International Journal of Management, Knowledge and Learning | ISSN 2232-5697

Does Higher Education Dropout in Slovenia Increase? An Empirical Assessment Using Robust and Bayesian Methods

dr. Rado Pezdir
International School for Social and Business Studies, Celje
rado.pezdir@mfdps.si

Purpose: The study aims to assess whether dropout rates in Slovenian higher education have increased over time, focusing on persistence and potential trends between 2011 and 2019. It also seeks to evaluate how robust and Bayesian statistical methods can improve dropout analysis when only limited data are available.

Study design/methodology/approach: The research uses aggregated dropout data from nine consecutive student cohorts (2011–2019) within the Targeted Research Programme (CRP V5-2360). The analysis combines descriptive statistics, robust methods (Theil–Sen estimator with bootstrap confidence intervals), and Bayesian regression modelling with weakly informative priors. This mixed methodological framework is particularly suitable for short time series and allows probabilistic interpretation of results.

Findings: The results reveal that dropout in Slovenia is persistently high, averaging around 51%. Two distinct phases were identified: a decline between 2011 and 2015, followed by renewed growth from 2016 onwards. While classical robust methods yielded inconclusive results regarding a linear trend, Bayesian analysis indicated with high posterior probability (>95%) that dropout has been increasing by about one percentage point per year since 2015.

Originality/value: This study contributes substantively by confirming that Slovenia belongs to the group of European countries with exceptionally high and persistent dropout levels. Methodologically, it demonstrates the added value of applying robust and Bayesian approaches to small-sample educational data, offering more reliable and nuanced insights than conventional frequentist techniques.

Keywords: higher education dropout, dropout persistence, trend analysis, Theil–Sen estimator, Bayesian approach, Slovenian higher education system

Introduction

Dropout in higher education constitutes one of the key challenges of modern education systems. It is a phenomenon with consequences at multiple levels: at the individual level, it represents a missed investment in human capital and often reduced future earnings; at the institutional level, it entails inefficient use of pedagogical and financial resources; and at the societal level, it results in the loss of human potential and reduced economic competitiveness. For these reasons, dropout has become a prominent topic in both research and policy debates.

International research shows that dropout rates across Europe range between 15% and 40% (HEDOCE, 2013; OECD, 2023). The highest values are typical of science and engineering programs, whereas health sciences and education disciplines are less exposed to dropout risks. In this context, Slovenia—with dropout rates averaging around 50%—stands out as one of the countries with the highest rates in Europe, confirming that this is a structural problem of the Slovenian higher education system.

This study draws on data prepared within the Targeted Research Programme (CRP V5-2360), which enables the monitoring of aggregate dropout rates at the first, second, and third cycles of study. The database covers nine consecutive student cohorts (2011–2019). The small number of observations presents a methodological challenge, limiting the use of classical statistical methods. Therefore, we opted for an approach that combines robust and Bayesian methods, which are suitable for the analysis of short time series.

The research addresses three key questions:

(i) What is the average level of dropout in Slovenian higher education in the period 2011–2019?
(ii) Is dropout persistent over time, i.e. does it remain at similarly high levels?

(iii) Are there indications of a positive trend in the latter part of the analysed period?

To answer these questions, we employ the Theil–Sen estimator as a robust measure of trend, bootstrap confidence intervals to estimate uncertainty, and a Bayesian approach that enables probabilistic interpretation of results even when data are limited.

The contribution of this study is twofold. Substantively, we confirm that Slovenia belongs to the group of European countries with persistently high dropout rates, with evidence of an upward trend after 2015. Methodologically, the study illustrates how the use of robust and Bayesian methods can enhance the reliability of analysis in conditions of limited data availability.

2 Literature Review

Research on higher education dropout in the last two decades demonstrates that it is a widespread and systemic phenomenon with significant consequences for individuals, society, and the economy. Dropout rates vary considerably between countries and across fields of study. While some disciplines, especially natural sciences and engineering, record markedly higher dropout rates, others—such as health sciences and education—are generally less exposed.

