# Total Least Squares Bias when Explanatory Variables are Correlated

Ross McKitrick
Department of Economics and Finance
University of Guelph
Guelph ON Canada N1G 2W1
ross.mckitrick@uoguelph.ca
ORCID: 0000-0002-2257-0675

March 22, 2022

Key points

- Total Least Squares (TLS) is widely used in climatological applications such as optimal fingerprinting
- Little is known about consistency and bias when variances are not equal and variables are correlated
- Monte Carl analysis reveals TLS may suffer extreme bias problems in applications that resemble optimal fingerprinting

## Abstract

Total Least Squares (TLS) or orthogonal regression is used to remedy attenuation bias in optimal fingerprinting regressions. Consistency properties in multivariate applications require strong assumptions about unobservable variance ratios. Monte Carlo analysis is used herein to examine coefficient biases when the explanatory variables are correlated and have heterogeneous error variances. Ordinary Least Squares (OLS) exhibits the expected attenuation bias patterns which vanish as the noise variances on the explanatory variable disappear. TLS is generally more biased than OLS except under homogeneous noise variances. When the explanatory variables are negatively correlated TLS imparts a large upward bias which gets worse as the noise variance on the explanatory variable gets smaller. In general without specific diagnostic information TLS should not be considered an improvement on OLS and can yield extremely biased coefficients.

## Plain Language Summary

Total Least Squares (TLS) or orthogonal regression is a regression technique widely used when explanatory variables are noisy, such as climate signal detection regressions, in order to remedy the downward bias associated with Ordinary Least Squares (OLS). The theory behind TLS was developed for univariate models but in multivariate applications like signal detection little is known about bias except in the special case when the noise variances on all variables are assumed to be equal. Monte Carlo analysis is used herein to study coefficient biases when the explanatory variables are correlated and the model variables have error variances that may differ. The bias pattern of OLS generally goes as expected, it tends to be relatively small and it vanishes as the noise variance on the explanatory variable goes to zero. TLS behaves in a very erratic way and unexpectedly tends to exhibit large biases that get worse as the noise variance on the explanatory variable goes to zero. When the explanatory variables are negatively correlated, as is the case with the sample climate signals examined herein, TLS imparts an upward bias. Valid interpretation of TLS coefficients requires specific diagnostic information otherwise they may be misleading.

Key words: Total Least Squares; Orthogonal Regression; Optimal Fingerprinting.

## 1. Introduction

When the explanatory variable in a univariate regression model is measured with error (denoted Errors-in-Variables or EIV) the ordinary least squares (OLS) slope coefficient is biased downwards, a phenomenon called attenuation bias (Wooldridge 2020). Orthogonal regression provides a correction in the univariate regression case as long as the regression model is correctly specified (Carroll and Ruppert 1996). The estimation technique is referred to as Total Least Squares (TLS) in the multivariate case (Gleser 1981, Markovsky and Van Huffel 2007). Since Allen and Stott (2003) TLS has been widely-used in climatology for the purpose of “optimal fingerprinting” regression, which forms the basis of causal attribution claims for observed climate changes, as well as estimating the magnitude of the “carbon budget”, or cumulative carbon dioxide emission limits consistent with climate warming targets (Gillett et al. 2013). Claims about the validity of TLS coefficient estimates, especially consistency and unbiasedness, typically require strong assumptions about unobservable error terms. As noted in Carroll and Ruppert (1996) in a univariate orthogonal regression there are more parameters to estimate than sufficient statistics available in the sample which requires imposing an assumption on the ratio of the unknown error variances. The need for a normalizing assumption carries over to the multivariate case. Gleser (1981) provides a thorough treatment of the consistency properties of TLS under the assumption that the error variances in all variables (dependent and explanatory) are equal and homoscedastic. If this assumption does not hold, he emphasizes (p. 43) that no strongly consistent estimator exists. Consequently little is known about bias in multivariate TLS applications with unknown noise variances.

The purpose of this study is to explore coefficient bias properties for a full range of correlation levels among explanatory variables using a Monte Carlo analysis. A two-variable EIV model is presented in which the regressors are allowed to be correlated and the ratio of the variances on dependent and independent variables is allowed to vary. The bias pattern in OLS follows the expected downward pattern when the true slope coefficient is positive but when the true coefficient is zero the bias follows the sign of the correlation between explanatory variables. In the TLS case biases are generally quite large and tend to be positive, especially when the explanatory variables are negatively correlated. Implications for interpreting results from optimal fingerprinting are discussed.

