Aïhounton, Ghislain B. D.; Henningsen, Arne

Working Paper

Units of Measurement and the Inverse Hyperbolic Sine Transformation

IFRO Working Paper, No. 2019/10

Provided in Cooperation with:
Department of Food and Resource Economics (IFRO), University of Copenhagen

Suggested Citation: Aïhounton, Ghislain B. D.; Henningsen, Arne (2019) : Units of Measurement and the Inverse Hyperbolic Sine Transformation, IFRO Working Paper, No. 2019/10, University of Copenhagen, Department of Food and Resource Economics (IFRO), Copenhagen

This Version is available at:
https://hdl.handle.net/10419/211087

DEPARTMENT OF FOOD AND RESOURCE ECONOMICS
UNIVERSITY OF COPENHAGEN

IFRO Working Paper

Units of Measurement and the Inverse Hyperbolic Sine Transformation

Ghislain B. D. Aihounton
Arne Henningsen

2019 / 10

IFRO Working Paper 2019 / 10

Units of Measurement and the Inverse Hyperbolic Sine Transformation

Authors: Ghislain B. D. Aihounton and Arne Henningsen

JEL-classification: C1, C5

Published December 2019

See the full series IFRO Working Paper here:
www.ifro.ku.dk/english/publications/ifro_series/working_papers/

Department of Food and Resource Economics (IFRO)
University of Copenhagen
Rolighedsvej 25
DK 1958 Frederiksberg DENMARK
www.ifro.ku.dk/english/

Units of Measurement and the Inverse Hyperbolic Sine Transformation

Ghislain B. D. Aihounton and Arne Henningsen

20th December 2019

Abstract

The inverse hyperbolic sine (IHS) transformation is frequently applied in econometric studies to transform right-skewed variables that include zero or negative values. We confirm a previous study that shows that regression results can largely depend on the units of measurement of IHS-transformed variables. Hence, arbitrary choices regarding the units of measurement for these variables can have a considerable effect on recommendations for policies or business decisions. In order to address this problem, we suggest a procedure for choosing units of measurement for IHS-transformed variables. A Monte Carlo simulation assesses this procedure under various scenarios and a replication of the study by Bellemare and Wichman (2019) illustrates the relevance and applicability of our suggested procedure.

Keywords: inverse hyperbolic sine, arcsinh, unit of measurement, scale factor

JEL codes: C1, C5

1 Introduction

The inverse hyperbolic sine (IHS or arcsinh) transformation, which empirical economists frequently apply to reduce the skewness of variables with zero or negative values has a major weakness in that it is not invariant to the unit of measurement of the transformed variable. Hence, the arbitrary choice of the unit of measurement (e.g., whether the variable is measured in Euros per year or in 1,000 USD per month) affects the regression results and, potentially, also the implications for policy recommendations or business decisions. This paper addresses the question about how to choose a suitable unit of measurement for IHS-transformed variables.

Although econometric methods do not require any specific distributional assumptions of explanatory or dependent variables,[1] empirical economists frequently apply the logarithmic (log) transformation in order to reduce the skewness and narrow the ranges of variables that have heavily right-skewed distributions. In many empirical applications, this has the advantage that the assumptions of regressions models (e.g., homoscedasticity) are more likely to be fulfilled, while the estimated coefficients are more robust to outliers and extreme values (Wooldridge, 2016, p. 172). As the log transformation can only be applied to strictly positive values, several studies add a "small" positive number to variables with zero values or replace zero values by a "small" positive number before they apply the log transformation. However, these arbitrary manipulations of the variables change the original structure of the data (Duan et al., 1983), and the arbitrary choice of a "small" positive value can substantially affect the empirical results (N'guessan et al., 2017).

In order to avoid these problems, the inverse hyperbolic sine transformation can be used instead of the log transformation (see, e.g., Johnson, 1949; Burbidge et al., 1988). This is because it can be applied to

[1] These methods demand that explanatory variables have a variance larger than zero. However, we do not consider this to be a distributional assumption.

zero and even negative values without any arbitrary manipulations of the original variable and it allows a similar interpretation of the regression results as the log transformation (see, e.g., Carroll et al., 2003; Carboni, 2012; Ravallion, 2017; Bahar and Rapoport, 2018; Bellemare and Wichman, 2019). Moreover, the IHS transformation has been used to overcome problems in regression analyses with right-skewed censored dependent variables (Carboni, 2012).

However, the way in which the IHS-transformation transforms the variable largely depends on the magnitude of the values of the transformed variable and, thus, its unit of measurement. More specifically, if one chooses the unit of measurement in a way that the values of the transformed variable are all rather small (e.g., weekly food expenditure of individual households in millions of USD), the IHS transformation has only a negligible effect on the variable (see left panel of Figure 1). In contrast, if one chooses the unit of measurement in a way that the values of the transformed variable are all rather large (e.g., annual food expenditure of individual households in USD), the IHS transformation is almost identical to a log transformation except for an upward shift by the value of log(2) (see right panel of Figure 1).

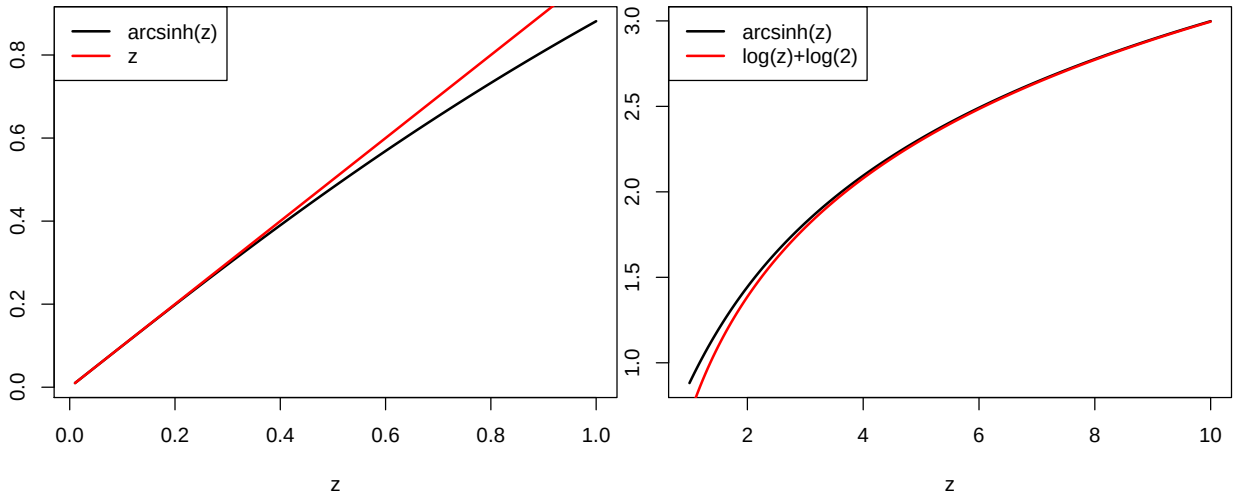

Figure 1: IHS transformation for small and large numbers

In order to investigate the dependence of the IHS transformation on the units of measurement, some studies (e.g., Pence, 2006; Bellemare et al., 2013; Bellemare and Wichman, 2019) use different scale factors to rescale IHS-transformed variables to different units of measurement before they apply the IHS transformation. While the main results of the empirical study of Bellemare et al. (2013) and of the simulation study of Bellemare and Wichman (2019, Table 1) are robust to the scale factor and, thus, the unit of measurement of the IHS-transformed variable, the results of the empirical analysis of Pence (2006) are substantially affected by the scale factor. In order to find the optimal scale factor, Carroll et al. (2003) and Pence (2006) estimate the scale factor along with the model coefficients in a Maximum-Likelihood estimation. Although this approach seems to be the best way to choose the scale of IHS-transformed variables, it has not been widely adopted in empirical studies. Reluctance to use this approach is probably due to its non-linearity in parameters so that standard econometric estimators that are implemented in many software packages (e.g., OLS, GLS, various time-series and panel-data estimators) cannot be applied.

Our paper suggests various simple-to-obtain criteria for choosing the unit of measurement for IHS-transformed variables and uses Monte-Carlo simulations and a real-world application to investigate the suitability of these criteria in empirical applications. Our results indicate that the R^2-value and the predictive R^2-value (Montgomery, 2012) of the regression are the most suitable general-purpose criteria for

choosing the units of measurement for IHS-transformed variables. However, under specific circumstances, some additional criteria can be used to increase the robustness and reliability of our suggested procedure.

The following section describes the IHS transformation and presents potential criteria for choosing units of measurement for IHS-transformed variables. Section 3 describes our Monte-Carlo simulations and presents the results of these simulations. Section 4 demonstrates our suggested procedure by applying it to an empirical application with real-world data. Finally, Section 5 concludes.

2 IHS transformation

Let us assume that we want to regress a dependent variable y_i on a set of covariates x_i, e.g., by:

$$y_i = \alpha + \beta x_i + \varepsilon_i, \tag{1}$$

where ε_i is the error term, subscript i indicates the observation, and α is an intercept and β is a vector of the slope coefficients to be estimated.

If the relationship between some of the covariates x_i and the dependent variable y_i is expected to be non-linear or some of the covariates or the dependent variable have a skewed distribution, applied economists frequently transform some or all of the variables:

$$\tilde{y}_i = f^y\left(y_i\right) \tag{2}$$

$$\tilde{x}_i = f^x\left(x_i\right), \tag{3}$$

where $f^y\left(\cdot\right)$ is a function and $f^x\left(\cdot\right)$ is a set of functions that transform some or all of the (dependent and explanatory) variables so that the regression model becomes:

$$\tilde{y}_i = \tilde{\alpha} + \tilde{\beta}\tilde{x}_i + \tilde{\varepsilon}_i, \tag{4}$$

where the tilde signs ($\tilde{}$) above α, β, and ε indicate that the transformation of the variables, in many cases, also affects the coefficients and the error term.

