as is typical for quantile regression methods (KOENKER, 2005). For instance, estimates of parameter vector $\boldsymbol{\alpha}^{full}(\tau)$ result from the following minimization problem:

$$\min_{\boldsymbol{\alpha}_{full}} \left(\sum_{r_i > 0} \tau \cdot r_i + \sum_{r_i < 0} (1 - \tau) \cdot |r_i| \right), \tag{11}$$

where underpredictions $r_i := \log(s_i) - Q_\tau(\log(s_i) | p_{e_i}, \mu_i, \mathbf{z}_i) = \log(s_i) - (\boldsymbol{\alpha}^{full})^T \mathbf{x}_i > 0$ are penalized by τ and overpredictions $r_i < 0$ by $1 - \tau$. This is reasonable, as for large τ one would not expect low estimates $\hat{Q}_\tau$ and vice versa, so that these incidences have to be penalized accordingly.

Just as ordinary least squares methods fit a linear function to the dependent variable by minimizing the expected squared error, quantile regression methods fit a linear model by minimizing the expected absolute error, using the asymmetric loss function $\rho_\tau(r) := 1(r > 0) \cdot \tau \cdot r + 1(r \leq 0) \cdot (1 - \tau) \cdot |r|$, where the indicator function $1(r > 0)$ indicates positive residuals r and $1(r \leq 0)$ non-positive residuals. $\rho_\tau(r)$ is called 'check' function, as its graph looks like a check mark.

For $\tau = 0.5$, in particular, the parameter estimates result from the minimization of the sum of the absolute deviations of r_i. This special case of a median regression is perfectly in line with the well-known statistical result that it is the median that minimizes the sum of the absolute deviations of a variable, whereas it is the mean that minimizes the sum of squared residuals, being a special case of OLS estimation. It is also well known that the median is more robust to outliers than the mean. This property translates to both median and quantile regressions in general, which have the advantage that they are more robust to outliers than mean (OLS) regression methods.

Conditional on p_e, μ, and $\mathbf{x}$, the conditional quantile functions (CQFs) given by (3) and (4) depend on the distribution of the corresponding error terms ε_i via the inverse distribution function $F^{-1}_{\varepsilon_i}(\tau)$. In the special case of homoscedasticity, that is, if the error terms ε_i were to be independent and identically distributed (iid) and, hence, the density of the errors and their inverse distribution function do not vary across observations ($f_i(\varepsilon_i) = f(\varepsilon_i)$ and likewise $F^{-1}_{\varepsilon_i}(\tau) = F^{-1}_{\varepsilon}(\tau)$), the CQFs exhibit common slopes, differing only in the intercepts $\alpha(\tau)$. In this case, there is no need for quantile regression methods if the focus is on marginal effects and elasticities, such as $\eta_{p_e}(s)$, as these are given by the invariant slope parameters, e. g. $\alpha_{p_e}(\tau) = \alpha_{p_e}$. In general, however, the CQFs Q_τ will differ at different values τ in more than just the intercept and may well be even non-linear in $\mathbf{x}$.

It also bears noting that in the special case of homoscedasticity, the asymptotic covariance matrix $\mathbf{V}$ would collapse to

$$\mathbf{V} = \frac{\tau(1-\tau)}{f^2(F^{-1}(\tau))}\mathbf{V_x}^{-1}. \tag{12}$$

This strongly reminds of the covariance matrix of an ordinary least squares estimator given by $\sigma^2\mathbf{V_x}^{-1}$. Note that in formula (12), the term $\tau(1-\tau)$ reflects the asymptotic variance of the check function ρ_τ. This term takes its maximum for $\tau = 0.5$, but gets small for percentiles close to 0 and 1. In this case, the term $\tau(1-\tau)$ may be dominated by factor $f^2(F^{-1}(\tau))$, leading to less precise parameter estimates, whereas the variance of the parameter estimates gets smaller for quantiles close to the median.

