Computational Methods in Econometrics: Assignment 2

By Group 15

Lanlan Hou	2801069
Rick van der Ploeg	2782774
Sebastiaan van Dijen	2769505
Wa Chak Sou	2796840

School of Business and Economics

17th October 2025

INTRODUCTION

In this report, the time series $\{Y_t\}$ represents the average annual temperature at the well-known The Royal Netherlands Meteorological Institute (KNMI) weather station located in De Bilt (near Utrecht) in the Netherlands from 1900 to 2014. The following linear regression model is considered:

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t, \quad t = 1, 2, \ldots, n \tag{0.1}$$

In this model, the time series is decomposed into a linear trend with intercept β_0 and slope β_1 and an error process $\{\varepsilon_t\}$. To test for the presence of a significant upward trend, we consider the pair of hypotheses $H_0 : \beta_1 = 0$ vs. $H_1 : \beta_1 \geq 0$. With the use of t-statistic

$$T_n(Y) = \frac{\hat{\beta}_1}{\sqrt{\widehat{\mathrm{Var}(\beta_1)}}}$$

where β_1 is the OLS estimator of β_1 and $\widehat{\mathrm{Var}(\beta_1)}$ denotes a consistent estimator of the variance of β_1. The significance level is given by α.

The assumption of the error term $\varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$ is very restrictive in practice and is likely not satisfied in the temperature data. The residuals from the trend estimation are likely to suffer from serial correlation and/or heteroskedasticity. In case of heteroskedasticity and autocorrelation, the famous HAC standard errors can be used, which are proposed by Newey & West, are robust to both correlation and heteroskedasticity.

Furthermore, several inference methods that are robust to either serial correlation or heteroskedasticity are performed by performing the bootstrap t-test using the four bootstrap methods: *Nonparametric residual bootstrap (i.i.d. bootstrap)*, *Block bootstrap*, *Wild bootstrap (using Standard Normal and Rademacher distribution)*, *Sieve bootstrap*.

Scientific evidence indicates that there is seasonal variation in the warming trends. Research has shown that winters tend to get warmer faster than summers. Given the yearly winter and summer averages, the bootstrap t-tests are performed to and give further indication for the question that *'Do the warming rates is faster in winter than in summer?'*.

Instead of performing t-test, a two-sided version of confidence intervals can be performed around the estimated coefficients. Bootstrap methods are being used for the construction of confidence intervals. The four different confidence intervals of the yearly average temperatures for each bootstrap method are constructed: *Equal-tailed percentile intervals*, *Equal-tailed percentile-t intervals*, *Symmetric percentile intervals*, *Symmetric percentile-t intervals*.

In order to decide which bootstrap algorithm or confidence interval should be used in practice, the efficiency of the bootstrap t-test and the constructon of confidence intervals should be evaluated. The results are heavily depending on the chosen regression models and the error structures. As a result, investigate the empirical coverages of the selected bootstrap algorithms and confidence intervals using Monte Carlo simulation. Error terms are generated as selected specification, the empirical coverages of two different bootstrap methods, with two different types of confidence intervals around the estimated slope coefficient are checked.

EXERCISE 1

a. Estimation of model and performing the HAC t-test

We begin by estimating the linear trend model

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t, \qquad t = 1, \ldots, n, \tag{1}$$

where y_t denotes the average annual temperature at the KNMI weather station in De Bilt for the period 1900–2014. The model decomposes the observed time series into a deterministic linear trend and a stochastic error component ε_t.

The parameters are estimated by ordinary least squares (OLS):

$$\widehat{\boldsymbol{\beta}} = (X'X)^{-1}X'y, \qquad \widehat{\boldsymbol{\varepsilon}} = y - X\widehat{\boldsymbol{\beta}},$$

where X contains a column of $\mathbb{1}$ (for the intercept) and a time index t (for the trend).