The most commonly used transformation for right-skewed variables is the logarithmic transformation, while the IHS transformation is usually recommended if the variables include some zero or even negative values (see, e.g., Bellemare and Wichman, 2019), e.g., income of individual persons, consumption of specific goods, or health expenditure. The IHS transformation of a variable z is defined as:

$$\widetilde{z} = \operatorname{arcsinh}\left(z\right) = \log\left(z + \sqrt{z^2 + 1}\right). \tag{5}$$

However, in contrast to a regression of equation (1) with linear variables or a regression of equation (4) with log-transformed variables, a regression with one or more IHS-transformed variables is not invariant to the units of measurement of these variables so that the arbitrary choice of the units of measurement affects the regression results and, potentially, recommendations for policies or business decisions.[2] As illustrated in the left panel of Figure 1, if one chooses the unit of measurement for a variable in a way that all values are rather small (e.g., smaller than 0.4), the IHS transformation has almost no effect and the estimation results

[2]Changing the unit of measurement of a linear dependent variable scales the intercept and all slope coefficients accordingly. Changing the units of measurement of a linear explanatory variable inversely scales the slope coefficient of this variable. Changing the unit of measurement of a log-transformed dependent or explanatory variable affects only the intercept. In spite of these effects on the coefficients, we consider these specifications to be invariant to units of measurement because none of them affects the practical interpretation of the results or unit-free measures such as elasticities or R^2 values.

are similar to those of a regression with linear variables (see also Bellemare and Wichman, 2019). Thus, if one wants to reduce the right-skewness of a variable, it does not make much sense to apply the IHS transformation to this variable if it has small values or if it is scaled to have small values because this has almost no effect on the skewness. In contrast, if one chooses the unit of measurement for a variable in a way that all values are rather large (e.g., larger than 3), the IHS transformation is almost identical to the log transformation (see right panel of Figure 1), and the estimation results are similar to those of a regression with log-transformed variables. However, if a variable contains zero values, it is impossible to scale this variable in a way that all values are rather large because zero values remain zero values no matter what unit of measurement is used. In the presence of zero values, changing the unit of measurement to larger units (i.e., multiplying the variable with a number smaller than one) moves the zero values "closer" to the non-zero values, while changing the unit of measurement to smaller units (i.e., multiplying the variable with a number larger than one) moves the zero values "further apart" from the non-zero values (see Figure 2).

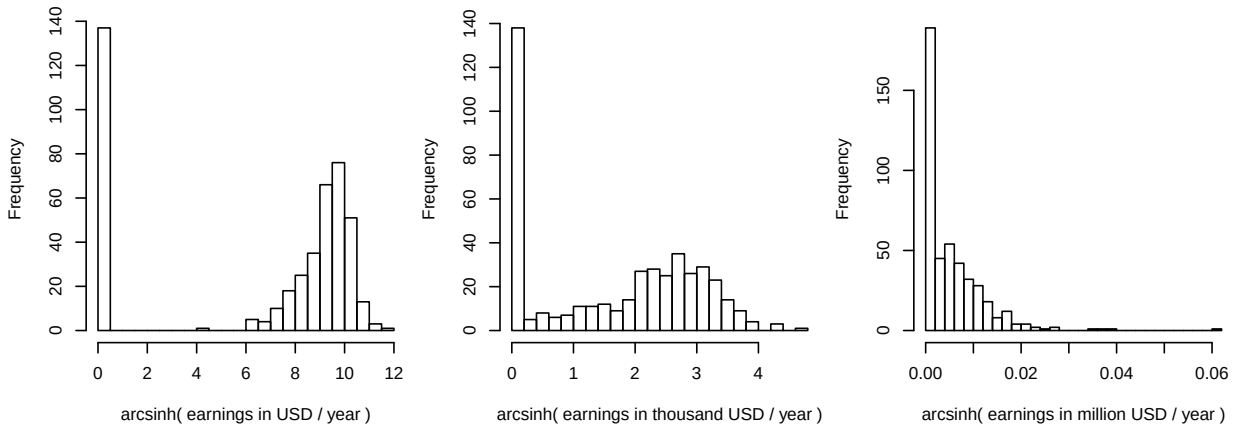

Figure 2: Histograms of IHS-transformed earnings in 1978 used in the empirical illustration in Section 4

As the unit of measurement for a variable clearly affects the IHS transformation (see Figures 1 and 2), it also affects the functional relationship that is implied by regression models with IHS-transformed variables (see Figures 3 and 4). Given that the assumed functional relationship affects the regression results, it is extremely important to carefully select the units of measurement for IHS-transformed variables. However, as the true functional relationship is usually unknown, the question arises of how an empirical analyst can choose suitable units of measurement for IHS-transformed variables.

Carroll et al. (2003) and Pence (2006) suggest including a scale factor for the IHS-transformed variable in the regression and estimating this scale factor jointly with the intercept and the slope coefficients. If both the dependent variable and the explanatory variable are IHS-transformed, one would get, for instance, the regression model:

$$\operatorname{arcsinh}(\theta_y y_i) = \tilde{\alpha} + \tilde{\beta} \operatorname{arcsinh}(\theta_x x_i) + \tilde{\varepsilon}_i, \tag{6}$$

where it is assumed that there is only a single explanatory variable (x_i), θ_y is the scale factor of the dependent variable, and θ_x is the scale factor of the explanatory variable. However, this regression specification is non-linear in parameters θ_y and θ_x, which is undesirable in many empirical applications.

A simpler approach would be to estimate the regression model with different sets of scale factors (θ_y, θ_x) and then to choose a set of scale factors based on one or more criteria. Table 1 lists potential criteria for choosing scale factors that determine the units of measurement for IHS-transformed variables. One group of these criteria assesses the "fit" of the regression model such as the R^2 value of the regression, the predictive R^2

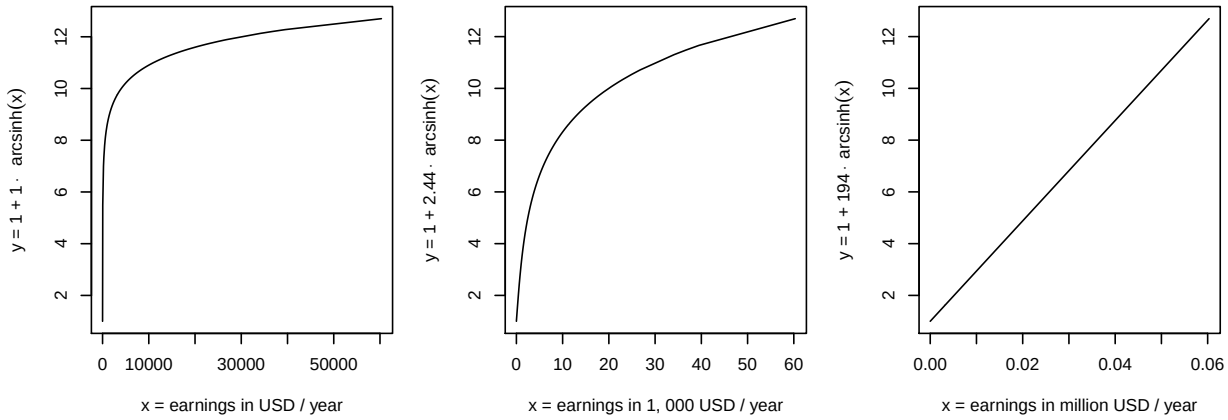

Figure 3: Functional relationships when using different units of measurement for an IHS-transformed explanatory variable (i.e., earnings in 1978 as used in the empirical illustration in Section 4; the slope coefficient is adjusted so that the range of the dependent variable is the same for all three units of measurement for the explanatory variable)

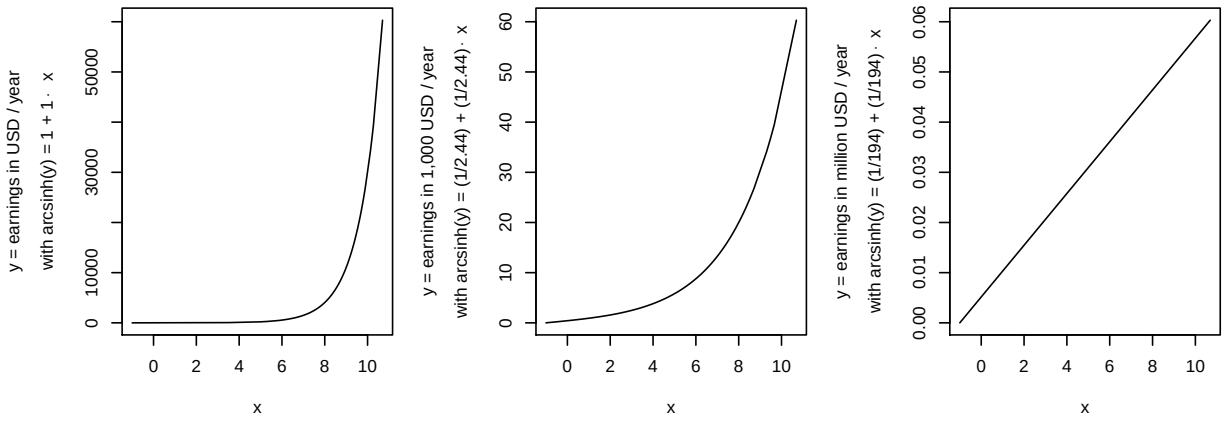

Figure 4: Functional relationships when using different units of measurement for an IHS-transformed dependent variable (i.e., earnings in 1978 as used in the empirical illustration in Section 4; the intercept and slope coefficient are adjusted so that the range of the explanatory variable is the same for all three units of measurement for the dependent variable)

value based on the "prediction error sum of squares" (PRESS) obtained by leave-one-out cross-validation (see, e.g., Montgomery, 2012), and the log-likelihood value of the regression. However, if the dependent variable is IHS-transformed, different scale factors θ_y imply different non-linear transformations of the dependent variable. Therefore,whether the above-mentioned goodness-of-fit criteria can be used to compare regression analyses with different scale factors applied to an IHS-transformed dependent variable is questionable. A prominent method to compare regression models with different non-linear transformations of the dependent variable is to adjust the log-likelihood value by adding the logarithm of the Jacobian of the transformation (Davidson and MacKinnon, 2004, p. 438–440). The adjusted log-likelihood value for regression analyses with an IHS-transformed dependent variable that is scaled by a factor θ_y is derived as:

$$logLikAdj = logLik - \frac{1}{2}\sum_{i=1}^{nObs} \ln\left(\theta_y^2 y_i^2 + 1\right) + nObs \cdot \ln\left(\theta_y\right), \tag{7}$$

where $logLikAdj$ is the adjusted log-likelihood value, $logLik$ is the (unadjusted) log-likelihood value, and $nObs$ is the number of observations (see, e.g., Carroll et al., 2003; Pence, 2006).[3] A second group of the potential criteria listed in Table 1 assesses the distribution of the regression residuals, e.g., how similar the distribution is to a normal distribution, how platykurtic or leptokurtic it is (Royston et al., 2011), or how symmetric it is.[4] Finally, a third group of criteria assesses the appropriateness of model assumptions, specifically homoscedasticity and the specified functional relationship between the explanatory variables and the dependent variable (see Medina et al., 2018).