Beyond structural differences between study programs, dropout is strongly influenced by institutional factors such as student support systems, tutoring, mentoring, and early warning mechanisms for at-risk students. Differences between countries and universities, therefore, reflect not only individual student characteristics but also the interplay of programmatic, institutional, and systemic policies.

The literature review is structured around three key themes:

(i) levels of dropout and international comparisons,
(ii) factors influencing dropout, and
(iii) the consequences of dropout for individuals and society.

This structured approach enables the positioning of Slovenian data within a broader European context and provides the foundation for the empirical analysis in the following sections.

2.1 Dropout Levels in Europe and Cross-Country Comparisons

Empirical studies confirm that dropout in higher education is widespread in Europe, with substantial variation across countries and fields of study. The HEDOCE project (2013), which included 35 European countries, provides a comprehensive comparative overview. Results show that dropout rates range between 15% and 40%, depending on country and discipline. Particularly high risks are found in science and engineering programs, while health sciences and education record substantially lower dropout levels. The project also highlighted that countries with well-developed student support systems (e.g. mentoring, tutoring, early warning mechanisms) typically report lower dropout rates.

Similar findings are reported in national empirical studies. An Italian analysis (Scarpati et al., 2024) finds that dropout is lowest in health sciences (around 5%) but much higher in economics and engineering (up to 40%). This dispersion suggests that dropout is not a uniform national phenomenon but the outcome of institutional, programmatic, and disciplinary factors. Portuguese data (OECD, 2023) also confirms high dropout levels—around 40% on average—placing the country among the European systems with above-average risks.

In this context, Slovenia is particularly problematic, with average dropout rates around 50% during 2011–2019. Although direct international comparisons are methodologically difficult (definitions of dropout vary across countries in terms of program switching, time limits for completion, and inclusion of different forms of study), the levels themselves indicate that Slovenia belongs to the European countries with exceptionally high dropout rates.

2.2 Determinants of Dropout

Studies consistently show that dropout results from the combination of individual, programmatic, institutional, and systemic factors. The most frequently cited clusters include:

First, the field of study plays a crucial role. Analyses (HEDOCE, 2013; Scarpati et al., 2024) demonstrate that science and engineering programs are far more exposed to dropout than health and education programs. Similarly, OECD (2023) reports that dropout risk in STEM fields is above average in most member countries, primarily due to demanding curricula and selective entry conditions.

Second, institutional factors have a significant impact. Countries and universities with well-developed student support systems—tutoring, mentoring, early warning mechanisms—generally report lower dropout rates (HEDOCE, 2013; Tinto, 1993; Yorke & Longden, 2004). Tinto's classical theory of student integration emphasises that academic and social support reduce dropout risk. Yorke and Longden (2004) further demonstrate that institutional measures are most effective when targeted at the early stages of study, when dropout risk is highest.

Third, socio-economic background matters. Students from families with lower socio-economic status face higher dropout risks (Heineck & Kucel, 2015; Thomas, 2002), as they encounter additional financial and social pressures. Similarly, OECD (2021) notes that students who must work while studying are more likely to drop out, as they struggle to balance academic requirements with employment.

Overall, empirical studies confirm that dropout is not merely the result of individual student decisions but primarily the outcome of structural and institutional conditions (Tinto, 1993; Yorke & Longden, 2004; OECD, 2023). Hence, policies to reduce dropout are most effective when they combine measures at the individual level (financial and social support), institutional level (mentoring, tutoring), and systemic level (state scholarship policies, quality assurance regulation).

2.3 Consequences of Dropout

Dropout in higher education has significant consequences at both the micro level (for individuals) and the macro level (for society and the economy).

At the individual level, research consistently shows that dropouts generally face weaker labour market prospects than graduates, although they tend to have better outcomes than those who never entered higher education. Heineck and Kucel (2015), analysing European data, find that dropouts are more likely to be employed than those without higher education, yet their earnings are lower, and job stability is reduced. Martins (2025) similarly reports that dropouts are more likely to hold temporary contracts, experience career interruptions, and have limited promotion opportunities.

At the societal level, dropout is associated with inefficient use of public resources, as higher education in most European countries is heavily state-funded (OECD, 2023). High dropout thus implies that part of the investment fails to achieve its primary objective—completion of studies. In addition, dropout reduces human capital, with long-term consequences for productivity, innovation, and competitiveness (Johnes & Johnes, 2004).