To test whether the trend is significantly positive, we consider the hypotheses

$$H_0 : \beta_1 = 0 \qquad \text{vs.} \qquad H_1 : \beta_1 \geq 0.$$

Because temperature data often exhibit serial correlation and heteroskedasticity, we compute *heteroskedasticity and autocorrelation consistent* (HAC) standard errors following Newey & West. The corresponding test statistic is defined as

$$T_n(Y) = \frac{\widehat{\beta}_1 - \hat{\beta}_1^{(0)}}{\sqrt{\widehat{\text{Var}}(\hat{\beta}_1)}} = \frac{\widehat{\beta}_1}{\sqrt{\widehat{\text{Var}}(\hat{\beta}_1)}}, \tag{2}$$

where $\widehat{\text{Var}(\hat{\beta}_1)}$ is a consistent estimator of the variance of the slope coefficient. Hence, the denominator represents the *HAC standard error*

$$\text{SE}_{HAC}(\widehat{\beta}_1) = \sqrt{\widehat{\text{Var}}(\hat{\beta}_1)}.$$

Under the null hypothesis, the restricted coefficient $\hat{\beta}_1^{(0)}$ is equal to 0 and $T_n(Y)$ approximately follows a *Standard Normal distribution*. Applying the t-test to the De Bilt data yields an estimated t-statistic of

$$T_n = 6.035.$$

At the 10% significance level, the two-sided critical region using standard normal distribution is given by $[-1.645,\ 1.645]$. Since the test statistic T_n lies well outside this interval, we reject the null hypothesis of no trend.

Result:

$$t = 6.035 \notin [-1.645, 1.645] \quad \Rightarrow \quad \text{Reject } H_0.$$

Interpretation: The estimated slope coefficient $\widehat{\beta}_1$ is positive and highly significant, indicating a strong upward trend in the average annual temperature at De Bilt over the 1900–2014 period. This provides robust statistical evidence of long-term warming at this location, even when accounting for possible serial dependence and heteroskedasticity in the residuals.

b. Bootstrap-based t-tests

After the classical HAC-based inference, we perform the test using bootstrap methods to obtain p-values that are robust to potential deviations from the classical assumptions of homoskedasticity and independence. In particular, we apply the *Nonparametric Residual (i.i.d.) bootstrap*, *Wild bootstrap*, *Sieve bootstrap* and *Block bootstrap*.

(i) Nonparametric Residual Bootstrap (i.i.d. bootstrap)

The *i.i.d. bootstrap* assumes that the regression residuals are independent and identically distributed. The method resamples the residuals with replacement to generate new bootstrap samples of the dependent variable.

Algorithm (under $H_0 \colon \beta_1 = 0$):

1. Estimate model (0.1) by OLS and obtain residuals $\hat{\varepsilon}_t$.
2. Impose H_0 by setting $\hat{\beta}_1^{(0)} = 0$ and compute fitted values $\widehat{y}_t^{(0)} = \hat{\beta}_0$.
3. For each replication $b = 1, \ldots, B$ (with $B = 999$):
   (a) Draw residuals $\hat{\varepsilon}_t^{*(b)}$ i.i.d. with replacement from $\{\hat{\varepsilon}_t\}$.
   (b) Generate bootstrap sample $y_t^{*(b)} = \hat{y}_t^{(0)} + \hat{\varepsilon}_t^{*(b)}$.
   (c) Re-estimate the regression model on $y_t^{*(b)}$ and compute the adapted t-test statistic
   $$T_b^* = \frac{\hat{\beta}_{1,b}^*}{\sqrt{\widehat{\mathrm{Var}}(\hat{\beta}_{1,b}^*)}}.$$
   where $\sqrt{\widehat{\mathrm{Var}}(\hat{\beta}_{1,b}^*)}$ is the *classical* standard error, under the assumption of independent and identically distributed residuals.
4. The one-sided bootstrap p-value is estimated as
   $$\hat{p}_{\text{i.i.d.}} = \frac{1}{B} \sum_{b=1}^{B} \mathbb{1}\{T_b^* \geq T_n\}.$$

Result: The *i.i.d. bootstrap* yields a p-value of $\hat{p}_{\text{i.i.d.}} = 0.0$, confirming that the null hypothesis H_0 is strongly rejected. Even though the *i.i.d. bootstrap* reproduces the strong significance of the trend, it relies on the independence assumption, which is restrictive for climate data.

(ii) Wild Bootstrap

The *Wild bootstrap* is designed to handle heteroskedastic errors by multiplying residuals with random mean-zero, unit-variance weights ("multipliers"). The residuals will not be reshuffled. The residual at point t will determine the bootstrap error at point t.