3 Monte Carlo simulation

We conduct a Monte Carlo simulation to assess the suitability of the suggested criteria for choosing the scale factor and, thereby, the units of measurement for IHS-transformed variables.[5]

3.1 Data generating process

We generate artificial data sets for the Monte Carlo simulation under different scenarios. The generation of each artificial data set entails the following three steps (the abbreviations of the simulation parameters that differ between the scenarios are explained in Table 2):

1. Generation of the explanatory variable
   We use pseudo-random numbers from a log-normal distribution to generate a vector with "nObs" $\in \{100, 1000, 5000\}$ elements, where the mean value of the corresponding normal distribution is set to zero and its variance is set to "xVar" $\in \{0.1, 1, 10\}$, where larger values of "xVar" result in a more right-skewed distribution of the explanatory variable with more 'outliers' on the right-hand side of the distribution. Then we obtain the "xZero" $\in \{0, 0.1, 0.3, 0.5\}$ quantile of this vector and subtract this

[3] Several studies that scale IHS-transformed variables (e.g., Carroll et al., 2003; Pence, 2006) divide the IHS-transformed value by the scale factor so that the transformed variable is $\tilde{z} = \operatorname{arcsinh}(\theta z)/\theta = \log\left(\theta z + \sqrt{\theta^2 z^2 + 1}\right)/\theta$, where θ denotes the scale factor. If this procedure is applied to the dependent variable, the last term of equation (7), i.e., $nObs \cdot \ln(\theta_y)$, must be omitted.

[4] We note that Ordinary Least Squares (OLS) and many other regression methods—besides the fact that the error terms are independent and identically distributed (iid) with an expectation of zero—do not require any further assumptions about the distribution of the error term, e.g., neither a symmetric distribution nor a normal distribution (except for inference in small samples).

[5] Both the Monte Carlo simulation and the empirical illustration are conducted with the statistical software "R" (R Core Team, 2019) using the add-on packages "DescTools" (Signorell, 2019), "haven" (Wickham and Miller, 2019), "lmtest" (Zeileis and Hothorn, 2002), "moments" (Komsta and Novomestky, 2015), and "xtable" (Dahl et al., 2019) and code from Hopper (2014).

Table 1: Potential criteria for choosing the units of measurement for IHS-transformed variables in regression analyses

Abbreviation	Description	Rationale for using the criterion
rSquared	R^2 value of the regression	a larger value indicates a better fit of the model
pSquared	predictive R^2 value	a larger value indicates a better out-of-sample prediction performance
logLik	log-likelihood value of the regression	a larger value indicates a better fit of the model
logLikAdj	log-likelihood value of the regression adjusted as described in equation (7)	a larger value indicates a better fit of the model
Kolmogorov-Smirnov	test statistic of the Kolmogorov-Smirnov test for normality applied to the residuals	a smaller value indicates that the distribution of the residuals is closer to a normal distribution
Shapiro-Wilk	test statistic of the Shapiro-Wilk test for normality in small samples applied to the residuals	a larger value indicates that the distribution of the residuals is closer to a normal distribution
Shapiro-Fancia	test statistic of the Shapiro-Fancia test for normality in large samples applied to the residuals	a larger value indicates that the distribution of the residuals is closer to a normal distribution
Anderson	test statistic of the Anderson-Darling test for normality applied to the residuals	a smaller value indicates that the distribution of the residuals is closer to a normal distribution
Jarque-Bera	test statistic of the Jarque-Bera test for normality applied to the residuals	a smaller value indicates that the distribution of the residuals is closer to a normal distribution
Pearson	test statistic of the Pearson test for normality applied to the residuals	a smaller value indicates that the distribution of the residuals is closer to a normal distribution
kurtosis	the kurtosis (= fourth moment of the distribution) of the residuals	a value closer to zero indicates that the "peakedness" of the distribution of the residuals is more similar to a normal distribution
skewness	the skewness (= third moment of the distribution) of the residuals	a value closer to zero indicates a more symmetric distribution of the residuals
Breusch-Pagan	test statistic of the Breusch-Pagan test of homoscedasticity	a smaller value indicates a higher degree of homoscedasticity
RESET	test statistic of the Regression Equation Specification Error Test suggested by Ramsey (1969)	a smaller value indicates a more appropriate specification of the regression model

quantile from each element of the vector. Finally, we set all negative values to zero. This procedure generates a vector x with values that have a right-skewed distribution, for which a larger value of "xVar" results in a more right-skewed distribution with more extreme 'outliers', and that are left-censored at zero with "xZero" being the proportion of values equal to zero.

2. Error term
   We use pseudo-random numbers from a normal distribution ("rDist" = n), a Student t-distribution with 3 degrees of freedom ("rDist" = t), or a skew normal distribution with shape parameter $\alpha = 4$ ("rDist" = sn) to generate a vector with "nObs" $\in \{100, 1000, 5000\}$ elements, where the location and shape parameters of these distributions are chosen so that the distributions have an expected value of zero and the resulting regressions have an R^2-value of approximately "R2" $\in \{0.1, 0.5, 0.8, 0.95\}$.[6] This procedure generates a vector $\tilde{\varepsilon}$ with normally distributed values (if "rDist" = n), with many 'outliers' (if "rDist" = t), or with a skewed distribution (if "rDist" = sn).

3. Dependent variable
   We calculate a vector of values by equation (4) with $\tilde{\alpha} = 1$, $\tilde{\beta} = 1$, $\tilde{x} = \operatorname{arcsinh}(x)$ with x as obtained in step 1, and $\tilde{\varepsilon}$ as obtained in step 2. Then we obtain the "yZero" $\in \{0, 0.1, 0.3, 0.5\}$ quantile of this vector and subtract this quantile from each element of the vector and set all negative values to zero, which results in the vector $\tilde{y}$ of IHS-transformed values of the dependent variable. Finally, we obtain the non-transformed values of the dependent variable by $y = \sinh(\tilde{y})$. This procedure generates a vector y with values that have a right-skewed distribution and that are left-censored at zero with "yZero" being the proportion of values equal to zero.

Table 2: Scenarios used in the Monte Carlo simulation

Abbreviation	Description	Base scenario	Alternative scenarios
nObs	number of observations	1000	100, 5000
xVar	variance of the normally distributed variable, from which the log-normally distributed explanatory variable is derived: right-skewness and prevalence of 'outliers' in the explanatory variable	1	0.1, 10
xZero	share of observations for which the explanatory variable is zero	0.1	0, 0.3, 0.5
yZero	share of observations for which the dependent variable is zero	0.1	0, 0.3, 0.5
rDist	distribution of the error term	normal (n)	Student t with 3 degrees of freedom (t), skew normal with shape parameter $\alpha = 4$ (sn)
R2	approximate R^2 value	0.80	0.1, 0.5, 0.95

[6] In order to obtain an R^2-value of approximately "R2", we set the variance of the error term to $\left(\text{R2}^{-1} - 1\right) \cdot \text{VAR}\left(\tilde{\alpha} + \tilde{\beta}\tilde{x}_i\right)$, where $\text{VAR}\left(\tilde{\alpha} + \tilde{\beta}\tilde{x}\right)$ is the variance of the "deterministic" part of the dependent variable, i.e., the dependent variable before adding the error term (see third step of this procedure).

3.2 Simulation procedure

For each generated data set, we calculate the 'true' elasticity of the dependent variable y with respect to the explanatory variable x at the mean values of these variables as derived by Bellemare and Wichman (2019):

$$\epsilon = \tilde{\beta} \, \frac{\theta_x \bar{x}}{\theta_y \bar{y}} \, \frac{\sqrt{\theta_y^2 \bar{y}^2 + 1}}{\sqrt{\theta_x^2 \bar{x}^2 + 1}}, \tag{8}$$

where $\tilde{\beta} = \theta_x = \theta_y = 1$, $\bar{y}$ is the mean value of y, and $\bar{x}$ is the mean value of x. Furthermore, we estimate equation (6) with the scale factors θ_y and θ_x set to given values so that the regression equation is linear in (the remaining) parameters. We estimate this equation with each combination of $\theta_y \in \{10^{-6}, 10^{-5}, \ldots, 10^{6}\}$, and $\theta_x \in \{10^{-6}, 10^{-5}, \ldots, 10^{6}\}$ by Ordinary Least Squares (OLS). For each of these 169 regression analyses, we calculate the estimated elasticity by equation (8), where $\tilde{\beta}$ is the estimated slope coefficient and θ_y and θ_x are the scale factors used in the respective regression. Finally, we obtain the values of the criteria listed in Table 1 to find out which scale factors are the most suitable according to each of these criteria.

We repeat the entire procedure with 5,000 artificially generated data sets[7] for the base scenario as well as for each of the 15 alternative scenarios defined in Table 2.[8] For each scenario and each criteria for choosing the scale factor(s), we calculate the bias and root mean squared error (RMSE) of the elasticity estimate:

$$bias_{sc} = \frac{1}{5000} \sum_{j=1}^{5000} \left(\epsilon^*_{jsc} - \epsilon_{js} \right)$$

$$RMSE_{sc} = \sqrt{\frac{1}{5000} \sum_{j=1}^{5000} \left(\epsilon^*_{jsc} - \epsilon_{js} \right)^2},$$

where $bias_{sc}$ is the bias in scenario s when using criteria c for choosing the scale factors, $RMSE_{sc}$ is the root mean squared error in scenario s when using criteria c for choosing the scale factors, ϵ^*_{jsc} is the elasticity obtained when using criteria c for choosing the scale factors in replication j of scenario s, and ϵ_{js} is the 'true' elasticity in replication j of scenario s.

3.3 Results

Tables 3, 4, 5, and 6 summarise the simulation results for four different setups. These tables present the biases and RMSEs of the elasticities that are obtained from the regression analyses with the scale factors that were chosen by the criteria listed in Table 1 under the scenarios described in Table 2. Additionally, these tables present the biases and RMSEs of the elasticities obtained by using the 'correct' scale factors, i.e., $\theta_x = \theta_y = 1$, so that we can use these values as benchmarks for assessing the performance of the various criteria for choosing the scale factors or unit of measurement for the IHS transformed variables.