From the perspective of social cohesion, dropout is problematic because it exacerbates inequality. Students from disadvantaged socio-economic backgrounds who drop out are more vulnerable to precarious employment and social exclusion (Thomas, 2002; Yorke & Longden, 2004). Thus, dropout is not only an academic challenge but also a broader societal issue with implications for equal opportunities.

2.4 Summary and the Slovenian context

The literature confirms that dropout in higher education is a complex phenomenon that must be understood as the outcome of interactions among individual student characteristics, program requirements, institutional arrangements, and systemic policies. Across Europe, dropout rates range from 15% to 40%. STEM fields generally record above-average dropout rates, whereas health and education programs are less affected. Institutional measures such as tutoring, mentoring, and early warning systems also significantly influence dropout levels, supporting the notion that dropout can be managed through combined academic and social support policies.

The consequences of dropout manifest at both micro and macro levels: dropouts generally earn less and hold less stable jobs, while societies face inefficient public spending and reduced human capital. Dropout also contributes to social inequality, disproportionately affecting students from disadvantaged socio-economic backgrounds.

Against this background, Slovenia's position is particularly concerning. Average dropout rates during 2011–2019 were around 50%, placing Slovenia among the European countries with the highest values. Although international comparisons are complicated by methodological differences, the level itself indicates a structural challenge for the Slovenian higher education system. This finding provides the starting point for the empirical analysis that follows, which examines the dynamics of dropout in 2011–2019 and its implications for policy.

3 Methodology

The methodological approach of this study is designed as a sequence of analytical steps that provide a comprehensive insight into the dynamics of dropout in Slovenian higher education during 2011–2019. Given the small number of observations ($n = 9$) and the specific nature of the data, special emphasis was placed on methods that are robust to outliers and suitable for short time series.

First, descriptive statistics and graphical representations were used to provide an initial overview of dropout levels, dispersion, and dynamics. This step allowed us to identify the basic features of the data (mean, variability, distributional symmetry) and to visually detect potential trends or breakpoints.

Next, the Theil–Sen estimator was employed as a classical robust method for estimating a linear trend. This nonparametric approach is less sensitive to outliers than ordinary least squares, a key advantage when dealing with small samples.

Since the Theil–Sen method with nine observations does not yield reliable confidence intervals, we complemented it with the bootstrap method. By repeatedly resampling with replacement, we constructed an empirical distribution of slope estimates and derived confidence intervals from it. This approach provides a more reliable assessment of uncertainty, although it remains constrained by the informational content of the original sample.

Finally, a Bayesian approach was applied, enabling the integration of prior knowledge with data in the formation of posterior distributions of parameters. We employed weakly informative priors, striking a balance between stabilising estimates and empirical flexibility. The Bayesian framework is particularly suitable for small samples, as it allows for probabilistic interpretation of results (e.g. the probability that the trend is positive), thereby overcoming limitations of classical methods.

The overall methodological framework thus enables a gradual deepening of analysis: from basic descriptive statistics, through robust trend estimation (Theil–Sen), to the assessment of uncertainty (bootstrap), and finally to Bayesian modelling, which allows a full probabilistic interpretation. This sequence of methods ensures that the results of the empirical analysis are

3.1 Data

The analysis relies on data prepared within the Targeted Research Programme (CRP V5-2360), which enables the monitoring of aggregate dropout rates at the first, second, and third cycles of study. Dropout is defined as the share of students in a given cohort who did not complete their studies within the expected time frame. This definition is methodologically important, as it ensures comparability across cohorts and levels of study.

For this analysis, we used aggregated data for nine consecutive student cohorts (2011–2019). This yields a short time series that allows us to identify trends and the dynamics of dropout across the period. The main limitation of this dataset is the relatively small number of observations (n = 9), which requires the use of methods appropriate for short time series and less sensitive to anomalies. This limitation provides the key methodological justification for applying robust and Bayesian approaches in the remainder of the study.