Algorithm (under $H_0 \colon \beta_1 = 0$):

1. Estimate model (0.1) by OLS and obtain residuals $\hat{\varepsilon}_t$.
2. Impose H_0 and compute fitted values $\hat{y}_t^{(0)} = \hat{\beta}_0$.
3. For each replication $b = 1, \ldots, B$:

(a) Draw random multipliers $w_t^{(b)}$ from the *Standard Normal* distribution $w_t^{(b)} \sim \mathcal{N}(0,1)$, or, draw random multipliers $v_t^{(b)}$ from the *Rademacher* distribution $v_t^{(b)} = \begin{cases} +1, & \text{w.p. } \frac{1}{2}, \\ -1, & \text{w.p. } \frac{1}{2}. \end{cases}$

(b) Generate the bootstrap sample

$$y_t^{*(b)} = \hat{y}_t^{(0)} + w_t^{(b)} \hat{\varepsilon}_t, \quad \text{or,} \quad y_t^{*(b)} = \hat{y}_t^{(0)} + v_t^{(b)} \hat{\varepsilon}_t.$$

(c) Re-estimate the regression model on the generated data and compute adapted t-statistic

$$T_b^* = \frac{\hat{\beta}_{1,b}^*}{\sqrt{\widehat{\text{Var}_{HC}}(\widehat{\beta}_{1,b}^*)}}.$$

where $\sqrt{\widehat{\text{Var}}_{HC}\left(\widehat{\beta}_{1,b}^*\right)}$ is the *heteroskedasticity consistent* standard error, under the assumption that the residuals are independent but have unequal variances.

4. Compute the one-sided p-value as mentioned above for the *i.i.d. bootstrap*.

Result: The *Wild bootstrap* (using both *Standard Normal* and *Rademacher* multipliers) also produces $\hat{p}_{\text{wild}} = 0.0$. Hence, the null hypothesis of no upward trend is again rejected.

(iii) Block Bootstrap

The *Block bootstrap* method is designed to handle dependent data by dividing the residuals into blocks and then resampling the blocks. Within the blocks, the dependence pattern is completely preserved. While, between blocks, the dependence pattern is destroyed.

Algorithm (under $H_0: \beta_1 = 0$):

1. Estimate model (0.1) by OLS and obtain residuals $\hat{\varepsilon}_t$.
2. Impose H_0 and compute fitted values $\hat{y}_t^{(0)} = \hat{\beta}_0$.
3. Choose a block length L. That will result in having $n - L + 1$ blocks that look like:

$$B_1 = [\hat{\varepsilon}_1, ..., \hat{\varepsilon}_L] \quad B_2 = [\hat{\varepsilon}_2, ..., \hat{\varepsilon}_{L+1}] \quad \ldots \quad B_{n-L+1} = [\hat{\varepsilon}_{n-L+1}, ..., \hat{\varepsilon}_n]$$

4. For each replication $b = 1, \ldots, B$:

(a) Choose $\lceil \frac{n}{L} \rceil$ blocks randomly and draw $B_1^*, ..., B_{n-L+1}^*$ with replacement from $B_1, ..., B_{n-L+1}$.

(b) The bootstrap errors are obtained by: $(\varepsilon_1^*, ..., \varepsilon_n^*)$=$(B_1^*, ..., B_{n-L+1}^*)$.

(c) Generate bootstrap sample

$$y_t^{*(b)} = \hat{y}_t^{(0)} + \varepsilon_t^{*(b)}$$

(d) Re-estimate the regression model on the generated data and compute adapted t-statistic

$$T_b^* = \frac{\widehat{\beta}_{1,b}^*}{\sqrt{\widehat{\text{Var}_{HAC}}(\widehat{\beta}_{1,b}^*)}}.$$

where $\sqrt{\widehat{\text{Var}_{HAC}}(\widehat{\beta}_{1,b}^*)}$ is the *heteroskedasticity and autocorrelation consistent* standard error, under the assumption that the residuals may be heteroskedastic and serially correlated.

5. Compute the one-sided p-value as mentioned above for the *i.i.d. bootstrap*.

Result: The *Block bootstrap* (using block length $\lceil \frac{n}{10} \rceil$) also produces $\hat{p}_{\text{block}} = 0.005$. Even though this p-value is a bit higher then earlier, the null hypothesis of no upward trend is again rejected.