Even when using the 'correct' scale factors, i.e., $\theta_x = \theta_y = 1$, none of the elasticity estimates are completely unbiased, which is caused by the censoring of the dependent variable (which gives a biased estimate of the slope coefficient, i.e., $E\left[\hat{\tilde{\beta}} - 1\right] \neq 0$, in all scenarios in which "yZero" is larger than zero) and a correlation

[7] We investigated how the initial state of the pseudo-random number generator and the number of replications affect the results and found out that the effect of the initial state of the pseudo-random number generator on the simulation results is negligibly small when the number of replications is 5,000 or more.

[8] For the alternative scenarios, we altered only one simulation parameter to an "alternative" value at a time and kept all other simulation parameters at their "base" values. This keeps our Monte Carlo simulation concise and easily comprehensible, while an analysis with all 1728 combinations of simulation parameters would make it impossible to present the results in a journal paper. We make the R code for our Monte Carlo simulation freely available so that everyone can re-run the analysis for any scenario they consider relevant.

between the estimate of the slope coefficient and the remaining part of the right-hand side of equation (8) (so that even if the estimate of the slope coefficient is unbiased, i.e., $E\left[\hat{\beta}-1\right]=0$, $E\left[\hat{\epsilon}\right]=E\left[\hat{\beta}\cdot\kappa\right]=E\left[1\cdot\kappa\right]+E\left[\left(\hat{\beta}-1\right)\cdot\kappa\right]=\epsilon+E\left[\left(\hat{\beta}-1\right)\cdot\kappa\right]\neq\epsilon$, where $\hat{\epsilon}$ is the estimated elasticity, $\hat{\beta}$ is the estimated slope coefficient, and $\kappa=(x/y)\left(\sqrt{\bar{y}^2+1}/\sqrt{\bar{x}^2+1}\right)$). However, given that the elasticities are in the order of magnitude of one, the biases are relatively small in most scenarios. Only the scenarios in which half of the values of the dependent variable are censored at zero[9] or the R^2-value of the regression is only 10% give substantially biased estimates even if the 'correct' scale factors are used.

Table 3 focuses on the scale factor of the *dependent* variable, i.e., θ_y, by always using the 'true' scale factor of the explanatory variable, i.e., $\theta_x=1$. This setup mimics the problem that analysts have when only the dependent variable is IHS-transformed. The main results are:

- Almost all criteria result in relatively small biases and RMSEs for a large proportion of scenarios, while criteria "rSquared" and "pSquared" are the only criteria that perform relatively well across all scenarios.
- The criteria that assess the similarity of the distribution of the residuals with the normal distribution outperform the criteria "rSquared" and "pSquared" in a few scenarios, but these criteria perform poorly when the error terms are not normally distributed. However, as one can never be sure in real-world empirical applications that the 'true' error terms are indeed normally-distributed, the criteria that assess the similarity of the distribution of the residuals with the normal distribution seem to be not well suited to real-world empirical applications.
- Criterion "logLikAdj", i.e., the adjusted log-likelihood value, generally performs equally well as criteria "rSquared" and "pSquared" in many scenarios, but it performs poorly in the scenarios with a very right-skewed explanatory variable and with an R^2-value of only 10%. Furthermore, it performs extremely poorly in scenarios with 30% or more censored values of the dependent variable. Hence, although criterion "logLikAdj" seems to be very well suited from a theoretical point of view, one needs to be cautious when using it in empirical applications.
- Criteria "skewness", "Breusch-Pagan", and "RESET" perform reasonably well in most scenarios and, thus, may be used as additional criteria in specific empirical applications.

Table 4 focuses on the scale factor of the *explanatory* variable, i.e., θ_x, by always using the 'true' scale factor of the dependent variable, i.e., $\theta_y=1$. This setup mimics the problem that analysts have when only an explanatory variable is IHS-transformed. In this setup, criteria "logLik" and "logLikAdj" are monotonic transformations of the R^2-value and, thus, always indicate the same scale factors as criterion "rSquared". Hence, these two criteria are omitted in Table 4. The main results are:

- Most criteria perform substantially worse in this setup than in the case where one needs to choose a scale factor for the *dependent* variable. Hence, it seems to be more difficult to find an appropriate scale factor for an explanatory variable than for a dependent variable.
- However, criteria "rSquared" (and, thus, "logLik" and "logLikAdj"), "pSquared", and "RESET" are also suitable for choosing the scale factor for an explanatory variable as they perform similarly well as in the case where one needs to choose a scale factor for the dependent variable.
- All other criteria perform substantially worse in all but a few specific scenarios and, thus, seem to be unsuitable for empirical applications.

[9] Hence, in empirical applications in which the dependent variable is censored at around half or more of the observations, we suggest using a regression method for censored dependent variables instead of OLS.

Table 3: Simulation results (IHS of the dependent variable)

	base	nObs=100	nObs=5000	xVar=0.1	xVar=10	xZero=0	xZero=0.3	xZero=0.5	yZero=0	yZero=0.3	yZero=0.5	rDist=t	rDist=sn	R2=0.1	R2=0.5	R2=0.95
nObs	1000	100	5000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
xVar	1.00	1.00	1.00	0.10	10.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
xZero	0.10	0.10	0.10	0.10	0.10	0.00	0.30	0.50	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
yZero	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.00	0.30	0.50	0.10	0.10	0.10	0.10	0.10
rDist	n	n	n	n	n	n	n	n	n	n	n	t	sn	n	n	n
R2	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.10	0.50	0.95
bias correct scale	-0.04	-0.04	-0.04	-0.03	-0.09	-0.01	-0.07	-0.10	-0.09	-0.10	-0.39	-0.02	-0.03	-0.18	-0.08	-0.02
bias rSquared	-0.04	-0.04	-0.04	-0.03	-0.13	-0.01	-0.07	-0.10	-0.09	-0.10	-0.34	-0.02	-0.03	-0.21	-0.08	-0.02
bias pSquared	-0.04	-0.04	-0.04	-0.03	-0.13	-0.01	-0.07	-0.10	-0.09	-0.10	-0.34	-0.02	-0.03	-0.20	-0.08	-0.02
bias logLik	0.13	0.10	0.13	-0.01	0.51	0.18	0.12	0.14	0.09	0.08	-0.21	0.11	0.13	0.01	0.11	0.14
bias logLikAdj	-0.04	-0.04	-0.04	-0.03	0.30	-0.01	-0.07	-0.10	-0.09	2.61	3.08	-0.01	-0.03	0.36	-0.06	-0.02
bias Kolmogorov-Smirnov	-0.04	-0.01	-0.04	-0.03	-0.09	-0.01	-0.07	-0.10	-0.09	-0.08	-0.23	0.03	0.15	-0.18	-0.08	-0.02
bias Shapiro-Wilk	-0.04	-0.03	-0.04	-0.03	-0.09	-0.01	-0.07	-0.10	-0.09	0.06	0.05	0.15	0.15	-0.12	-0.08	-0.02
bias Shapiro-Fancia	-0.04	-0.02	-0.04	-0.03	-0.09	-0.01	-0.07	-0.10	-0.09	0.08	0.07	0.16	0.15	-0.12	-0.08	-0.02
bias Anderson	-0.04	-0.03	-0.04	-0.03	-0.09	-0.01	-0.07	-0.10	-0.09	0.01	-0.01	0.08	0.15	-0.15	-0.08	-0.02
bias Jarque-Bera	-0.04	-0.01	-0.04	-0.02	-0.09	-0.01	-0.07	-0.10	-0.09	0.15	0.50	0.17	0.15	-0.10	-0.08	-0.02
bias Pearson	-0.04	-0.01	-0.04	-0.02	-0.08	-0.01	-0.07	-0.10	-0.10	0.06	0.30	0.01	0.13	-0.15	-0.08	-0.02
bias kurtosis	0.10	0.07	0.12	0.16	-0.02	0.14	0.01	-0.08	-0.09	2.61	3.08	0.21	0.15	-0.17	-0.03	0.18
bias skewness	-0.04	-0.01	-0.04	-0.03	-0.13	-0.01	-0.07	-0.10	-0.09	0.37	-0.15	0.12	0.15	-0.11	-0.08	-0.02
bias Breusch-Pagan	-0.04	-0.03	-0.04	-0.03	-0.07	-0.01	-0.07	-0.10	-0.10	0.17	2.88	-0.00	-0.03	-0.14	-0.07	-0.02
bias RESET	-0.04	-0.01	-0.04	-0.03	-0.08	-0.01	-0.07	-0.10	-0.10	0.02	-0.17	-0.03	-0.03	-0.03	-0.06	-0.02
RMSE correct scale	0.04	0.07	0.04	0.03	0.09	0.02	0.07	0.10	0.09	0.10	0.39	0.04	0.04	0.19	0.08	0.02
RMSE rSquared	0.04	0.07	0.04	0.03	0.14	0.02	0.07	0.10	0.09	0.10	0.35	0.04	0.04	0.22	0.08	0.02
RMSE pSquared	0.04	0.07	0.04	0.03	0.14	0.02	0.07	0.10	0.09	0.10	0.35	0.04	0.04	0.22	0.08	0.02
RMSE logLik	0.15	0.19	0.14	0.02	0.57	0.19	0.14	0.15	0.11	0.11	0.23	0.22	0.15	0.32	0.14	0.15
RMSE logLikAdj	0.04	0.07	0.04	0.03	0.33	0.02	0.07	0.10	0.09	2.62	3.08	0.08	0.04	0.38	0.08	0.02
RMSE Kolmogorov-Smirnov	0.04	0.09	0.04	0.04	0.09	0.02	0.07	0.10	0.10	0.11	0.25	0.11	0.15	0.19	0.08	0.02
RMSE Shapiro-Wilk	0.04	0.08	0.04	0.04	0.09	0.02	0.07	0.10	0.09	0.15	0.29	0.17	0.15	0.15	0.08	0.02
RMSE Shapiro-Fancia	0.04	0.08	0.04	0.04	0.09	0.02	0.07	0.10	0.09	0.15	0.30	0.17	0.15	0.15	0.08	0.02
RMSE Anderson	0.04	0.08	0.04	0.03	0.09	0.02	0.07	0.10	0.09	0.13	0.29	0.13	0.15	0.17	0.08	0.02
RMSE Jarque-Bera	0.04	0.09	0.04	0.05	0.09	0.02	0.07	0.10	0.09	0.17	0.67	0.19	0.15	0.13	0.08	0.02
RMSE Pearson	0.04	0.09	0.04	0.04	0.09	0.02	0.07	0.10	0.10	0.20	0.83	0.09	0.15	0.17	0.08	0.02
RMSE kurtosis	0.12	0.13	0.13	0.16	0.04	0.17	0.05	0.08	0.09	2.62	3.08	0.26	0.15	0.18	0.08	0.18
RMSE skewness	0.04	0.09	0.04	0.04	0.14	0.02	0.07	0.10	0.09	0.48	0.25	0.16	0.15	0.13	0.08	0.02
RMSE Breusch-Pagan	0.05	0.08	0.05	0.04	0.12	0.03	0.07	0.10	0.10	0.17	2.94	0.10	0.04	0.18	0.09	0.02
RMSE RESET	0.04	0.15	0.04	0.03	0.08	0.02	0.07	0.10	0.10	0.14	0.20	0.13	0.04	0.25	0.08	0.02

Note: the abbreviations are explained in Tables 1 and 2.