## 2. Monte Carlo Simulation

In what follows a bold-face letter (e.g. **X**) denotes a vector or matrix, a variable with a single numerical subscript (e.g. $x_1$) denotes a column of a matrix and a lower-case variable (e.g. $x_{i,j}$) with two subscripts *i,j* refer to the *i,j*-th element of the corresponding matrix. We are interested in estimating a linear model in which a dependent variable **y** is regressed on explanatory variables $x_1, x_2$ where the sample size is *N.* The model is thus

$$\mathbf{y} = \mathbf{Xb} + \mathbf{e} \qquad [1]$$

where $\mathbf{e}$ is a homoscedastic random variable with $E(\mathbf{e}|\mathbf{X}) = 0$. Assume we cannot observe $\mathbf{X}$ directly, instead we observe $\mathbf{W} = \mathbf{X} + \mathbf{U}$ with elements $w_{i,j}$ where $\mathbf{U}$ is an $N \times 2$ matrix of column-wise zero-mean error terms. The OLS estimator $\widehat{\mathbf{b}}_{OLS} = (\mathbf{W}^{\mathbf{T}}\mathbf{W})^{-\mathbf{1}}\mathbf{W}^{\mathbf{T}}\mathbf{y}$ is a biased and inconsistent estimator of $\mathbf{b}$ (Davidson and MacKinnon 2004, p. 313). TLS regression coefficients are found using the method of Markovsky and Van Huffel (2007). If the singular value decomposition of the $N \times 3$ matrix $[\mathbf{W}\ \mathbf{y}]$ is denoted $\mathbf{V\Sigma H}$ and the final column of $\mathbf{H}$ is denoted $h_3$ then the TLS estimate $\hat{\mathbf{b}}_{TLS}$ is:

$$\hat{\mathbf{b}}_{TLS} = \left[-\frac{h_{1,3}}{h_{3,3}}, -\frac{h_{2,3}}{h_{3,3}}\right]^T \qquad [2]$$

The Gauss-Markov theorem implies that if $\mathbf{e}$ in equation [1] satisfies the classical assumptions (Koop 2008), which include the absence of randomness in $\mathbf{W}$, $\widehat{\mathbf{b}}_{OLS}$ is the most efficient estimator of $\mathbf{b}$. When $\mathbf{W}$ is random TLS is supposed to trade off efficiency for a reduction in bias, but as we will see bias in most cases increases.

We use $N = 200$. In optimal fingerprinting applications the sample sizes are typically much smaller than this (Allen and Tett 2003, Jones et al. 2016) so it is a conservative parameter selection. The first simulated *x* variable is constructed as a uniform random draw $x_{1i} = v_{1i}$ where $i = 1, \dots, N$ and $v_{1i}$ are draws from a uniform distribution with bounds $\pm\sqrt{12}/2$ in order to yield an expected variance of 1. The second *x* variable is defined using $x_{2i} = v_{2i} + cx_{1i}$ where $c$ is a constant that induces correlation between the *x*'s and will vary between -

0.9 and +0.9. The covariance between $x_1$ and $x_2$ is $E(x_1 v_2) + c\sigma_{x1}^2$ where $\sigma_{x1}^2$ is the variance of $x_1$.

Define three error vectors $u_j \sim N(0,1), j = (1, 2, \; y)$, representing independent zero-mean Gaussian noise terms associated with, respectively, the two explanatory variables and the dependent variable, having respective associated variances $\sigma_{x1}^2, \sigma_{x2}^2$ and $\sigma_y^2$. We observe the noise-contaminated explanatory variables $w_j = x_j + u_j s$ $(j = 1,2)$ where $s$ is a parameter we will use to scale the variance of $u_j$ relative to that of $x_j$. All variables are zero-centered then we construct pseudo-observations $y_i$ using:

$$y_i = \beta x_{1i} + x_{2i} + u_{iy} \qquad [3]$$

using an assumed true value of $\beta$ which will be, in turn, 0.0 or 1.0. Note that equation [3] implies there is no omitted variables bias or model error (in the sense of Fuller 1987 and Carroll and Ruppert 1996) the presence of which is known to cause TLS to overcorrect attenuation bias.

Estimation is done by applying both OLS and TLS to the regression model

$$y_i = \beta w_{i1} + \alpha w_{i2} + e_i \qquad [4]$$

yielding, respectively, $\hat{\beta}_{OLS}$ and $\hat{\beta}_{TLS}$. The constant term is omitted. We will ignore $\hat{\alpha}$ and confine the discussion to the values of $\hat{\beta}$. The simulations were run for 19 values of $c$

running sequentially from -0.9 to 0.9 in steps of 0.1, 10 values of $s$ running from 0.0 to 1.8 in steps of 0.2 and, within each parameter pair, 500 repetitions were run to obtain the mean values of $\hat{\beta}_{OLS}$ and $\hat{\beta}_{TLS}$. The values of $c$ span the range from highly anticorrelated to highly correlated. The values of $s$ determine the ratio of the variance of the error term on the $x$'s relative to that on $y$. When $s = 0$ the $x$'s are measured without error and OLS can be expected to be unbiased. When $s = 1$ the variances are equal which corresponds to the cases in Gleser (1981) and elsewhere in which TLS is unbiased. When $s > 1$ the explanatory variables are noisier than the dependent variable. In practice, while a researcher can estimate the correlation $r_w$ of $w_1$ and $w_2$ and thereby infer the likely value of $c$, without a measurement of the variances it will be unknown which value of $s$ best describes the regression being run.

All simulations were done using R version 4.0.2 (R Core Team 2020). The code file that generates all results shown herein is in the Supplementary Information file accompanying this paper.

## 3. Results

### *3.1 True value of $\beta = 0$*

The results for $\beta = 0$ are shown in Figure 1 and Tables 1—3. Table 1 reports the mean estimated values of $\hat{\beta}_{OLS}$ and Table 2 reports the same for $\hat{\beta}_{TLS}$. Since true $\beta = 0$ the table entries are all estimates of coefficient bias. Table 3 reports the correlations between $w_1$ and $w_2$, denoted $r_w$, associated with each pair of $c$ and $s$ coefficients.