(iv) Sieve Bootstrap

The *Sieve bootstrap* method is designed to handle correlation in regression approximating their independence with an auto regressive process. Instead of assuming independent errors, the sieve bootstrap fits an AR(p) model to the residuals and generates new bootstrap samples from the estimated process and preserves temporal dependence.

Algorithm (under $H_0 : \beta_1 = 0$):

1. Estimate model (0.1) by OLS and obtain residuals $\hat{\varepsilon}_t$.
2. Impose H_0 and compute fitted values $\hat{y}_t^{(0)} = \hat{\beta}_0$.
3. Fit an AR(p) model to the residuals:
$$\hat{\varepsilon}_t = \phi_0 + \phi_1 \hat{\varepsilon}_{t-1} + \cdots + \phi_p \hat{\varepsilon}_{t-p} + \hat{\eta}_t,$$
where $\hat{\eta}_t$ are the estimated residuals.
4. For each replication $b = 1, \ldots, B$:
   (a) Resample residuals $\eta_t^{*(b)}$ with replacement from $\{\hat{\eta}_t\}$.
   (b) Generate new bootstrap innovations recursively:
   $$\varepsilon_t^{*(b)} = \phi_0 + \phi_1 \varepsilon_{t-1}^{*(b)} + \cdots + \phi_p \varepsilon_{t-p}^{*(b)} + \eta_t^{*(b)}.$$
   (c) Construct the bootstrap sample:
   $$y_t^{*(b)} = \hat{y}_t^{(0)} + \varepsilon_t^{*(b)}.$$
   (d) Re-estimate the regression model on the bootstrap sample and compute the adapted t-statistic:
   $$T_b^* = \frac{\hat{\beta}_{1,b}^*}{\sqrt{\widehat{\text{Var}_{HAC}(\hat{\beta}_{1,b}^*)}}},$$
   where $\sqrt{\widehat{\text{Var}_{HAC}(\hat{\beta}_{1,b}^*)}}$ is the *heteroskedasticity and autocorrelation consistent* standard error, under the assumption that the residuals may be heteroskedastic and serially correlated.
5. Compute the one-sided p-value as mentioned above for the *i.i.d. bootstrap*.

Result: The *Sieve bootstrap* (with default $p = 5$) yields a p-value of $\hat{p}_{\text{sieve}} = 0.038$, so the null hypothesis of no upward trend is rejected.

Method	Test Statistic / p-Value	Decision
HAC t-test	$t = 6.035$	Reject H_0
i.i.d. bootstrap	$p = 0.000$	Reject H_0
Wild bootstrap (Standard Normal)	$p = 0.000$	Reject H_0
Wild bootstrap (Rademacher)	$p = 0.000$	Reject H_0
Block bootstrap	$p = 0.002$	Reject H_0
Sieve bootstrap	$p = 0.081$	Reject H_0

Table 1.1: The results of HAC t-test and different bootstrap methods for testing β_1

c. Results and findings

Table 1.1 shows the results of testing the significant of trend using t-test and bootstrap t-test.The HAC t-test and most of the bootstrap methods yield highly significant results ($p \approx 0.000\check{~}0.002$). It indicates a strong positive warming trend. Only the *Sieve bootstrap*, which considers about the autocorrelation, gives a slightly higher p-value $= 0.081$. It suggests the significance at the 10% level. Overall, the results support consistently that reject the null hypothesis and confirms a statistically significant upward temperature trend.

EXERCISE 2

a. Bootstrap-Based t-tests with Seasonal Averages

We now extend the analysis to the *seasonal averages* in order to examine whether the previously identified trend also holds within individual seasons. Specifically, we estimate separate regression models for the *summer* and *winter* averages, allowing us to compare the strength and significance of temperature trends across different parts of the year. We will once again apply the same four bootstrap methods used for the annual data, the *i.i.d. residual bootstrap*, *wild bootstrap*, *sieve bootstrap*, and *block bootstrap*. Table 2.1 demonstrates the estimation of coefficents and the results of t-test and different bootstrap-based t-tests.

(i) Estimation of model and HAC t-test

First, estimate the model with both *summer* and *winter average* using OLS. Nest, compute the corresponding t-statistic using HAC standard error.

Results:

- *Summer:* $\beta_0 = 15.683,\ \beta_1 = 0.012,\ t(\beta_1) = 5.265 \Rightarrow$ Reject H_0.
- *Winter:* $\beta_0 = 2.038,\ \beta_1 = 0.012,\ t(\beta_1) = 2.723 \Rightarrow$ Reject H_0.