Table 4: Simulation results (IHS of the explanatory variable)

	base	nObs=100	nObs=5000	xVar=0.1	xVar=10	xZero=0	xZero=0.3	xZero=0.5	yZero=0	yZero=0.3	yZero=0.5	rDist=t	rDist=sn	R2=0.1	R2=0.5	R2=0.95
nObs	1000	100	5000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
xVar	1.00	1.00	1.00	0.10	10.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
xZero	0.10	0.10	0.10	0.10	0.10	0.00	0.30	0.50	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
yZero	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.00	0.30	0.50	0.10	0.10	0.10	0.10	0.10
rDist	n	n	n	n	n	n	n	n	n	n	n	t	sn	n	n	n
R2	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.10	0.50	0.95
bias correct scale	-0.04	-0.04	-0.04	-0.03	-0.09	-0.01	-0.07	-0.10	-0.09	-0.10	-0.39	-0.02	-0.03	-0.18	-0.08	-0.02
bias rSquared	-0.04	-0.04	-0.04	-0.03	-0.09	-0.01	-0.07	-0.10	-0.09	-0.10	-0.50	-0.02	-0.03	-0.19	-0.08	-0.02
bias pSquared	-0.04	-0.04	-0.04	-0.03	-0.09	-0.01	-0.07	-0.10	-0.09	-0.10	-0.49	-0.02	-0.03	-0.19	-0.08	-0.02
bias Kolmogorov-Smirnov	-0.28	-0.28	-0.30	-0.22	-0.31	-0.09	-0.38	-0.13	-0.16	-0.43	-0.82	-0.31	-0.32	-0.39	-0.39	-0.02
bias Shapiro-Wilk	-0.11	-0.18	-0.05	-0.05	-0.26	-0.03	-0.25	-0.24	-0.11	-0.10	-0.39	-0.40	-0.29	-0.25	-0.25	-0.02
bias Shapiro-Fancia	-0.11	-0.19	-0.05	-0.05	-0.27	-0.03	-0.24	-0.24	-0.11	-0.11	-0.39	-0.41	-0.30	-0.25	-0.26	-0.02
bias Anderson	-0.26	-0.23	-0.30	-0.17	-0.27	-0.07	-0.32	-0.28	-0.13	-0.22	-0.39	-0.30	-0.30	-0.37	-0.31	-0.02
bias Jarque-Bera	-0.11	-0.21	-0.05	-0.06	-0.25	-0.06	-0.15	-0.17	-0.11	-0.26	-0.39	-0.63	-0.32	-0.35	-0.28	-0.02
bias Pearson	-0.20	-0.34	-0.20	-0.14	-0.31	-0.09	-0.41	-0.38	-0.21	-0.27	-0.48	-0.30	-0.30	-0.35	-0.28	-0.02
bias kurtosis	-0.27	-0.34	-0.25	-0.17	-0.29	-0.10	-0.10	-0.10	-0.23	-0.40	-0.52	-0.72	-0.48	-0.50	-0.34	-0.22
bias skewness	-0.15	-0.22	-0.10	-0.04	-0.26	-0.12	-0.22	-0.25	-0.11	-0.21	-0.47	-0.28	-0.25	-0.37	-0.31	-0.02
bias Breusch-Pagan	-0.23	-0.27	-0.19	-0.16	-0.48	-0.16	-0.07	-0.10	-0.17	-0.84	-1.48	-0.10	-0.09	-0.46	-0.22	-0.33
bias RESET	-0.04	-0.04	-0.04	-0.03	-0.09	-0.01	-0.07	-0.10	-0.09	-0.10	-0.55	-0.02	-0.03	-0.19	-0.08	-0.02
RMSE correct scale	0.04	0.07	0.04	0.03	0.09	0.02	0.07	0.10	0.09	0.10	0.39	0.04	0.04	0.19	0.08	0.02
RMSE rSquared	0.04	0.07	0.04	0.04	0.09	0.02	0.07	0.10	0.09	0.10	0.52	0.04	0.04	0.20	0.08	0.02
RMSE pSquared	0.04	0.07	0.04	0.04	0.09	0.02	0.07	0.10	0.09	0.10	0.51	0.04	0.04	0.20	0.08	0.02
RMSE Kolmogorov-Smirnov	0.29	0.35	0.30	0.27	0.32	0.12	0.44	0.15	0.19	0.44	0.83	0.32	0.33	0.42	0.41	0.02
RMSE Shapiro-Wilk	0.16	0.24	0.08	0.08	0.27	0.06	0.27	0.25	0.13	0.12	0.39	0.44	0.30	0.27	0.27	0.02
RMSE Shapiro-Fancia	0.17	0.25	0.08	0.08	0.27	0.06	0.27	0.25	0.13	0.12	0.39	0.45	0.30	0.27	0.28	0.02
RMSE Anderson	0.28	0.28	0.30	0.22	0.27	0.10	0.34	0.29	0.15	0.28	0.40	0.31	0.30	0.38	0.31	0.02
RMSE Jarque-Bera	0.16	0.26	0.07	0.09	0.26	0.11	0.18	0.19	0.13	0.32	0.40	0.65	0.34	0.38	0.29	0.03
RMSE Pearson	0.24	0.43	0.24	0.20	0.32	0.12	0.46	0.42	0.23	0.32	0.51	0.30	0.31	0.36	0.30	0.03
RMSE kurtosis	0.35	0.42	0.31	0.30	0.35	0.12	0.13	0.10	0.29	0.42	0.57	0.72	0.50	0.53	0.40	0.29
RMSE skewness	0.19	0.27	0.13	0.06	0.27	0.17	0.24	0.25	0.12	0.24	0.50	0.31	0.26	0.40	0.32	0.02
RMSE Breusch-Pagan	0.33	0.38	0.29	0.29	0.53	0.17	0.07	0.10	0.23	0.84	1.48	0.21	0.17	0.51	0.30	0.42
RMSE RESET	0.04	0.07	0.04	0.04	0.09	0.02	0.07	0.10	0.09	0.10	0.56	0.04	0.04	0.20	0.08	0.02

Note: the abbreviations are explained in Tables 1 and 2.

Table 5: Simulation results (IHS of both the explanatory and the dependent variable)

	base	nObs=100	nObs=5000	xVar=0.1	xVar=10	xZero=0	xZero=0.3	xZero=0.5	yZero=0	yZero=0.3	yZero=0.5	rDist=t	rDist=sn	R2=0.1	R2=0.5	R2=0.95
nObs	1000	100	5000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
xVar	1.00	1.00	1.00	0.10	10.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
xZero	0.10	0.10	0.10	0.10	0.10	0.00	0.30	0.50	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
yZero	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.00	0.30	0.50	0.10	0.10	0.10	0.10	0.10
rDist	n	n	n	n	n	n	n	n	n	n	n	t	sn	n	n	n
R2	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.10	0.50	0.95
bias correct scale	-0.04	-0.04	-0.04	-0.03	-0.09	-0.01	-0.07	-0.10	-0.09	-0.10	-0.39	-0.02	-0.03	-0.18	-0.08	-0.02
bias rSquared	-0.03	-0.02	-0.03	-0.02	-0.06	-0.00	-0.04	-0.04	-0.08	-0.08	-0.37	-0.02	-0.03	-0.18	-0.08	-0.01
bias pSquared	-0.03	-0.03	-0.03	-0.02	-0.09	-0.01	-0.05	-0.04	-0.09	-0.08	-0.37	-0.02	-0.03	-0.22	-0.09	-0.01
bias logLik	0.02	0.02	0.01	-0.02	0.11	0.03	0.02	0.03	0.01	-0.04	-0.33	0.00	0.03	0.02	0.04	0.00
bias logLikAdj	-0.04	-0.04	-0.04	-0.03	0.16	-0.01	-0.07	-0.10	-0.09	1.99	3.00	-0.03	-0.03	0.29	-0.06	-0.02
bias Kolmogorov-Smirnov	-0.28	-0.27	-0.30	-0.21	-0.29	-0.09	-0.45	-0.20	-0.17	-0.43	-0.79	-0.31	-0.30	-0.39	-0.39	-0.02
bias Shapiro-Wilk	-0.11	-0.17	-0.05	-0.05	-0.26	-0.03	-0.25	-0.24	-0.12	0.06	0.05	-0.53	-0.02	-0.23	-0.25	-0.02
bias Shapiro-Fancia	-0.11	-0.19	-0.05	-0.05	-0.26	-0.03	-0.25	-0.24	-0.12	0.08	0.07	-0.55	-0.01	-0.24	-0.26	-0.02
bias Anderson	-0.26	-0.22	-0.30	-0.17	-0.27	-0.07	-0.33	-0.28	-0.14	-0.11	-0.01	-0.32	-0.19	-0.35	-0.31	-0.02
bias Jarque-Bera	-0.12	-0.20	-0.05	-0.05	-0.70	-0.06	-0.15	-0.18	-0.13	0.14	0.09	-0.59	0.14	-0.40	-0.28	-0.02
bias Pearson	-0.20	-0.34	-0.20	-0.10	-0.29	-0.09	-0.44	-0.44	-0.25	-0.10	0.29	-0.31	-0.16	-0.32	-0.28	-0.04
bias kurtosis	-0.05	-0.24	0.08	0.14	-0.38	0.06	0.01	-0.13	-0.23	1.97	-0.94	-0.53	0.11	-0.48	-0.29	0.15
bias skewness	-0.16	-0.19	-0.11	-0.04	-0.63	-0.05	-0.22	-0.23	-0.44	0.22	-0.04	-0.21	-0.17	-0.41	-0.31	-0.02
bias Breusch-Pagan	-0.12	-0.34	-0.09	-0.74	-0.66	-0.16	-0.38	-0.39	-0.49	-0.23	0.33	-0.15	-0.09	-0.47	-0.20	-0.14
bias RESET	0.08	0.18	0.02	0.11	0.02	0.02	0.10	0.06	-0.09	0.09	-0.33	0.04	0.02	-0.06	0.09	0.01
RMSE correct scale	0.04	0.07	0.04	0.03	0.09	0.02	0.07	0.10	0.09	0.10	0.39	0.04	0.04	0.19	0.08	0.02
RMSE rSquared	0.05	0.08	0.04	0.03	0.14	0.05	0.06	0.06	0.11	0.09	0.38	0.04	0.05	0.33	0.12	0.02
RMSE pSquared	0.05	0.08	0.04	0.03	0.12	0.04	0.06	0.06	0.11	0.09	0.38	0.04	0.04	0.23	0.11	0.02
RMSE logLik	0.07	0.12	0.04	0.03	0.26	0.07	0.06	0.06	0.06	0.08	0.35	0.22	0.08	0.47	0.13	0.03
RMSE logLikAdj	0.04	0.07	0.04	0.04	0.16	0.02	0.07	0.10	0.09	1.99	3.02	0.05	0.04	0.33	0.08	0.02
RMSE Kolmogorov-Smirnov	0.29	0.35	0.30	0.27	0.30	0.12	0.50	0.27	0.21	0.44	0.81	0.33	0.32	0.41	0.41	0.07
RMSE Shapiro-Wilk	0.16	0.24	0.08	0.08	0.26	0.06	0.27	0.25	0.14	0.15	0.29	0.58	0.15	0.25	0.27	0.02
RMSE Shapiro-Fancia	0.17	0.26	0.08	0.08	0.26	0.06	0.27	0.25	0.14	0.15	0.30	0.59	0.15	0.26	0.28	0.02
RMSE Anderson	0.28	0.28	0.30	0.21	0.27	0.10	0.35	0.29	0.16	0.25	0.29	0.34	0.25	0.35	0.31	0.02
RMSE Jarque-Bera	0.16	0.27	0.07	0.09	0.74	0.11	0.19	0.21	0.17	0.20	0.53	0.65	0.16	0.47	0.29	0.05
RMSE Pearson	0.24	0.43	0.24	0.17	0.30	0.12	0.48	0.46	0.29	0.29	0.82	0.32	0.22	0.33	0.30	0.12
RMSE kurtosis	0.26	0.38	0.17	0.16	0.47	0.17	0.06	0.18	0.29	2.08	0.95	0.63	0.20	0.52	0.37	0.20
RMSE skewness	0.19	0.26	0.14	0.06	0.66	0.14	0.24	0.25	0.50	0.45	0.88	0.26	0.20	0.48	0.32	0.04
RMSE Breusch-Pagan	0.15	0.44	0.11	0.76	0.69	0.19	0.46	0.43	0.53	0.82	0.98	0.26	0.11	0.51	0.23	0.29
RMSE RESET	0.17	0.31	0.11	0.19	0.10	0.10	0.16	0.10	0.10	0.54	0.35	0.15	0.10	0.26	0.26	0.08