An important step in obtaining estimates of FIC^0 and FIC^{full} is to find suitable estimators for the matrix $\mathbf{I}$. To this end, smoothing techniques can be applied. BEHL, CLAESKENS, and DETTE (2014), as well as KIM and WHITE (2003), propose to use the estimator

$$\widehat{\mathbf{I}} = \frac{1}{2\hat{c}_n n}\sum_{i=1}^{n} \mathbf{1}_{\{-\hat{c}_n \le \hat{\varepsilon}_i \le \hat{c}_n\}}\mathbf{x}_i\mathbf{x}_i^T, \tag{13}$$

where $\hat{c}_n$ denotes a bandwidth that has to be determined by data-driven procedures, such as Cross Validation, and n denotes sample size.

4 Empirical Illustration

The data used in this illustrating example is drawn from regular surveys on the mobility behavior of German households (MOP, 2016). Households that participate in a survey are requested to fill out a questionnaire eliciting general household information, such as household income and the number of employed household members, person-related characteristics, and relevant aspects of everyday travel behavior. In addition, for a period of six weeks in the spring, households are requested to record detailed travel information for every car in the household, such as the price paid for fuel with each visit to a gas station, the liters of fuel consumed, and the kilometers driven. (For more details on the database, see FRONDEL, RITTER, and VANCE, 2012.)

We use this travel survey information to derive both the regressors and the dependent variable s, which is given by the total monthly distance driven in kilometers. On the basis of survey information that covers thirteen years, spanning 1997 through 2009, a period during which real fuel prices rose 1.97% per annum on average, we estimate the focus parameter $\Lambda(\tau) = -\eta_{p_e}(s) = -\alpha_{p_e}$ using quantile regression methods, thereby obtaining estimates of the rebound effect that depend on the percentile τ (Table 1).

Table 1: **Quantile Regression Estimates on the Rebound Effect given by Focus Parameter $\Lambda(\tau) = -\eta_{p_e}(s)$ resulting from the null model (3) and the full model (4).**

τ	$\widehat{\Lambda}^0(\tau)$	$\widehat{\Lambda}^{full}(\tau)$
0.1	0.898 (0.114)	0.869 (0.114)
0.3	0.714 (0.076)	0.686 (0.076)
0.7	0.551 (0.068)	0.493 (0.068)
0.9	0.561 (0.080)	0.551 (0.080)
Number of obs.	4,097	4,097

Note: Standard errors are in parentheses.

In line with FRONDEL, RITTER, and VANCE (2012), who estimated the rebound effect on the basis of the null model, we also find for the full model substantially smaller rebound effects for households with a high travel intensity, irrespective of the model specification. This outcome is in perfect accord with our expectations: To the extent that those who drive more are more dependent on car travel, we would expect them to exhibit less responsiveness to changes in fuel prices than those who drive less. Yet another source of heterogeneity in the rebound estimates is the kind of model specification: although the discrepancies across the null and the full model are not statistically significant, the magnitudes of the rebound estimates differ substantially, indicating that model selection is of great importance in our example.

In this respect, we now employ the FIC formulae presented in Section 2 to decide on whether efficiency variable μ should be included in specifications employed to estimate focus parameter $\Lambda(\tau) = -\eta_{p_e}(s)$, that is, the direct rebound effect. The unanimous recommendation of the FIC across all percentiles (Table 2), including those not reported, is to account for variable μ in the estimation of the rebound effect. After all, the FIC values for the full model (4) are always lower than those for the null model (3), with the difference of the FIC values between both models being larger for the lower percentiles than for higher τ. In qualitative terms, the same recommendation with respect to model selection results from the most common classic model selection criterion, the AIC, which is defined for a linear quantile regression of y on $\mathbf{x}$ as follows:

$$AIC(\tau) = n \log \left(\frac{1}{n} \sum_{i=1}^{n} \rho_\tau (y_i - \mathbf{x}_i^T \boldsymbol{\beta}) \right) + p,$$

where p denotes the number of coefficients belonging to vector $\boldsymbol{\beta}$.