Interpretation. Both *summer* and *winter average* results a similar estimate slope coefficient 0.012. At significant level $\alpha = 10\%$, both t test statistic are outside the accepted region. Therefore, reject the null hypothesis for both seasonal averages.

(ii) Nonparametric Residual Bootstrap (i.i.d. bootstrap)

Following the same procedure as in Exercise 1b, we resample OLS residuals with replacement under $H_0 : \beta_1 = 0$ to obtain a bootstrap distribution of the t-statistic.

Results: *Summer:* $p = 0.0$ and *Winter:* $p = 0.012$.

Interpretation. The i.i.d. bootstrap confirms strong significance for both seasons; results are robust to non-normality/small-sample effects, though the i.i.d. assumption of independent errors may be restrictive for climate series.

(iii) Wild Bootstrap

We next apply the wild bootstrap, again following the procedure in Section 1b, to allow for heteroskedasticity by multiplying residuals with random mean-zero, unit-variance weights. The point estimates remain the same as above.

Results (Standard Normal): *Summer:* $p = 0.0$ and *Winter:* $p = 0.013$.
Results (Rademacher): *Summer:* $p = 0.0$ and *Winter:* $p = 0.014$.

Interpretation. Both seasons remain statistically significant under heteroskedasticity-robust resampling. The winter trend is *significant at the 5% level* with a wild-bootstrap $p \approx 3.7\%$, indicating somewhat weaker (but still meaningful) evidence compared with summer.

(iv) Block Bootstrap

To account for possible short-term dependence in the residuals, implement the block bootstrap by resampling consecutive blocks of observations rather than individual residuals. This approach preserves the local temporal correlation structure within each block while maintaining the overall time ordering of the data.

Results: *Summer:* $p = 0.003$ and *Winter:* $p = 0.032$.

Interpretation. Both seasonal trends remain statistically significant after accounting for short-term serial dependence. The slightly higher p-values compared with the i.i.d. and wild bootstraps suggest modestly greater uncertainty, but the positive warming trend persists in both seasons.

(v) Sieve Bootstrap

Apply the sieve bootstrap for possible serial correlation in the residuals, following the procedures mentioned above. The method fits an AR(p) model to the OLS residuals, resamples the estimated innovations, and regenerates bootstrap errors with dependency.

Results: *Summer:* $p = 0.036$ and *Winter:* $p = 0.096$.

Interpretation. When we correct for reasonable autocorrelation in the residuals using the sieve bootstrap, both seasonal estimates of the temperature trend remain statistically significant. The increase in p-values reflects the adjustment for dependence, but does not change the overall inference, which is the trend in both winter and summer temperatures persists under even more realistic time series sampling.

b. Results and findings

Table 2.1 illustrates the comparison of the estimated coefficients of the linear regression models, t-test results and corresponding bootstrap p-values for both the *winter* and *summer* in De Bilt. Both seasons tend to be positive and statistically significant for warming trends. The estimated increase of 0.012 °C in each year. However, the significance of the summer trend is slightly stronger than the winter trend. The p-values remain significantly small using all bootstrap methods for *summer average*, while the *winter average* results a $p = 0.096$ using the *Sieve bootstrap*. It may indicate for possible autocorrelation . The results reveal that both seasons have the similar warming rates, whereas the statistical evidence for winter is less robust. To conclude, there is no clear evidence that the warming rate is faster in winter based on the similar warming rates.

Statistics/Methods	Winter Average	Summer Average
$\hat{\beta}_0$	2.038	15.683
$\hat{\beta}_1$	0.012	0.012
t-statistic for $\hat{\beta}_1$	2.723	5.265
t-test decision	Reject H_0	Reject H_0
i.i.d. bootstrap	$p = 0.012$ (Reject H_0)	$p = 0.000$ (Reject H_0)
Wild bootstrap (Standard Normal)	$p = 0.013$ (Reject H_0)	$p = 0.000$ (Reject H_0)
Wild bootstrap (Rademacher)	$p = 0.004$ (Reject H_0)	$p = 0.000$ (Reject H_0)
Block bootstrap	$p = 0.032$ (Reject H_0)	$p = 0.003$ (Reject H_0)
Sieve bootstrap	$p = 0.096$ (Reject H_0)	$p = 0.036$ (Reject H_0)

Table 2.1: Comparison of estimated coefficients, t-statistics, and bootstrap p-values for Winter and Summer averages

EXERCISE 3

a. Bootstrap Confidence Intervals with Yearly Average

In addition to hypothesis testing, bootstrap methods are being used for the construction of confidence intervals. Four different confidence intervals of the yearly average temperatures for each bootstrap method are constructed.