Table 1: Dropout by year of enrolment

Year	**Dropout (%)**
2011	52,7
2012	52,5
2013	51,4
2014	49,5
2015	47,7
2016	47,9
2017	49,4
2018	53,8
2019	53,8

3.2 Descriptive Statistical Approach

The first step of the analysis is based on descriptive statistics and graphical representations, which provide an initial overview of the dataset. The aim of this step is to capture the level of dropout, its variability, and potential patterns over time. Measures of central tendency (arithmetic mean, median) and dispersion (standard deviation, range) enable a rapid assessment of the extent to which changes in dropout are driven by long-term trends or short-term fluctuations.

In addition to numerical indicators, we employed graphical methods that complement descriptive statistics and facilitate intuitive interpretation. A line chart was used to illustrate dropout dynamics over the entire period and to detect possible changes in its trajectory. A boxplot was employed to examine dispersion, identify the median, and detect outliers, thereby assessing whether the data distribution is consistent with numerical findings.

This approach represents the methodological foundation for subsequent modelling. While descriptive statistics alone cannot confirm hypotheses, they provide a crucial orientation step for selecting appropriate analytical methods in the following chapters.

3.3 Classical Trend Estimation Method: Theil–Sen

To estimate the linear trend in the data, we applied the Theil–Sen estimator (Theil, 1950; Sen, 1968), one of the most widely used robust regression methods. Instead of relying on the mean, the method uses the median of pairwise slopes between observations, making it substantially less sensitive to outliers and anomalies than the classical ordinary least squares (OLS) method. This feature is particularly important in our case, where the number of observations is small (n = 9), and individual deviations could disproportionately affect OLS results.

Formally, the slope β is estimated as:

$$\hat{\beta}_{TS} = median\left(\frac{y_j - y_i}{t_j - t_i}\right) : 1 \leq i < j \leq n$$

where yty_tyt is the dropout rate in year ttt, ttt denotes the time variable, and iii and jjj index the years.

The intercept α is then calculated as the median of all differences between the actual value and the product of the estimated slope and time:

$$\hat{\alpha}_{TS} = median\left(y_i - \hat{\beta}_{TS} t_i\right)$$

Thus, the regression equation is:

$$\hat{y}_t = \hat{\alpha}_{TS} + \hat{\beta}_{TS} t$$

which provides a robust estimate of the dropout trend over time.

The Theil–Sen method is frequently used in the analysis of short time series and environmental data, where samples are small and may include local anomalies (e.g. temperature trends, pollution, demographic indicators). In our case, it is suitable because it enables the estimation of the underlying linear dropout trend without strong sensitivity to breakpoints (such as the shift from declining dropout up to 2015 to rising dropout after 2015).

The limitation of this method is that, despite its robustness, it cannot provide confidence intervals from classical theory when the sample is very small. For this reason, we extended the analysis with a bootstrap approach in the next step, which allows an empirical assessment of the uncertainty surrounding the estimated slope.

Since the Theil–Sen method, with only nine observations, does not allow for reliable confidence interval estimation under classical asymptotic theory, we employed the bootstrap method (Efron, 1979). This nonparametric technique of repeated resampling with replacement from the original dataset enables an empirical estimation of the distribution of the estimator.

The basic idea is to draw repeated bootstrap samples$D^{*(b)}$ of size n from the original dataset $D = \{(t_i, y_i)\}_{i=1}^{n}$, with replacement, and calculate the Theil–Sen slope estimator for each sample. Formally, for *b* = 1, … ,*B* (*B* = 5000) in our case:

$$\hat{\beta}_{TS} = median\left(\frac{y_j^{*(b)} - y_i^{*(b)}}{t_j^{*(b)} - t_i^{*(b)}}\right) : 1 \leq i < j \leq n$$

This produces a set of bootstrap estimates $\left\{\hat{\beta}_{TS}^{*(1)}, \hat{\beta}_{TS}^{*(2)}, \ldots, \hat{\beta}_{TS}^{*(B)}\right\}$ which represents the empirical distribution of slope estimates. From this distribution, confidence intervals can be constructed. The most common is the percentile interval:

$$CI_{1-\alpha} = \left[\hat{\beta}^{*(\alpha/2)}, \hat{\beta}^{*(1-\alpha/2)}\right]$$

where $\hat{\beta}^{*p}$ denotes the p-th percentile of the bootstrap distribution.