The lines in Figure 1 are colour-coded based on the value of $s$. Red indicates $s = 1$, implying the noise variance on the $x$'s matches that on $\boldsymbol{y}$. The range from gray to black corresponds to $s$ going towards zero, which corresponds to the explanatory variable error terms disappearing. $s > 1$ is coded with the coloured lines with blue representing the maximum value of 1.8. Starting with OLS results (right panel) we see that when the signals are negatively correlated, corresponding with $c < 0$, the coefficient bias is uniformly negative, and vice versa. Note that the standard EIV attenuation bias is multiplicative so when true $\beta = 0$ OLS is unbiased in the univariate case. Here we see that OLS is also unbiased in the multivariate case but only when the $x$'s are uncorrelated. (OLS is also unbiased when $s = 0$ but this is the trivial case because the $x$'s are non-random.) When the $x$'s are positively correlated OLS is positively biased. Regardless of the sign of $c$ Table 1 shows that the magnitude of the bias is maximized when the noise ratio $s$ is about 0.6-0.8 which in the diagram is near the red line. It shrinks as $s \to 0$ (black line) as expected since the noise component of the $x$'s disappears.

The TLS results in the left panel look very different in several respects. Only when the $x$'s are uncorrelated ($c = 0$) is TLS unbiased. When the $x$'s are negatively correlated the coefficient bias is generally positive and can be extremely large. Moreover it gets worse rather than better as $s$ approaches 0 (the line shading goes from gray to black). Likewise when $c > 0$ the bias is generally negative although unstable and gets larger as $s \to 0$. When $s = 1$ (red line) the bias stays near zero when $c$ falls in the range $[0, 0.5]$ (see Table 2) which, from Table 3, can be seen to correspond to values of $r_w$ in the range $[0, 0.3]$. Outside of that interval the bias is unpredictable. Note, however, that the entire shape of the red line was unstable upon repetitions with different random number seeds so the profile shown is not consistently observed. When $s > 1$ as it gets larger the bias pattern begins to resemble that for OLS but as shown in Tables 1 and 2 it is uniformly larger in magnitude.

*3.2 True value of $\beta = 1$*

Since $\beta$ no longer equals 0 OLS can be expected to exhibit attenuation bias. The results are shown in Figure 2 and Tables S1 and S2 in the supplement. The correlation values from Table 3 are nearly identical and are not repeated. Looking at the right panel in Figure 2, When $s = 0$ OLS is, of course, unbiased, but as $s$ gets larger the downward bias grows, and also gets worse as $c$ declines. In all cases the coefficient estimates remain between 0 and 1.

Looking at the left panel the TLS results are pretty dismal. When $s = 1$ and $c$ lies within $[-0.4, 0.2]$ the bias is relatively small, although again the shape of the red line can change dramatically upon repetition. As $s$ goes to zero the bias quickly becomes large and positive for $c < 0$ and of indeterminate sign for $c > 0$. When $s > 1$ the bias pattern is more stable and is uniformly negative, resembling the OLS pattern but with smaller magnitudes for most values of $c$.

For both cases examined herein, it is unreasonable to assume in practice that $c$ will be zero but there is an argument for assuming $s = 1$. It can be shown (Gleser 1981) that if $s \neq 1$ but its value is known the model [1] can be transformed into another form in which the noise variances are equalized and the biases would therefore correspond to those shown in red in Figures 1 and 2. The desired slope coefficients can be recovered using an inverse operation. However, if *s* is unknown this remedy is unavailable.

*3.3 Application: Optimal Fingerprinting*

TLS is widely-used as part of the optimal fingerprinting or signal detection methodology (Allen and Tett 2003, DelSole et al. 2019). The dependent variable is a measurement of a climate pattern to be explained, such as a vector of observed temperature trends in spatial gridcells over the Earth's surface. The explanatory variables are climate model-generated analogues or "signals" for the same time interval with the model run under different assumptions such as anthropogenic greenhouse gas (GHG) forcing only and natural (NAT) forcing only. A pre-whitening operator is applied to remove heteroskedasticity associated with spatial patterns of natural climate variability, which we assume has been done herein. If the regression coefficient associated with a signal is significantly greater than zero the signal is said to have been "detected" and if it is not significantly different from unity the result is said to have passed a model consistency test. Consequently the estimation of coefficients in the [0, 1] interval is deemed to be of considerable scientific interest in attributing observed climate change to GHGs. Use of two explanatory patterns in a fingerprinting regression, such as GHG and NAT, is typical, although some authors have attempted to identify three signals at a time (e.g. Jones et al 2016).

The rationale for assuming EIV is that the GHG and NAT signals are generated by climate models that have internal representations of large-scale weather systems and any run of such a model will have sampling noise, in the sense that the same model re-run with nearly identical initial conditions would generate slightly different results. Consequently while the model output is observed without error, it is a potentially noisy observation of the "true" underlying signal.