Note: the abbreviations are explained in Tables 1 and 2.

Table 6: Simulation results (IHS of both the explanatory and the dependent variable, same scale factor)

	base	nObs=100	nObs=5000	xVar=0.1	xVar=10	xZero=0	xZero=0.3	xZero=0.5	yZero=0	yZero=0.3	yZero=0.5	rDist=t	rDist=sn	R2=0.1	R2=0.5	R2=0.95
nObs	1000	100	5000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
xVar	1.00	1.00	1.00	0.10	10.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
xZero	0.10	0.10	0.10	0.10	0.10	0.00	0.30	0.50	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
yZero	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.00	0.30	0.50	0.10	0.10	0.10	0.10	0.10
rDist	n	n	n	n	n	n	n	n	n	n	n	t	sn	n	n	n
R2	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.10	0.50	0.95
bias correct scale	-0.04	-0.04	-0.04	-0.03	-0.09	-0.01	-0.07	-0.10	-0.09	-0.10	-0.39	-0.02	-0.03	-0.18	-0.08	-0.02
bias rSquared	-0.03	-0.01	-0.03	-0.02	-0.08	-0.00	-0.04	-0.04	-0.09	-0.06	-0.32	-0.02	-0.03	-0.17	-0.08	-0.01
bias pSquared	-0.03	-0.02	-0.04	-0.02	-0.10	-0.00	-0.05	-0.04	-0.10	-0.07	-0.32	-0.02	-0.03	-0.21	-0.09	-0.01
bias logLik	0.00	-0.01	0.00	-0.02	-0.10	0.00	0.00	0.01	-0.01	-0.07	-0.36	-0.02	-0.00	-0.11	0.00	-0.00
bias logLikAdj	-0.04	-0.04	-0.04	-0.03	-0.20	-0.01	-0.07	-0.10	-0.09	-0.36	-0.84	-0.03	-0.03	-0.61	-0.08	-0.02
bias Kolmogorov-Smirnov	-0.04	-0.05	-0.04	-0.03	-0.15	-0.01	-0.07	-0.10	-0.13	-0.10	-0.50	-0.03	-0.09	-0.22	-0.08	-0.02
bias Shapiro-Wilk	-0.04	-0.05	-0.04	-0.03	-0.14	-0.01	-0.07	-0.10	-0.09	-0.10	-0.39	-0.07	-0.09	-0.25	-0.08	-0.02
bias Shapiro-Fancia	-0.04	-0.05	-0.04	-0.03	-0.14	-0.01	-0.07	-0.10	-0.09	-0.10	-0.39	-0.07	-0.09	-0.25	-0.08	-0.02
bias Anderson	-0.04	-0.05	-0.04	-0.03	-0.15	-0.01	-0.08	-0.10	-0.11	-0.10	-0.39	-0.04	-0.09	-0.22	-0.08	-0.02
bias Jarque-Bera	-0.04	-0.04	-0.04	-0.03	-0.11	-0.01	-0.07	-0.10	-0.09	-0.11	-0.51	-0.08	-0.09	-0.28	-0.08	-0.02
bias Pearson	-0.04	-0.06	-0.04	-0.03	-0.18	-0.01	-0.13	-0.14	-0.18	-0.10	-0.39	-0.03	-0.09	-0.26	-0.08	-0.02
bias kurtosis	-0.04	-0.04	-0.04	-0.03	-0.12	-0.01	-0.08	-0.16	-0.09	-0.36	-0.84	-0.11	-0.05	-0.19	-0.08	-0.02
bias skewness	-0.04	-0.05	-0.04	-0.03	-0.11	-0.01	-0.08	-0.10	-0.10	-0.10	-0.43	-0.07	-0.09	-0.27	-0.08	-0.02
bias Breusch-Pagan	-0.04	-0.07	-0.04	-0.03	-0.11	-0.01	-0.07	-0.10	-0.33	-0.29	-0.67	-0.03	-0.03	-0.30	-0.08	-0.02
bias RESET	-0.03	-0.04	-0.03	-0.03	-0.09	0.00	-0.06	-0.09	-0.09	-0.06	-0.33	-0.04	-0.03	-0.22	-0.08	-0.01
RMSE correct scale	0.04	0.07	0.04	0.03	0.09	0.02	0.07	0.10	0.09	0.10	0.39	0.04	0.04	0.19	0.08	0.02
RMSE rSquared	0.05	0.10	0.04	0.03	0.14	0.05	0.06	0.06	0.11	0.09	0.33	0.04	0.05	0.33	0.12	0.02
RMSE pSquared	0.05	0.08	0.04	0.03	0.12	0.04	0.06	0.06	0.12	0.08	0.33	0.04	0.04	0.22	0.11	0.02
RMSE logLik	0.07	0.14	0.04	0.03	0.34	0.08	0.07	0.07	0.07	0.11	0.38	0.21	0.08	0.51	0.15	0.04
RMSE logLikAdj	0.04	0.07	0.04	0.03	0.26	0.02	0.07	0.10	0.09	0.36	0.84	0.04	0.04	0.62	0.09	0.02
RMSE Kolmogorov-Smirnov	0.05	0.08	0.04	0.03	0.16	0.02	0.08	0.10	0.15	0.10	0.52	0.04	0.09	0.23	0.08	0.02
RMSE Shapiro-Wilk	0.04	0.08	0.04	0.03	0.15	0.02	0.08	0.10	0.09	0.10	0.39	0.08	0.09	0.26	0.08	0.02
RMSE Shapiro-Fancia	0.04	0.07	0.04	0.03	0.15	0.02	0.08	0.10	0.09	0.10	0.39	0.08	0.09	0.27	0.08	0.02
RMSE Anderson	0.04	0.08	0.04	0.03	0.16	0.02	0.08	0.10	0.12	0.10	0.39	0.05	0.09	0.24	0.08	0.02
RMSE Jarque-Bera	0.04	0.07	0.04	0.03	0.12	0.02	0.08	0.10	0.09	0.11	0.53	0.09	0.09	0.28	0.08	0.02
RMSE Pearson	0.05	0.09	0.04	0.03	0.18	0.02	0.14	0.15	0.22	0.10	0.40	0.04	0.09	0.27	0.09	0.02
RMSE kurtosis	0.04	0.07	0.04	0.03	0.13	0.02	0.09	0.17	0.09	0.36	0.84	0.13	0.06	0.21	0.08	0.02
RMSE skewness	0.05	0.09	0.04	0.04	0.13	0.03	0.09	0.11	0.12	0.10	0.44	0.08	0.09	0.28	0.09	0.03
RMSE Breusch-Pagan	0.04	0.15	0.04	0.03	0.14	0.02	0.07	0.10	0.44	0.30	0.68	0.08	0.04	0.33	0.09	0.02
RMSE RESET	0.04	0.08	0.04	0.03	0.09	0.06	0.07	0.10	0.09	0.08	0.34	0.11	0.04	0.25	0.08	0.02

Note: the abbreviations are explained in Tables 1 and 2.