Table 2: **FIC and AIC Values for the Quantile Regression of the null model (3) and the full model (4).**

	FIC		AIC	
	Null	Full	Null	Full
τ	Model	Model	Model	Model
0.1	55.03	**53.47**	-8,872.67	**-8,955.82**
0.3	25.25	**23.41**	-6,379.66	**-6,475.37**
0.7	20.00	**18.87**	-6,679.76	**-6,749.86**
0.9	26.46	**26.09**	-9,627.59	**-9,653.20**

To demonstrate that FIC and AIC may well yield divergent recommendations, we now present another example in which we compare four models, where in one model both μ and the household income are omitted, as, frequently, income information is lacking in empirical studies. While the recommendations of both FIC and AIC are the same for the majority of percentiles (Table 3), they differ for $\tau = 0.9$, for which the AIC prefers the full model (4), whereas the FIC recommends selecting the model in which income is omitted. That the AIC chooses the full model (4) does not come as a surprise, as AIC tends to select exhaustive models for large sample sizes. Furthermore, there are also divergent recommendations across percentiles τ: for low percentiles, the FIC selects the full model (4) including both μ and household income, but it prefers the model without household income for higher percentiles.

Table 3: **Quantile Regression Estimates of the FIC, AIC, and the Rebound Effect given by Focus Parameter $\Lambda(\tau) = -\eta_{p_e}(s)$.**

	$\widehat{\Lambda}(\tau)$	$\widehat{Var}(\widehat{\Lambda}(\tau))$	$\widehat{Bias^2}$	FIC	AIC
$\boldsymbol{\tau = 0.1}$:					
Model without Income nor μ	1.278	0.1536	0.1226	95.03	-803.81
Model without Income, but with μ	0.904	0.1624	0.0003	55.98	-806.32
Model (3) with Income, but without μ	1.381	0.1537	0.1120	91.41	-819.51
Model (4) with both Income and μ	0.708	0.1625	0	**55.91**	**-824.39**
$\boldsymbol{\tau = 0.3}$:					
Model without Income nor μ	0.865	0.0876	0.1503	81.84	-567.27
Model without Income, but with μ	0.599	0.0926	0.0002	31.92	-584.61
Model (3) with Income, but without μ	0.853	0.0877	0.1421	79.02	-566.50
Model (4) with both Income and μ	0.566	0.0927	0	**31.89**	**-585.88**
$\boldsymbol{\tau = 0.7}$:					
Model without Income nor μ	0.590	0.0724	0.1013	59.73	-578.03
Model without Income, but with μ	0.584	0.0765	0.0000	**26.31**	**-600.68**
Model (3) with Income, but without μ	0.645	0.0724	0.0988	58.91	-577.47
Model (4) with both Income and μ	0.556	0.0766	0	26.34	-600.27
$\boldsymbol{\tau = 0.9}$:					
Model without Income nor μ	1.030	0.1002	0.0438	49.53	-836.17
Model without Income, but with μ	0.916	0.1059	0.0000	**36.42**	-844.87
Model (3) with Income, but without μ	0.933	0.1002	0.0417	48.82	-837.08
Model (4) with both Income and μ	0.819	0.1060	0	36.46	**-845.68**

Note: These results are based on MOP survey information for the years 1997 and 1998 and a total of $n = 344$ observations.

5 Summary and Conclusion

The well-known direct rebound effect captures the behaviorally induced offset in the reduction of energy consumption following efficiency improvements. To investigate the heterogeneity of the rebound effect in mobility demand across different percentiles of the distribution of distance traveled, we have used quantile regression methods and the Focused Information Criterion (FIC) introduced by BEHL, CLAESKENS and DETTE (2014) for quantile regression analysis to choose between competing model specifications in which the endogenous variable energy efficiency μ is either omitted or included.

The FIC is conceived for targeted model searches, whereas conventional model selection criteria, such as the Akaike criterion (AIC), are mainly designed to find a model that is optimal in a general sense, regardless of a specific purpose of the data analysis. The empirical example presented in the previous section has illustrated that, first, the recommendations of the FIC differ across percentiles and, second, they deviate from those of the AIC in that the FIC selects a smaller model for the estimation of the rebound effect for high percentiles. Given that the FIC is precisely designed for purpose-specific model selection, we follow this recommendation, arguing that whenever a model is to be chosen that is optimal for the estimation of a certain parameter of interest, such as the rebound effect here, the FIC is a good choice.

References

ANGRIST, J. D. , PISCHKE, J.-S. (2009) *Mostly Harmless Econometrics – An Empiricists Companion*. Princeton University Press.