Correlations between model-generated climate signals are not typically discussed or reported in optimal fingerprinting applications. Neither the magnitude of the error variance on the signals from a single climate model run nor its relative magnitude ($s$) to $\sigma_y^2$ can be estimated directly so there is no assurance that $s = 1$. Moreover it is common in optimal fingerprinting applications to use multiple climate models and compare results from individual models to those using the ensemble average signals so even if it were the case that $s = 1$ for all individual model regressions, it would shrink towards zero for the ensemble average.

We obtained forcing patterns (anthropogenic GHG and natural) from nine climate models archived as part of the Fifth Coupled Model Intercomparison Project (CMIP5) archive and were taken as-is from the Koninklijk Nederlands Meteorologisch Instituut Climate Data Explorer site (van Oldenborg, 2016). The signals were defined as linear temperature trends over the 1950-2005 interval by grid cell associated with each forcing pattern, and for each model the correlation between the GHG and natural forcing signal patterns were computed. All nine correlations were negative with magnitudes ranging from about -0.2 to -0.9. This approximately corresponds with the upper left quadrant of the left panel of Figure 1 although not all combinations of $s$ and $c$ are compatible with these values of $r_w$. Figure 3 redraws this section with the appropriate truncations and all values mapped against corresponding $r_w$ values. In the region $-0.4 < r_w < -0.2$ when $s = 1$ (red line) the bias values are positive, they become negative if $s$ increases and positive if $s$ decreases. For $r_w < -0.4$ the bias is negative if $s = 1$, it becomes very erratic if $s = 0.8$ and is uniformly positive and generally >1 if $s < 0.8$.

As a specific example suppose true $\beta = 0$ and $r_w = -0.4$. Table 3 indicates that $s$ must therefore be less than 1.2. If $s = 1$ then $c$ must be about -0.8 which according to Table 2 is associated with a bias of about -0.2. If $s = 0.8$ then $c$ must be about -0.6 and the bias is about +1.2. If $s = 0.6$ then $c$ must be about -0.4 and the bias is about +1.3. If $s = 0.2$ then $c \cong 0.25$ and the bias equals about +1.0. These are the values indicated in Figure 3 for $r_w = -0.4$.

It is often supposed that ensemble averaging allows a climate signal to emerge more strongly from background noise. The results in Jones et al. (2016) illustrate this: TLS-based fingerprinting coefficients from 15 individual climate models are very unstable and do not, as a group, yield a clear conclusion about the detectability of GHG's on the climate, whereas averaging the model-generated signals yields a GHG coefficient close to 1.0, supporting an inference of causal detection. However the problem revealed by the present analysis is that when model-generated signals are averaged together, since $\sigma_y^2$ remains constant $s \to 0$ and the bias pattern converges towards the black lines shown in Figures 1—3. A signal coefficient near 1.0 when $s$ is known to be close to zero is consistent with a true $\beta = 0$.

For the purpose of diagnosing likely bias in TLS regressions It is useful to compute $r_w$ and to compare TLS and OLS coefficient values. If $r_w < 0$ and the TLS and OLS coefficient estimates are both negative it is likely the case that true $\beta = 0$ and $s \geq 1$ in which case OLS is the preferred estimator since its downward bias is smaller. If $\hat{\beta}_{TLS} > 0$ and $\hat{\beta}_{OLS} < 0$ it is still likely that true $\beta = 0$, $s < 1$ and TLS exhibits a potentially large upward bias, implying OLS is again the preferred estimator. If both coefficients are between 0 and 1 for individual model signals while for the ensemble average the OLS coefficient → 1 from below while the TLS coefficient exceeds 1 and potentially becomes large, this indicates the likely true value of $\beta > 0$ and while TLS is probably valid for individual model estimate, OLS would be better for the ensemble mean.

Although OLS turns out often to be preferred when compared to TLS, in general the researcher should consider using an Instrumental Variables estimator since it provides a consistent estimator in the presence of EIV (Davidson and MacKinnon 2004).

## 4. Conclusion

A Monte Carlo analysis allowing explanatory variables to be correlated and variances to differ shows serious potential problems with TLS as compared to OLS. OLS exhibits the expected attenuation bias but TLS coefficients are typically biased even more and exhibit extreme instability depending on the correlation of the explanatory variables. If the explanatory variables are negatively correlated TLS has the particularly undesirable property that as the EIV problem declines (the noise variance on the *x* variables → 0) the positive bias gets larger. Practitioners of TLS should always report the correlation of the explanatory variables and use a comparison with OLS to assess the nature of any bias that may be present with TLS. In the absence of such diagnostics an apparently positive slope coefficient in a TLS regression is not particularly meaningful since it can easily arise even when the true value of the coefficient is zero.

**Declarations**

Funding: The author did not receive support from any organization for the submitted work.

Conflict of interest/Competing interests: The author is a Senior Fellow of the Fraser Institute and a member of the Academic Advisory Council of the Global Warming Policy Foundation. Neither organization had any knowledge of, involvement with or input into this research. The author has provided paid or unpaid advisory services to various government entities and to private sector entities in the law, manufacturing, distilling, communications, policy analysis and technology sectors. None of these entities had any knowledge of, involvement with or input into this research.