Table 5 simultaneously looks at the scale factor of the *dependent* and the *explanatory* variable, i.e., both θ_y and θ_x. This setup mimics the problem that analysts have when both the dependent variable and an explanatory variable are IHS-transformed. The main results are:

- The biases and RMSEs are similar to or higher than the biases and RMSEs for the case where one needs to choose a scale factor for the explanatory variable. This indicates that finding suitable values for two scale factors simultaneously is slightly more difficult than finding a suitable value for the scale factor for just an explanatory variable and much more difficult than finding a suitable value for the scale factor of the dependent variable only.
- Criteria "rSquared" and "pSquared" again generally result in relatively small biases and RMSEs.
- Criterion "logLikAdj", i.e., the adjusted log-likelihood value, performs equally well as criteria "rSquared" and "pSquared" in many scenarios, but it performs extremely poorly in scenarios with 30% or more censored values of the dependent variable.
- If 10% or less of the values of the dependent variable are censored, we suggest using criterion "logLikAdj" in addition to criteria "rSquared" and "pSquared", while all other criteria perform substantially worse in most scenarios and, thus, seem to be unsuitable for empirical applications with both an IHS-transformed dependent variable and one or more IHS-transformed explanatory variables.

Table 6 also simultaneously looks at the scale factor of the *dependent* and the *explanatory* variable. However, in contrast to the previous setup, the two scale factors are restricted to be equal, i.e., $\theta_y = \theta_x$. This setup mimics the problem that analysts have when both the dependent variable and an explanatory variable are IHS-transformed and it is reasonable to assume that the scale factors of the dependent variable and the explanatory variable are equal, e.g., when the IHS-transformed explanatory variable consists of lagged values of an IHS-transformed dependent variable. The main results are:

- The results for this setup are generally similar to those of the first setup that focuses on the scale factor for the dependent variable.
- Criteria "rSquared", "pSquared", and "RESET" generally perform the best across all scenarios.
- Criterion "skewness" and the criteria that compare the distribution of the residuals with a normal distribution also perform rather well across all scenarios, even in the scenarios with non-normally distributed error terms.
- Criteria "logLik", "logLikAdj", "kurtosis", and "Breusch-Pagan" perform poorly in some of the scenarios.

Overall, in all four different setups regarding the IHS-transformed variables and in almost all scenarios, criteria "rSquared" and "pSquared" result in biases and RMSEs that are among the smallest among all criteria and that are very similar to those obtained by using the 'correct' scale factors. Thus, criteria "rSquared" and "pSquared" can be considered the best general-purpose criteria for choosing the scale factors for IHS-transformed variables. Depending on the empirical specifications and the data used, criteria "RESET", "logLikAdj", "Breusch-Pagan", "skewness" and those that compare the distribution of the residuals with a normal distribution can be used as additional criteria, e.g., for robustness checks.

4 Empirical illustration

As an empirical illustration, we replicate and extend an empirical example used in Bellemare and Wichman (2019) who use data from Dehejia and Wahba (1999) to analyse how a randomised treatment with the National Supported Work (NSW) program affects annual income as initially done by LaLonde (1986). The (simplified) specification used by Bellemare and Wichman (2019) is:

$$\operatorname{arcsinh}(\theta_y y_i) = \alpha + \delta D_i + \beta \operatorname{arcsinh}(\theta_x x_i) + \tilde{\varepsilon}_i, \tag{9}$$

where y_i is person i's earnings (in USD) in 1978, D_i is a dummy variable that indicates whether person i received the treatment or not, x_i is person i's earnings (in USD) in 1975, which is used as a control variable, $\tilde{\varepsilon}_i$ is the error term, α, β, and δ are the coefficients to be estimated, and θ_y and θ_x are scale factors that are not applied by Bellemare and Wichman (2019), which means $\theta_y = \theta_x = 1$. The pre-treatment earnings (i.e., variable x_i) are zero in 65% of the observations, while the post-treatment earnings (i.e., variable y_i) are zero in 31% of the observations.

We estimate equation (9) by OLS using scale factors $\theta_y, \theta_x \in \{10^{-9}, 10^{-8}, \ldots, 10^{9}\}$, where we set $\theta_y = \theta_x$, because variables y_i and x_i are the same variables with the same unit of measurement that are just observed in different years. For each scale factor, Table 7 presents: (i) the estimates of coefficients δ and β; (ii)˜three different semi-elasticities, which measure the effect of the treatment on the post-treatment earnings using equations (10), (11), and (12) of Bellemare and Wichman (2019) for calculations without approximation, approximate calculations, and approximate calculations with small-sample correction, respectively; (iii) the elasticity that measures the effect of the pre-treatment earnings on the post-treatment earnings using equation (16) of Bellemare and Wichman (2019), and; (iv) the various criteria that can potentially be used for choosing the scale factor. When we apply scale factor $\theta_y = \theta_x = 1$, we obtain the same results that are reported in column (4) of Table 3 of Bellemare and Wichman (2019) because the two specifications are identical. When the scale factor $\theta_y = \theta_x$ approaches zero, the values of the scaled variables $\theta_y y_i$ and $\theta_x x_i$ become so small that the IHS transformation approaches a linear transformation (see left panel of Figure 1) so that the estimation approaches a regression with non-transformed variables and, thus, we obtain basically the same results as reported in column (1) of Table 3 of Bellemare and Wichman (2019), e.g., the same intercept and coefficient of the treatment dummy after scaling them by the inverse of the scale factor, the same coefficient and elasticity of the pre-treatment earnings, the same t-statistic of all three coefficients, the same adjusted and unadjusted R^2-values, and the same adjusted log-likelihood values.[10]

The semi-elasticity that quantifies the effect of the treatment on the post-treatment earnings calculated without approximation by equation (10) of Bellemare and Wichman (2019) largely depends on the scale factor. With an arbitrarily chosen unit of measurement and scale factor, the estimated effect of participation in the program on the earnings could be between 31% and 2,451% (Table 7) or even higher with a higher scale factor or a smaller unit of measurement. The semi-elasticities that are calculated with the approximate equations (11) or (12) of Bellemare and Wichman (2019) also largely depend on the scale factor and—given that they assume 'large' numbers—give poor approximations when very small scale factors are used. The elasticity that indicates the effect of the pre-treatment earnings on the post-treatment earnings varies much less than the semi-elasticity, but it still varies substantially between 0.043 and 0.091 for different units of

[10] The semi-elasticity that measures the effect of the treatment on the earnings that is reported in column (1) of Table 3 of Bellemare and Wichman (2019) differs from our semi-elasticity in Table 7 because it is calculated in a different way, i.e., by using the average post-treatment earnings over the entire sample rather than the expected post-treatment earnings of the non-treated persons as denominator. If one performs a linear regression with non-transformed variables and divides the estimated coefficient of the treatment dummy by the expected post-treatment earnings of a non-treated person with all other variables equal to the sample mean, one obtains the same semi-elasticity as reported for the smallest scale factor in Table 7.

Table 7: OLS empirical estimation results (IHS of both dependent and independent variables)

scale factor	unit of measurement	coef treatment	ela treatment (10)	ela treatment (11)	ela treatment (12)	coef earnings	ela earnings (16)	rSquared	pSquared	logLik	logLikAdj	Kolmogorov-Smirnov	Shapiro-Wilk	Shapiro-Fancia	Anderson	Jarque-Bera	Pearson	kurtosis	skewness	Breusch-Pagan	RESET
10^{-9}	10^{9} USD	$1.75 \cdot 10^{-6}$	0.38	0.00	0.00	0.17	0.043	0.0241	0.0092	**4681**	-4541	0.15	0.80	0.79	18.3	4815.7	477	15.66	2.54	4.39	0.40
10^{-8}	10^{8} USD	$1.75 \cdot 10^{-5}$	0.38	0.00	0.00	0.17	0.043	0.0241	0.0092	3656	-4541	0.15	0.80	0.79	18.3	4815.7	477	15.66	2.54	4.39	0.40
10^{-7}	10^{7} USD	$1.75 \cdot 10^{-4}$	0.38	0.00	0.00	0.17	0.043	0.0241	0.0092	2631	-4541	0.15	0.80	0.79	18.3	4815.5	477	15.66	2.54	4.39	0.40
10^{-6}	10^{6} USD	$1.75 \cdot 10^{-3}$	0.38	0.00	0.00	0.17	0.043	0.0241	0.0092	1607	-4541	0.15	0.80	0.79	18.3	4798.0	477	15.64	2.54	4.40	0.40
10^{-5}	10^{5} USD	$1.72 \cdot 10^{-2}$	0.38	0.02	0.02	0.17	0.043	**0.0242**	**0.0093**	589	-4536	**0.14**	0.81	0.80	18.0	3486.2	445	13.94	2.38	4.62	0.39
10^{-4}	10^{4} USD	$1.14 \cdot 10^{-1}$	0.31	0.12	0.12	0.15	0.044	0.0232	0.0091	-281	-4456	0.15	0.90	0.90	**13.3**	57.5	463	3.57	0.95	5.86	**0.05**
10^{-3}	10^{3} USD	$2.99 \cdot 10^{-1}$	0.38	0.35	0.34	0.11	0.091	0.0202	0.0070	-754	-4364	0.18	**0.91**	**0.91**	16.4	**28.1**	**368**	**1.56**	**-0.06**	**0.22**	0.32
10^{-2}	10^{2} USD	$5.43 \cdot 10^{-1}$	0.72	0.72	0.68	0.09	0.091	0.0210	0.0080	-997	-4263	0.20	0.85	0.85	28.1	32.5	572	1.58	-0.47	4.28	0.16
10^{-1}	10^{1} USD	$7.89 \cdot 10^{-1}$	1.20	1.20	1.09	0.08	0.078	0.0202	0.0072	-1160	-4108	0.22	0.80	0.80	38.7	41.3	689	1.66	-0.63	6.37	0.20
10^{0}	USD	1.03	1.81	1.81	1.58	0.07	0.070	0.0194	0.0064	-1281	-3914	0.24	0.77	0.77	45.8	50.3	762	1.69	-0.70	7.19	0.25
10^{1}	10^{-1} USD	1.28	2.59	2.59	2.14	0.07	0.066	0.0189	0.0059	-1377	-3695	0.26	0.75	0.76	50.6	57.7	854	1.71	-0.74	7.64	0.27
10^{2}	10^{-2} USD	1.52	3.59	3.59	2.79	0.06	0.063	0.0186	0.0056	-1457	-3459	0.27	0.74	0.74	53.8	63.2	930	1.72	-0.76	7.92	0.29
10^{3}	10^{-3} USD	1.77	4.87	4.87	3.53	0.06	0.061	0.0183	0.0053	-1524	-3211	0.28	0.73	0.73	56.1	67.5	983	1.72	-0.77	8.11	0.30
10^{4}	10^{-4} USD	2.01	6.49	6.49	4.35	0.06	0.060	0.0181	0.0051	-1583	-2954	0.29	0.72	0.73	57.8	70.8	1459	1.72	-0.78	8.25	0.31
10^{5}	10^{-5} USD	2.26	8.57	8.57	5.25	0.06	0.059	0.0180	0.0049	-1635	-2690	0.30	0.72	0.72	59.0	73.4	1510	1.73	-0.79	8.35	0.31
10^{6}	10^{-6} USD	2.50	11.23	11.23	6.23	0.06	0.058	0.0178	0.0048	-1681	-2422	0.30	0.72	0.72	60.0	75.5	1499	1.73	-0.79	8.42	0.32
10^{7}	10^{-7} USD	2.75	14.63	14.63	7.28	0.06	0.057	0.0177	0.0047	-1723	-2148	0.31	0.71	0.71	60.8	77.1	1517	1.73	-0.79	8.49	0.32
10^{8}	10^{-8} USD	2.99	18.97	18.97	8.38	0.06	0.056	0.0176	0.0046	-1762	-1871	0.31	0.71	0.71	61.4	78.3	1529	1.73	-0.80	8.54	0.32
10^{9}	10^{-9} USD	3.24	24.51	24.51	9.52	0.06	0.056	0.0176	0.0045	-1797	**-1591**	0.32	0.71	0.71	61.9	79.4	1523	1.73	-0.80	8.58	0.33