(i) Equal-tailed percentile intervals

The bootstrap confidence interval is obtain from inverting a standard t-statistic:

$$C^*(Y) = [\hat{\theta}_n - c^*_{1-\alpha/2}\sqrt{\widehat{\mathrm{Var}(\theta_n)}}, \hat{\theta}_n - c^*_{\alpha/2}\sqrt{\widehat{\mathrm{Var}(\theta_n)}}]$$

where the critical values are obtained from the bootstrap distribution of

$$T^*_n(Y^*) = \frac{\hat{\theta}^*_n - \hat{\theta}_n}{\sqrt{\widehat{\mathrm{Var}(\hat{\theta}^*_n)}}}$$

(ii) Equal-tailed percentile-t intervals

Based an interval directly on the centered quantity $(\hat{\theta}_n - \theta)$ and obtain:

$$C^*(Y) = [\hat{\theta}_n - c^*_{1-\alpha/2}, \hat{\theta}_n - c^*_{\alpha/2}]$$

where c^*_α is the α-quantile of the distribution of

$$(\hat{\theta}^*_n - \hat{\theta}_n)$$

(iii) Symmetric percentile-t intervals

By the assumption of symmetry, a symmetric percentile-t interval can be constructed as:

$$C^*(Y) = [\hat{\theta}_n - c^*_{1-\alpha/2}\sqrt{\widehat{\mathrm{Var}(\hat{\theta}^*_n)}}, \hat{\theta}_n + c^*_{1-\alpha/2}\sqrt{\widehat{\mathrm{Var}(\hat{\theta}^*_n)}}]$$

where the critical values are now based on the bootstrap distribution of the t-ratio

$$|T_n^*(Y^*)| = \left| \frac{\hat{\theta}_n^* - \hat{\theta}_n}{\sqrt{\widehat{\mathrm{Var}}(\hat{\theta}_n^*)}} \right|$$

(iv) Symmetric percentile intervals

By the assumption of symmetry, the absolute value of the centered deviations can be used to obtain the critical values:

$$C^*(Y) = [\hat{\theta}_n - c_{1-\alpha/2}^*, \hat{\theta}_n + c_{1-\alpha/2}^*]$$

where c_α^* denotes the α-quantile of the distribution of

$$|\hat{\theta}_n^* - \hat{\theta}_n|$$

Compare to performing t-test using bootstrap algorithms mentioned in Exercise 1b, the procedures are very similar while the estimations are slightly different in the construction of confidence intervals. For bootstrap algorithms under the null, the restricted version $\hat{\beta}^R = (\hat{\beta}_0, \hat{\beta}_1^R)$ is being used to fit the model and estimate the residuals under the H_0. In constrast, there is no restriction for the coefficients as $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1)$ is being used directly. Consequently, the estimation of model and residuals are completely different from the bootstrap t-tests algorithms. Therefore, the simulations of bootstrap residuals are also different, which result a different empirical distribution of both the bootstrap estimate coefficients and t-statistics. One thing should be noticed and be careful about is the use of standard error. One should use the correct standard error to construct the confidence intervals correspond to the chosen bootstrap algorithm as there are three different types of standard errors ($\sqrt{\widehat{\mathrm{Var}}(\hat{\theta}_n^*)}$, $\sqrt{\widehat{\mathrm{Var}}_{HC}(\hat{\theta}_n^*)}$ and $\sqrt{\widehat{\mathrm{Var}}_{HAC}(\hat{\theta}_n^*)}$).

b. Results and findings

Intervals/Methods	*i.i.d.*	*Wild (Standard Normal)*	*Wild (Rademacher)*
Equal-tailed percentile	[0.01, 0.016]	[0.01, 0.016]	[0.01, 0.016]
Equal-tailed percentile-t	[0.01, 0.016]	[0.01, 0.016]	[0.01, 0.016]
Symmetric percentile	[0.009, 0.017]	[0.009, 0.017]	[0.01 0.017]
Symmetric percentile-t	[0.009, 0.017]	[0.009, 0.017]	[0.009 0.017]