Availability of data and material: not applicable

Code availability: Submitted with manuscript in the Supplementary Information

Author's contribution: 100%

## 7. References


Allen, M.R. and P.A. Stott (2003) Estimating Signal Amplitudes in Optimal Finger-Printing, Part I: Theory. *Climate Dynamics* 21:477—491. DOI 10.1007/s00382-003-0313-9

Carroll, R.J. and David Ruppert (1996) The Use and Misuse of Orthogonal Regression in Linear Errors-in-Variables Models. *The American Statistician* 50(1) 1—6.

Davidson, Russell and James MacKinnon (2004) *Econometric Theory and Methods*. New York: Oxford University Press.

DelSole, T., L. Trenary, X. Yan and M.K. Tippett (2019) Confidence intervals in optimal fingerprinting. *Climate Dynamics* 52:4111—4126. https://doi.org/10.1007/s00382-018-4356-3

Fuller, W (1987) *Measurement Error Models*. New York: John Wiley and Sons.

Gillett, N. P., Arora, V. K., Matthews, H. D. & Allen, M. R. Constraining the ratio of global warming to cumulative CO2 emissions using CMIP5 simulations. *Journal of Climate* 26, 6844–6858 (2013)

Gleser, Leon J. (1981) "Estimation in a Multivariate "Errors in Variables" Regression Model: Large Sample Results." *The Annals of Statistics* Vol. 9, No. 1 (Jan., 1981), pp. 24-44 https://www.jstor.org/stable/2240867

Jones, Gareth, Peter A. Stott and John F.B. Mitchell (2016) Uncertainties in the Attribution of Greenhouse Gas Warming and Implications for Climate Prediction. *Journal of Geophysical Research-Atmospheres* 121(12) 6969—6992 https://doi.org/10.1002/2015JD024337.

Koop, Gary (2008) *Introduction to Econometrics*. Chichester: Wiley.

Markovsky, I. and S. Van Huffel (2007) Overview of total least squares methods. *Signal Processing*, vol. 87, pp. 2283–2302, 2007 https://doi.org/10.1016/j.sigpro.2007.04.004

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

van Oldenborgh, G.J., (2016). Climate Data Explorer. Data available at http://climexp.knmi.nl/selectfield_co2.cgi?someone@somewhere

Wooldridge, Jeffrey (2019) *Introductory Econometrics: A Modern Approach 7th ed.* Boston: Cengage.

TABLES

| | Noise scaling ($s$) on $x$'s | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ***c*** | **0** | **0.2** | **0.4** | **0.6** | **0.8** | **1** | **1.2** | **1.4** | **1.6** | **1.8** |
| **-0.9** | -0.007 | -0.110 | -0.231 | -0.295 | -0.293 | -0.270 | -0.245 | -0.216 | -0.187 | -0.173 |
| **-0.8** | -0.010 | -0.098 | -0.227 | -0.273 | -0.274 | -0.259 | -0.225 | -0.200 | -0.175 | -0.156 |
| **-0.7** | 0.007 | -0.078 | -0.203 | -0.251 | -0.252 | -0.235 | -0.210 | -0.179 | -0.161 | -0.138 |
| **-0.6** | 0.007 | -0.076 | -0.174 | -0.223 | -0.224 | -0.208 | -0.187 | -0.161 | -0.143 | -0.122 |
| **-0.5** | 0.000 | -0.057 | -0.156 | -0.196 | -0.196 | -0.186 | -0.158 | -0.138 | -0.122 | -0.105 |
| **-0.4** | -0.006 | -0.049 | -0.134 | -0.163 | -0.163 | -0.151 | -0.133 | -0.117 | -0.098 | -0.084 |
| **-0.3** | 0.000 | -0.042 | -0.101 | -0.130 | -0.125 | -0.113 | -0.104 | -0.086 | -0.076 | -0.065 |
| **-0.2** | 0.000 | -0.027 | -0.061 | -0.086 | -0.088 | -0.077 | -0.069 | -0.060 | -0.048 | -0.043 |
| **-0.1** | 0.001 | -0.011 | -0.034 | -0.041 | -0.043 | -0.043 | -0.036 | -0.034 | -0.025 | -0.020 |
| **0** | 0.001 | 0.003 | -0.001 | 0.000 | 0.001 | 0.002 | -0.002 | 0.001 | 0.000 | 0.002 |
| **0.1** | -0.003 | 0.014 | 0.032 | 0.047 | 0.044 | 0.037 | 0.035 | 0.033 | 0.025 | 0.021 |
| **0.2** | 0.002 | 0.035 | 0.062 | 0.087 | 0.087 | 0.079 | 0.068 | 0.059 | 0.050 | 0.044 |
| **0.3** | 0.001 | 0.040 | 0.098 | 0.124 | 0.126 | 0.118 | 0.104 | 0.086 | 0.073 | 0.067 |
| **0.4** | -0.009 | 0.050 | 0.127 | 0.169 | 0.166 | 0.153 | 0.134 | 0.114 | 0.099 | 0.084 |
| **0.5** | -0.002 | 0.062 | 0.156 | 0.198 | 0.200 | 0.185 | 0.164 | 0.142 | 0.118 | 0.101 |
| **0.6** | -0.012 | 0.069 | 0.175 | 0.223 | 0.225 | 0.206 | 0.186 | 0.161 | 0.143 | 0.122 |
| **0.7** | -0.006 | 0.086 | 0.208 | 0.255 | 0.252 | 0.233 | 0.211 | 0.181 | 0.163 | 0.139 |
| **0.8** | 0.005 | 0.103 | 0.221 | 0.276 | 0.269 | 0.257 | 0.232 | 0.203 | 0.174 | 0.155 |
| **0.9** | 0.001 | 0.107 | 0.235 | 0.286 | 0.299 | 0.268 | 0.248 | 0.217 | 0.193 | 0.166 |