In our analysis, 5000 bootstrap replications were performed, ensuring a stable empirical distribution of the slope estimates. The resulting confidence interval was wide and included zero, suggesting that the robust classical framework does not allow us to conclude the existence of a clear trend. A histogram of bootstrap slopes further confirmed the symmetry of the distribution around zero, with no substantial bias.

The main advantage of the bootstrap method is that it enables uncertainty estimation even when sample sizes are small and theoretical distributions of the statistic are not available. Its limitation, however, is that it is based solely on resampling from the existing data and thus cannot extend beyond the informational content of the original sample. For this reason, the next step of the analysis employed a Bayesian approach, which incorporates prior knowledge and yields posterior distributions of the parameters.

3.4 Bayesian Approach

As an extension of the robust classical methods, we employed a Bayesian regression model, which enables inference about parameters in the form of posterior distributions. Unlike the frequentist approach, which relies on point estimates and confidence intervals, Bayesian analysis combines information from the data with prior distributions, which is particularly relevant when sample sizes are small, as in our case (n = 9).

The basic model assumes a linear relationship between dropout rate and time:

$y_t = \alpha + \beta t + \epsilon_t,\ \varepsilon_t \sim (0, \sigma^2)$

where y_t is the dropout rate in year t, α is the intercept, β is the slope measuring annual change in dropout, and ε_t is the error term.

In the Bayesian framework, priors must be specified for the parameters. In our case, we used weakly informative priors to stabilise estimates while avoiding overly restrictive assumptions:

$\alpha \sim (50{,}10), \quad \beta \sim (0{,}5), \quad \sigma \sim Half - Cauchy(0{,}5)$

The intercept prior N(50,10) reflects the fact that dropout hovered around 50% in the period, with a standard deviation of 10, allowing a sufficiently wide range (approximately 30% to 70%). For the slope β, we used N(0,5), implying no strong prior expectation of a trend (positive or negative), but permitting a broad range of possibilities. For the error standard deviation σ\sigmaσ, we used a weakly informative Half-Cauchy prior, which allows for realistic variance estimation and avoids numerical instability in small samples.

Posterior distributions are obtained via Bayes' rule:

$$p(\alpha, \beta, \sigma \mid \mathrm{y}) \propto p(\mathrm{y} \mid \alpha, \beta, \sigma) \cdot p(\alpha) \cdot p(\beta) \cdot p(\sigma)$$

where $p(\mathrm{y} \mid \alpha, \beta, \sigma)$ is the likelihood function and $p(\alpha) \cdot p(\beta) \cdot p(\sigma)$ represents the prior probabilities. The result is a posterior distribution that combines prior knowledge with empirical data.

We estimated the posterior distributions using Markov Chain Monte Carlo (MCMC) sampling via the *rstanarm* package. This produced a large number of draws from the posterior distribution, enabling us to compute posterior means, medians, standard errors, and credible intervals for the parameters.

The advantage of the Bayesian approach lies in its probabilistic interpretation of results. For instance, instead of stating that the slope is "not statistically different from zero," the posterior distribution allows us to say that there is a greater than 95% probability that the dropout trend during the analysed period is positive. This interpretation is more intuitive and informative, particularly in small-sample settings where frequentist methods often remain inconclusive.

4 Results

4.1 Descriptive Statistics and Basic Visualisations

The first step of the empirical analysis focused on descriptive statistics, providing a basic overview of the level and dispersion of dropout rates as well as the identification of potential peculiarities in the data. The average dropout rate during 2011–2019 was 50.97%, meaning that, on average, approximately half of students withdrew or failed to complete their studies. The median was 51.4%, very close to the arithmetic mean, suggesting that the data are relatively symmetrically distributed without pronounced skewness. The lowest value was recorded in 2015 (47.7%), while the highest values were observed in 2018 and 2019 (53.8%). The standard deviation of 2.4 percentage points indicates that dropout fluctuated within a relatively narrow range during the period and was not subject to extreme annual variations.