Intervals/Methods	*Block*	*Sieve*
Equal-tailed percentile	[0.008 0.017]	[0.011 0.016]
Equal-tailed percentile-t	[0.008 0.017]	[0.011 0.017]
Symmetric percentile	[0.008 0.018]	[0.01 0.016]
Symmetric percentile-t	[0.007 0.019]	[0.009 0.017]

Table 3.1: The different bootstrap confidence intervals for different types of bootstrap methods

Table 3.1 reveals the bootstraps confidence intervals for corresponding bootstrap methods. All methods yield remarkably similar intervals and centered around 0.013. These results indicate a statistically significant positive temperature trend. The *i.i.d.* and *Wild bootstraps* produce identical results, implying that heteroskedasticity is negligible. The *Block* and *Sieve bootstraps* yield slightly wider intervals which suggest tiny serial correlation in the residuals. Among all the confidence intervals, both the *Symmetric percentile* and *Symmetric percentile-t* are slightly more conservative as the expected nature. Overall, the results confirm the robustness of the estimated warming trend to different bootstrap methods and construction of confidence intervals.

EXERCISE 4

To evaluate the efficiency of bootstrap methods and confidence intervals, Monte Carlo simulation is performed with generated error terms to test the empirical coverages of confidence intervals. The selected specification of error terms is unconditional heteroskedasticity, which is introduced by

$$\varepsilon_t = \sigma_t u_t, \quad u_t \sim NID(0,1), \quad \sigma_t^2 = 1 + 2t + 4t^2$$

This specification makes the variance of the error terms depend on time. Using the estimated coefficients in Exercise 1a to generate the fitted values $\hat{y}_t$, combining the generated heteroskedastic error terms, a new set of data $\{y_t\}$ is simulated. Monte Carlo simulation study is applied with $B = 500$ replications, and a significance level at $\alpha = 10\%$ is used for constructing the confidence intervals. *Nonparametric residual bootstrap (i.i.d. bootstrap)* and *Wild bootstrap* are selected to compare the performances. For two selected bootstrap confidence intervals, *Equal-tailed percentile-t intervals* and *Symmetric percentile intervals* are selected to check the empirical coverages. The detailed bootstrap algorithms and construction of confidence intervals are explained Exercise 1b and Exercise 3.

Methods/Intervals	Equal-tailed percentile-t	Symmetric percentile
i.i.d. Bootstrap	0.874	0.934
Wild Bootstrap	0.902	0.948

Table 4.1: The empirical coverages of the *Equal-tailed percentile-t intervals* and *Symmetric percentile intervals* for *i.i.d. bootstrap* and *Wild bootstrap*

Table 4.1 illustrates the results of empirical coverages of the selected bootstrap confidence intervals with the corresponding bootstrap methods. Since the error terms are generated with unconditional heteroskedasticity, the assumption in *i.i.d. Bootstrap* that the data is homoskedasticity is inappropriate. Consequently, the estimation of standard error and t-statistic are inconsistent. It reflects on the slightly undercoverage than the nominal level in the *Equal-Tailed Percentile-t intervals*. With *Wild Bootstrap*, it successfully preserves the heteroskedastic variance pattern that was present in the residuals. The heteroskedasticity-consistent covariance estimator gives corrected the standard errors and t-statistics. 90.2% empirical coverage of the *Equal-Tailed Percentile-t interval* is close to the nominal level which aligns with the expectation. For *Symmetric Percentile interval*, the estimation does not take into account the estimated standard error and t-statistic, while using the absolute values of centered distribution. This approach makes the intervals to be less sensitive to skewness and tends to be more conservative. Both *i.i.d. Bootstrap* and *Wild Bootstrap* result slighty wider coverages which are 0.934 and 0.948 respectively. These results confirm that expected conservative nature of *Symmetric Percentile interval*. At the same time, *Wild Bootstrap* again shows slightly better calibration than the *i.i.d. Bootstrap*. Overall, these results demonstrate that *Wild Bootstrap* is more robust to heteroskedastic data than *i.i.d. Bootstrap*. Among the interval, *Equal-Tailed Percentile-t interval* provides better performance than the more conservative *Symmetric Percentile interval*.