**Table 1:** Estimated value of $\beta_{OLS}$ when true $\beta = 0$.

| | **Noise scaling ($s$) on $x$'s** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ***c*** | **0** | **0.2** | **0.4** | **0.6** | **0.8** | **1** | **1.2** | **1.4** | **1.6** | **1.8** |
| **-0.9** | 7.933 | 7.282 | 6.682 | 396.725 | 3.759 | -1.407 | -0.192 | -0.297 | -0.256 | -0.239 |
| **-0.8** | 5.680 | 5.666 | 6.220 | 1.339 | -2.383 | -0.211 | -0.247 | -0.269 | -0.245 | -0.217 |
| **-0.7** | 4.375 | 4.220 | 3.764 | 3.394 | 4.238 | 0.165 | -0.214 | -0.241 | -0.226 | -0.190 |
| **-0.6** | 3.393 | 3.207 | 3.035 | 2.500 | 1.153 | 0.065 | -0.208 | -0.224 | -0.200 | -0.168 |
| **-0.5** | 2.592 | 2.496 | 2.215 | 1.970 | -3.352 | 0.385 | -0.180 | -0.191 | -0.172 | -0.144 |
| **-0.4** | 1.952 | 1.827 | 1.602 | 1.308 | 0.695 | 0.036 | -0.153 | -0.166 | -0.137 | -0.115 |
| **-0.3** | 1.350 | 1.276 | 1.111 | 0.829 | 6.235 | 0.085 | -0.120 | -0.121 | -0.107 | -0.089 |
| **-0.2** | 0.833 | 0.828 | 0.713 | 0.510 | 0.253 | -0.019 | -0.085 | -0.086 | -0.068 | -0.059 |
| **-0.1** | 0.402 | 0.392 | 0.344 | 0.246 | 0.136 | 0.015 | -0.042 | -0.050 | -0.036 | -0.027 |
| **0** | 0.004 | -0.001 | -0.001 | -0.006 | 0.010 | 0.004 | -0.004 | 0.002 | 0.000 | 0.002 |
| **0.1** | -0.421 | -0.397 | -0.355 | -0.241 | -0.126 | -0.005 | 0.037 | 0.049 | 0.034 | 0.028 |
| **0.2** | -0.858 | -0.813 | -0.729 | -0.529 | -0.276 | -0.009 | 0.080 | 0.082 | 0.071 | 0.059 |
| **0.3** | -1.321 | -1.290 | -1.139 | -0.777 | -0.495 | -0.007 | 0.140 | 0.119 | 0.104 | 0.092 |
| **0.4** | -1.880 | -1.853 | -1.678 | -1.280 | -0.659 | -0.062 | 0.159 | 0.163 | 0.140 | 0.115 |
| **0.5** | -2.580 | -2.442 | -2.233 | -1.632 | -0.668 | 0.063 | 0.187 | 0.200 | 0.165 | 0.137 |
| **0.6** | -3.304 | -3.283 | -3.010 | -1.798 | -1.577 | -0.067 | 0.204 | 0.220 | 0.202 | 0.166 |
| **0.7** | -4.319 | -4.323 | -5.892 | -3.102 | -1.747 | -0.064 | 0.219 | 0.247 | 0.231 | 0.191 |
| **0.8** | -5.734 | -6.358 | -6.587 | -4.462 | -2.496 | 4.852 | 0.263 | 0.276 | 0.241 | 0.213 |
| **0.9** | -7.422 | -7.883 | -6.781 | -4.934 | -1.446 | 1.432 | 0.293 | 0.285 | 0.269 | 0.226 |