Graphical representations provide additional insight into these findings. The boxplot confirms that the values are clustered around the median, with no pronounced outliers, indicating that dropout rates varied within expected bounds across years. The line graph, however, clearly shows that the data dynamics do not follow a strictly linear pattern. Between 2011 and 2015, dropout gradually decreased (from 52.7% to 47.7%), reflecting an improvement in student progression. In the second half of the period, from 2016 to 2019, the trend reversed and dropout increased again, reaching its highest values in the observed interval.

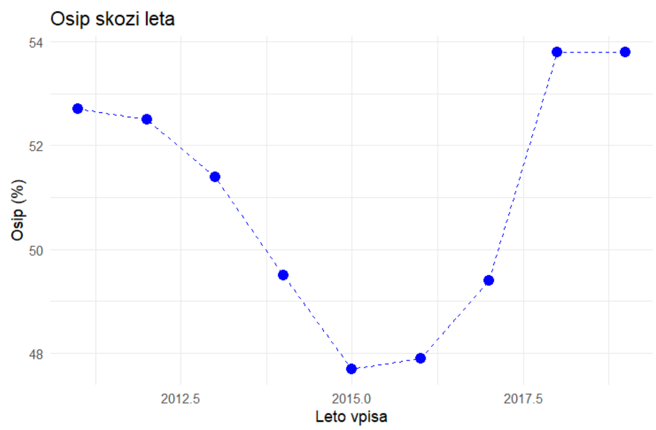

Figure 1: Dropout in the Slovenian Higher Education System (2011–2019)

Figure 1 clearly illustrates the two-phase movement of dropout: a period of decline up to 2015, followed by renewed growth after 2015. Such dynamics explain why average linear estimates struggle to capture the actual course of change.

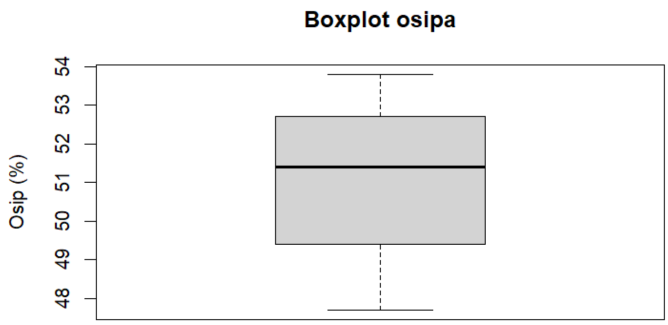

Figure 2: Boxplot of dropout rates in the period 2011–2019

The boxplot (Figure 2) confirms that values are symmetrically distributed around the median, with no pronounced outliers. This indicates that the dynamics of dropout are primarily determined by underlying trends rather than by exceptional values in individual years. Based on these descriptive results, the entire period can be divided into two phases: (1) 2011–2015, when dropout declined, and (2) 2016–2019, when dropout rose again. This suggests that a simple linear approach to trend estimation may be insufficient, as the average slope masks the dynamics of two opposing movements. For this reason, in the subsequent analysis, we applied robust trend estimation methods, which are less sensitive to local changes and allow for more stable conclusions even with a small number of observations.

4.2 Theil–Sen Estimate of the Robust Trend

For the initial assessment of the trend in the data, we applied the Theil–Sen estimator, one of the most widely recognised robust methods of linear regression (Theil, 1950; Sen, 1968). In contrast to the classical ordinary least squares (OLS) method, which is sensitive to outliers and assumes normally distributed errors, the Theil–Sen method is nonparametric and therefore suitable for analysing time series with a limited number of observations and potential anomalies. The method estimates the slope of the regression line as the median of all possible pairwise slopes between observations, which provides robustness in short and noisy data series. For this reason, it is frequently employed in environmental and social science research, where time series are short and deviations are common.

In our case, the Theil–Sen slope estimate was –0.0167. The negative sign could suggest a slight decline in dropout over the period, but the change is extremely small and statistically insignificant. The test statistic showed no deviation from zero (p-value = 1.0), and due to the small sample (nine observations), it was not possible to calculate a reliable confidence interval. This means that the method does not confirm the existence of a clear linear trend in the period 2011–2019.