Note: 'coef treatment' is the coefficient of the treatment dummy, 'ela treatment (10)' is the semi-elasticity of the treatment calculated without approximation by equation (10) of Bellemare and Wichman (2019), 'ela treatment (11)' is the approximate semi-elasticity of the treatment calculated by equation (11) of Bellemare and Wichman (2019), 'ela treatment (12)' is the approximate semi-elasticity of the treatment with small-sample correction calculated by equation (12) of Bellemare and Wichman (2019), 'coef earnings' is the coefficient of the IHS-transformed earnings in 1975, 'ela earnings (16)' is the elasticity of the earnings in 1975 calculated by equation (16) of Bellemare and Wichman (2019), and all other column headings are criteria for choosing the optimal scale as described in Table 1.

measurement. So the question arises, which unit of measurement and, thus, which corresponding semi-elasticity and elasticity should we choose?

Criterion "logLik" monotonically increases with decreasing scale factors, while criterion "logLikAdj" monotonically increases with increasing scale factors, even far beyond the range of the scale factors that are presented in Table 7. Hence, these two criteria seem to be unsuitable in the specific case of our empirical illustration, which could be related to the result of our Monte-Carlo simulation where these two criteria perform very poorly in some of the scenarios, particularly when there is a high proportion of zero values in the dependent variable as is the case in our empirical illustration. Criteria "rSquared", "pSquared", and "Kolmogorov-Smirnov" indicate that measuring earnings in 100,000 USD per year (i.e., scaling the original variable by 10^{-5}) is most appropriate, while criteria "Anderson" and "RESET" suggest that earnings should be measured in 10,000 USD per year (i.e., scaling the original variable by 10^{-4}), while criterion "Breusch-Pagan" and the remaining criteria that assess the distribution of the residuals point to measuring earnings in 1,000 USD per year (i.e., scaling the original variable by 10^{-3}) (Table 7). Hence, the semi-elasticity of the treatment is likely between 0.31 and 0.38 and the elasticity of the pre-treatment earnings is likely between 0.043 and 0.091. Given that our Monte-Carlo simulation indicates that the criteria "rSquared", "pSquared", and "RESET" are the the most reliable criteria, we can narrow down the range of the elasticity of the pre-treatment earnings to 0.043 to 0.044.

In a real-world application, we recommend repeating the search procedure with scale factors in the range of those scale factors that are most appropriate according to relevant criteria and with narrower distances between the scale factors (e.g., with scale factors $10^{-5.5}$, $10^{-5.25}$, 10^{-5}, $10^{-4.75}$, ..., $10^{-2.5}$ in our empirical application), at least if the regression results substantially differ between the scale factors that are pointed out to be suitable by relevant criteria. Furthermore, we emphasize that the search procedure should be conducted with all covariates that are used in the final regression model because the functional relationship between the covariates and the dependent variable and, thus, the suitability of the scale factors of IHS-transformed variables can depend on the covariates that are used in the regression model.

5 Conclusion

The inverse hyperbolic sine (IHS) transformation is frequently used in econometric analyses to transform right-skewed variables when the logarithmic transformation cannot be applied due to zero or negative values. We confirm the results of Pence (2006) who finds that the unit of measurement of IHS-transformed variables can substantially affect the regression results. Our Monte Carlo simulation shows that one can repeat the regression analysis with different units of measurement and then use the R^2-value and the predictive R^2-value (Montgomery, 2012) of the regression to choose suitable units of measurement for IHS-transformed variables. Depending on the empirical specifications and the data used, several other criteria can be used additionally to increase the robustness and reliability of the choice of the units of measurement. An empirical illustration with real-word data demonstrates the applicability of our approach. Both our Monte Carlo simulation and our empirical illustration cast some doubt on the suitability of using the adjusted log-likelihood value for choosing the units of measurement for IHS-transformed variables and, thus, the one-step procedure suggested by Carroll et al. (2003) and Pence (2006). Given that our analysis includes only a limited number of Monte Carlo scenarios and only one application with real-world data, we suggest that empirical analysts and applied econometricians extend our study to cover their specifications and data sets, e.g., by using and adjusting the code for our Monte Carlo analysis that we provide as an online supplement to this paper. Given the substantial dependence of regression results on the units of measurement for IHS-transformed variables, using

our approach to choose suitable units of measurement for IHS-transformed variables can contribute to more reliable estimates of econometric analyses and, thus, to better policies and business decisions.

Acknowledgements

We gratefully acknowledge the financial support provided by the Ministry of Foreign Affairs of Denmark (Grant: 14-02KU). The authors would like to thank Géraldine Henningsen and Marc F. Bellemare for their valuable comments and suggestions on an earlier draft of this article. Of course, the authors take full responsibility for any remaining errors.

References

Bahar, D. and Rapoport, H. (2018). Migration, knowledge diffusion and the comparative advantage of nations. *The Economic Journal*, 128(612):F273–F305.

Bellemare, M. F., Barrett, C. B., and Just, D. R. (2013). The welfare impacts of commodity price volatility: Evidence from rural Ethiopia. *American Journal of Agricultural Economics*, 95(4):877–899.

Bellemare, M. F. and Wichman, C. J. (2019). Elasticities and the inverse hyperbolic sine transformation. *Oxford Bulletin of Economics and Statistics*, forthcoming.

Burbidge, J. B., Magee, L., and Robb, A. L. (1988). Alternative transformations to handle extreme values of the dependent variable. *Journal of the American Statistical Association*, 83(401):123–127.

Carboni, O. A. (2012). An empirical investigation of the determinants of r&d cooperation: An application of the inverse hyperbolic sine transformation. *Research in Economics*, 66(2):131–141.

Carroll, C. D., Dynan, K. E., and Krane, S. D. (2003). Unemployment risk and precautionary wealth: Evidence from households' balance sheets. *Review of Economics and Statistics*, 85:586–604.

Dahl, D. B., Scott, D., Roosen, C., Magnusson, A., and Swinton, J. (2019). *xtable: Export Tables to LaTeX or HTML*. R package version 1.8-4, https://CRAN.R-project.org/package=xtable.

Davidson, R. and MacKinnon, J. G. (2004). *Econometric Theory and Methods*. Oxford University Press, New York.

Dehejia, R. H. and Wahba, S. (1999). Causal effects in non-experimental studies: Re-evaluating the evaluation of training programs. *Journal of the American Statistical Association*, 94(448):1053–1062.

Duan, N., Manning, W. G., Morris, C. N., and Newhouse, J. P. (1983). A comparison of alternative models for the demand for medical care. *Journal of business & economic statistics*, 1(2):115–126.

Hopper, T. (2014). Can we do better than R-squared? Blog post at: https://tomhopper.me/2014/05/16/can-we-do-better-than-r-squared/ [accessed October 22, 2019].

Johnson, N. L. (1949). Systems of frequency curves generated by methods of translation. *Biometrika*, 36(1/2):149–176.

Komsta, L. and Novomestky, F. (2015). *moments: Moments, cumulants, skewness, kurtosis and related tests*. R package version 0.14, https://CRAN.R-project.org/package=moments.

LaLonde, R. J. (1986). Evaluating the econometric evaluations of training programs with experimental data. *The American Economic Review*, 76(4):604–620.

Medina, L., Castro, P., Kreutzmann, A.-K., and Rojas-Perilla, N. (2018). *The R Package trafo for Transforming Linear Regression Models. Vignette to the R package trafo, CRAN, https://CRAN.R-project.org/package=trafo.*

Montgomery, D. C. (2012). *Design and Analysis of Experiments.* John Wiley & Sons, New York, eighth edition edition.

N'guessan, Y. G., Featherstone, A., Odeh, O., and Upendram, S. (2017). Choice of the empirical definition of zero in the translog multiproduct cost functional form. *Applied Economics Letters*, 24(15):1112–1120.

Pence, K. M. (2006). The role of wealth transformations: An application to estimating the effect of tax incentives on saving. *The BE Journal of Economic Analysis & Policy*, 5(1).

R Core Team (2019). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

Ramsey, J. B. (1969). Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society. Series B (Methodological)*, 31(2):350–371.

Ravallion, M. (2017). A concave log-like transformation allowing non-positive values. *Economics Letters*, 161:130–132.

Royston, P., Lambert, P. C., et al. (2011). Flexible parametric survival analysis using stata: beyond the cox model.

Signorell, A. (2019). *DescTools: Tools for Descriptive Statistics.* R package version 0.99.30, https://CRAN.R-project.org/package=DescTools.

Wickham, H. and Miller, E. (2019). *haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files.* R package version 2.1.1, https://CRAN.R-project.org/package=haven.

Wooldridge, J. M. (2016). *Introductory Econometrics. A Modern Approach.* Cengage Learning, 6th edition.

Zeileis, A. and Hothorn, T. (2002). Diagnostic checking in regression relationships. *R News*, 2(3):7–10.