AKAIKE, H. (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19 (6), 716-723.

BEHL, P. , CLAESKENS, G. , DETTE H. (2014) Focussed Model Selection in Quantile Regression. *Statistica Sinica* 24, 601-624.

BEHL, P. , DETTE, H. , FRONDEL M. , TAUCHMANN, H. (2012) Choice is Suffering: A Focused Information Criterion for Model Selection. *Economic Modelling* 29 (3), 817-822.

BEHL, P. , DETTE, H. , FRONDEL M. , TAUCHMANN, H. (2013) Energy Substitution: When Model Selection Depends on the Focus. *Energy Economics* 39, 233-238.

BERKHOUT, P. H. G. , MUSKENS, J. C. , VELTHUISJEN, J. W. (2000) Defining the Rebound Effect. *Energy Policy* **28**, 425-432.

BROWNLEES, C. T. , GALLO, G. M. (2008) On Variable Selection for Volatility Forecasting: The Role of Focused Selection Criteria, *Journal of Financial Econometrics* 6 (4), 513-539.

CLAESKENS, G. , CROUX, C. , VAN KERCKHOVEN, J. (2007) Prediction-Focused Model Selection for Autoregressive Models. *Australian & New Zealand Journal of Statistics* 49 (4), 359-379.

CLAESKENS, G. , HJORT, N. L. (2003) The Focused Information Criterion. *Journal of the American Statistical Association* 98, 900-916.

CLAESKENS, G. , HJORT, N. L. (2008) Minimising average risk in regression models. *Econometric Theory* 24, 493-527.

DETTE, H. (1999) A Consistent Test for the Functional Form of a Regression based on a Difference of Variance Estimators. *Annals of Statistics* 27, 1012-1040.

DETTE, H., PODOLSKIJ, M. , VETTER, M. (2006) Estimation of Integrated Volatility in Continuous Time Financial Models With Applications to Goodness-of-Fit Testing. *Scandinavian Journal of Statistics* 33, 259-278.

FRONDEL, M., PETERS, J., VANCE, C. (2008) Identifying the Rebound: Evidence from a German Household Panel. *The Energy Journal* 29 (4), 154-163.

FRONDEL, M., RITTER, N. , VANCE, C. (2012) Heterogeneity in the Rebound: Further Evidence for Germany. *Energy Economics* 34, 461-467.

FRONDEL, M., VANCE, C. (2013) Re-Identifying the Rebound: What About Asymmetry? *The Energy Journal* 34 (4), 43-54.

HJORT, N. L. , CLAESKENS, G. (2006) Focussed Information Criteria and Model Averaging for Cox's Hazard Regression Model, *Journal of the American Statistical Association* 101, 1449-1464.

KIM, T. , WHITE, H. (2003) Estimation, Inference, and Specification Testing for possibly Misspecified Quantile Regression. In: Thomas B. Fomby, R. Carter Hill (ed.) *Maximum Likelihood Estimation of Misspecified Models: Twenty Years Later*, Advances in Econometrics, Volume 17. Emerald Group Publishing Limited, 107-132.

KOENKER, R. (2005) Quantile Regression. *Econometric Society Monographs No. 38*, Cambridge University Press, New York.

KOENKER, R. , BASSETT, G. (1978) Regression Quantiles. *Econometrica* 46 (1), 33-50.

MOP (2016) German Mobility Panel. http://www.ifv.uni-karlsruhe.de/MOP.html

PODOLSKIJ, M. DETTE, H. (2008) Testing the Parametric Form of the Volatility in

Continuous Time Diffusion Models - a Stochastic Process Approach. *Journal of Econometrics* 143, 56-73.

SCHWARZ, G. (1978) Estimating the Dimension of a Model. *Annals of Statistics* 6, 461-464.

SORRELL, S. , DIMITROUPOULOS, J. (2008) The Rebound Effect: Microeconomic Definitions, Limitations, and Extensions. *Ecological Economics* 65 (3), 636-649.

SORRELL, S., DIMITROUPOULOS, J. , SOMMERVILLE, M. (2009) Empirical Estimates of the Direct Rebound Effect: A Review. *Energy Policy* 37, 1356-1371.