**Table 2:** Estimated value of $\beta_{TLS}$ when true $\beta = 0$.

| | Noise scaling ($s$) on $x$'s | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ***c*** | **0** | **0.2** | **0.4** | **0.6** | **0.8** | **1** | **1.2** | **1.4** | **1.6** | **1.8** |
| **-0.9** | -0.87 | -0.84 | -0.76 | -0.65 | -0.54 | -0.45 | -0.36 | -0.30 | -0.25 | -0.21 |
| **-0.8** | -0.85 | -0.82 | -0.73 | -0.61 | -0.50 | -0.41 | -0.34 | -0.28 | -0.23 | -0.19 |
| **-0.7** | -0.81 | -0.78 | -0.69 | -0.57 | -0.46 | -0.38 | -0.31 | -0.25 | -0.21 | -0.17 |
| **-0.6** | -0.77 | -0.73 | -0.63 | -0.52 | -0.41 | -0.34 | -0.27 | -0.22 | -0.18 | -0.15 |
| **-0.5** | -0.71 | -0.67 | -0.57 | -0.47 | -0.36 | -0.29 | -0.23 | -0.18 | -0.15 | -0.12 |
| **-0.4** | -0.62 | -0.58 | -0.49 | -0.39 | -0.31 | -0.23 | -0.18 | -0.15 | -0.12 | -0.10 |
| **-0.3** | -0.51 | -0.48 | -0.39 | -0.30 | -0.24 | -0.19 | -0.14 | -0.12 | -0.09 | -0.07 |
| **-0.2** | -0.37 | -0.34 | -0.28 | -0.21 | -0.16 | -0.12 | -0.09 | -0.08 | -0.06 | -0.05 |
| **-0.1** | -0.19 | -0.18 | -0.14 | -0.10 | -0.08 | -0.07 | -0.05 | -0.04 | -0.03 | -0.02 |
| **0** | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| **0.1** | 0.20 | 0.18 | 0.14 | 0.11 | 0.08 | 0.06 | 0.06 | 0.04 | 0.03 | 0.03 |
| **0.2** | 0.37 | 0.34 | 0.28 | 0.21 | 0.16 | 0.12 | 0.10 | 0.08 | 0.07 | 0.05 |
| **0.3** | 0.51 | 0.48 | 0.40 | 0.30 | 0.24 | 0.18 | 0.14 | 0.11 | 0.09 | 0.07 |
| **0.4** | 0.62 | 0.59 | 0.50 | 0.39 | 0.31 | 0.24 | 0.18 | 0.15 | 0.12 | 0.10 |
| **0.5** | 0.71 | 0.67 | 0.57 | 0.46 | 0.37 | 0.29 | 0.24 | 0.19 | 0.15 | 0.13 |
| **0.6** | 0.77 | 0.73 | 0.64 | 0.52 | 0.42 | 0.33 | 0.27 | 0.22 | 0.18 | 0.15 |
| **0.7** | 0.81 | 0.78 | 0.69 | 0.57 | 0.47 | 0.37 | 0.30 | 0.24 | 0.20 | 0.17 |
| **0.8** | 0.85 | 0.82 | 0.73 | 0.61 | 0.51 | 0.41 | 0.34 | 0.28 | 0.22 | 0.19 |
| **0.9** | 0.87 | 0.84 | 0.76 | 0.65 | 0.54 | 0.45 | 0.36 | 0.31 | 0.25 | 0.21 |

**Table 3:** Correlations between $w_1$ and $w_2$ for indicated values of $s$ and $c$.

FIGURES

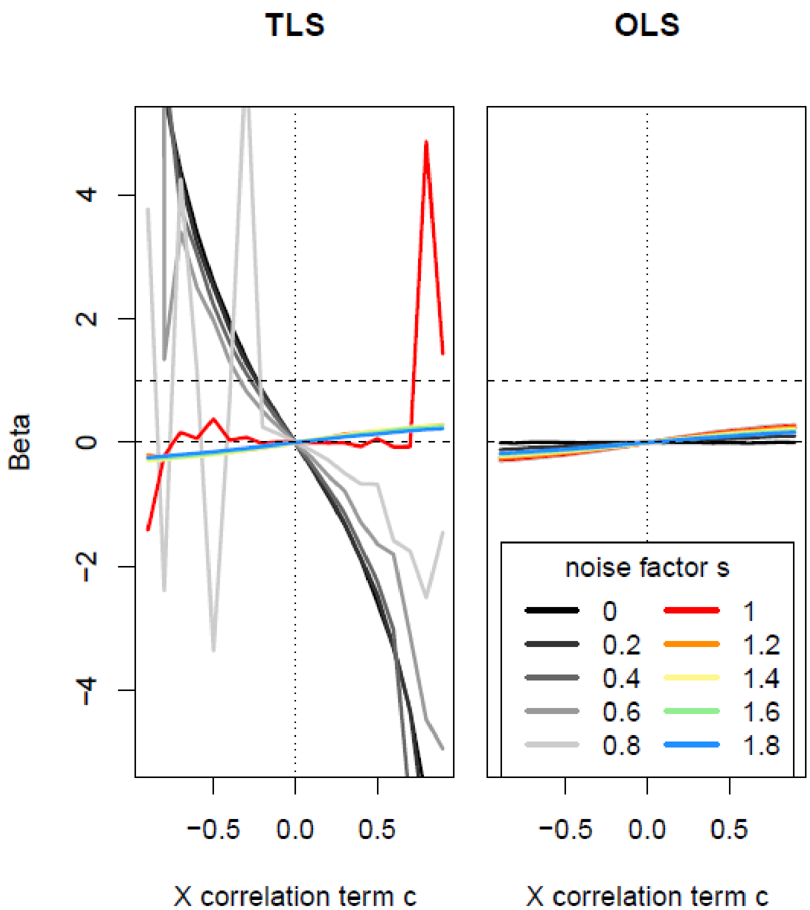


**Figure 1.** Mean estimated values of $\hat{\beta}$ using TLS (left) or OLS (right) when the true value of $\beta = 0$.

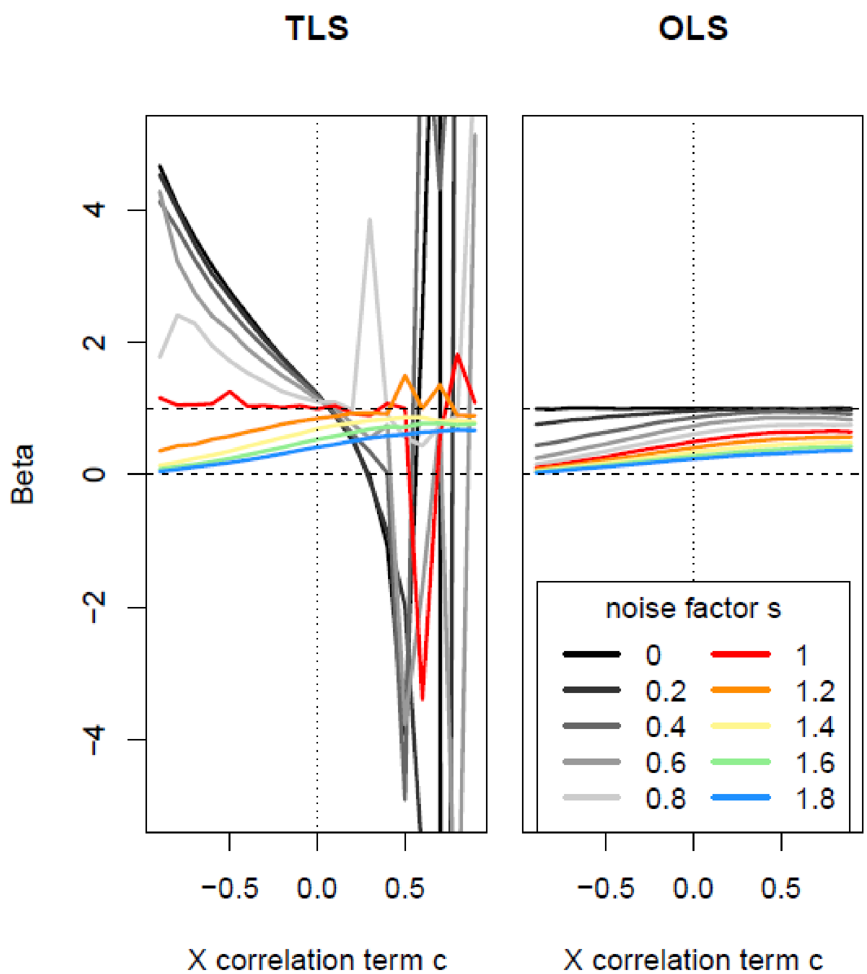


**Figure 2.** Mean estimated values of $\hat{\beta}$ using TLS (left) or OLS (right) when the true value of $\beta = 1$.

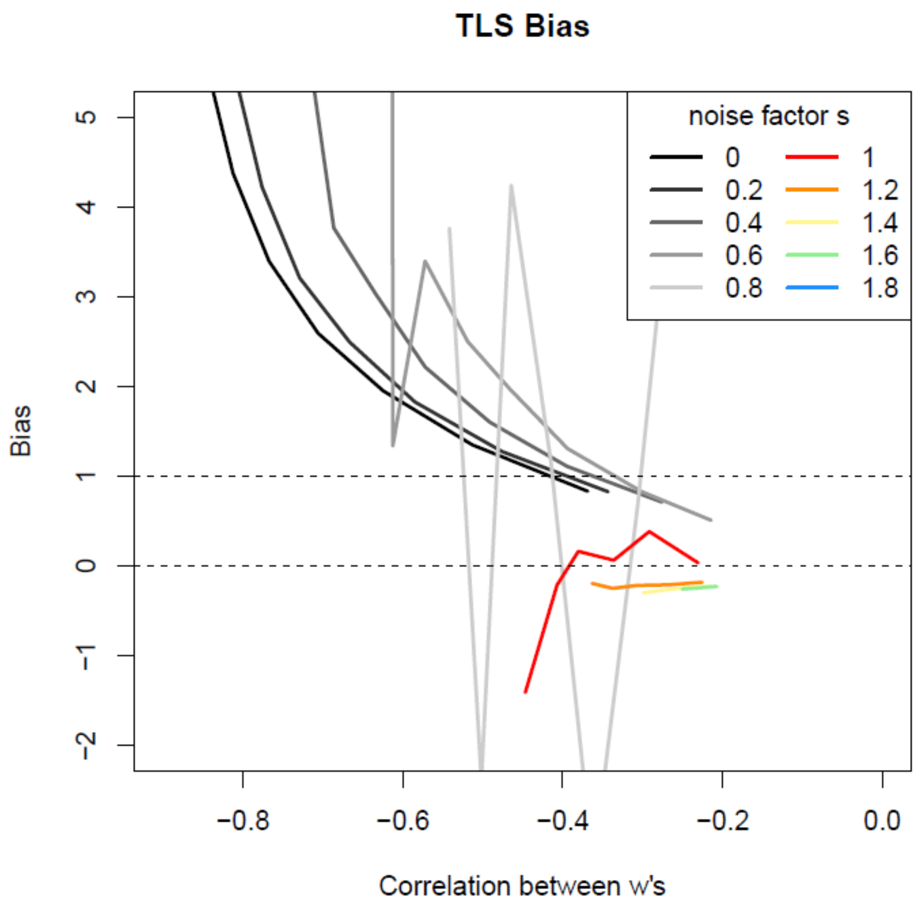


**Figure 3.** Bias of TLS estimator in region ($-0.9 < r_w < -0.2$) when true $\beta = 0$.