The result can also be understood in connection with the descriptive statistics and graphical representations presented in Section 4.1. It was evident that dropout decreased between 2011 and 2015, but increased again in the period 2016–2019. The Theil–Sen method, which seeks an average linear trend across the entire period, does not capture these two phases separately. Consequently, the slope estimate remains close to zero, giving the impression of no trend, even though visual analysis suggests a breakpoint dynamic.

As an additional robust measure of central tendency, we used the Hodges–Lehmann estimator, which amounted to 50.95%. This result confirms that dropout values were mostly centred around 51% and that the linear trend estimate does not deviate significantly from this central value.

The limitation of the Theil–Sen estimator in our case is primarily linked to the small sample size. With only nine observations, the slope estimate is sensitive to minor changes in individual values, so even slightly different dynamics in a single year may significantly affect the result. Because of these limitations, in the next step, we employed the bootstrap approach, which allows for the simulation of the slope distribution and the derivation of confidence intervals even in small samples where classical methods fail.

4.3 Bootstrap Confidence Intervals for the Theil–Sen Estimator

Since the Theil–Sen slope estimate in our case was not statistically significant and the small number of observations precluded reliable classical confidence interval estimation, we employed the bootstrap method. The bootstrap approach is based on repeated resampling with replacement from the original dataset and enables empirical estimation of the distribution of the estimator without relying on theoretical asymptotic properties. This is particularly useful for short time series such as the one considered here (n = 9).

In our case, 5000 bootstrap replications were performed. The results show that the 95% confidence interval for the Theil–Sen slope was (–0.59 ; +0.61). Since the interval includes zero, we conclude that within the classical robust framework, there is no statistically significant trend. This implies that the slope could be either slightly negative or slightly positive, leading to a methodologically inconclusive result.

The histogram of bootstrap slopes confirms this finding: the distribution is symmetric and centred around zero, without marked skewness in either direction. This visualisation clearly illustrates that classical robust approaches do not provide sufficient information in our case to reach a reliable conclusion.

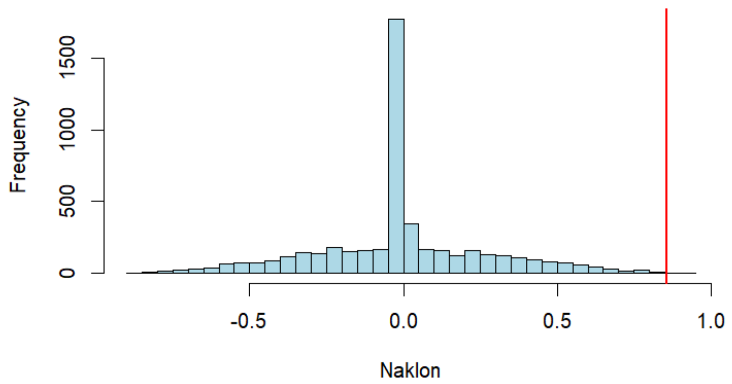

Figure 3: Bootstrap distribution of Theil–Sen slopes

Relating this to the previous section (4.2), it can be said that the Theil–Sen slope estimate remains close to zero because of the breakpoint dynamics (decline until 2015, followed by an increase until 2019). The bootstrap confirms this with a wide uncertainty interval. This means that classical robust results are insufficient for a clear interpretation of the trend.

This limitation justifies the methodological transition to a Bayesian approach, which, instead of a point estimate and a wide confidence interval, provides the full posterior distribution of the slope. Bayesian analysis not only delivers point estimates and intervals but also allows probabilistic interpretation (e.g. the probability that the trend is positive). In this way, the Bayesian framework goes beyond classical robust methods and offers more informative conclusions in small-sample settings.

4.4 Bayesian Analysis of the Linear Trend

After both the Theil–Sen estimator and its bootstrap framework indicated that no clear trend could be confirmed within the classical robust approach, we proceeded with a Bayesian regression model. The Bayesian framework differs fundamentally from frequentist methods in that it produces posterior parameter distributions, combining information from the data with prior assumptions. This approach is particularly appropriate when the number of observations is small, as in our case (n = 9), since it does not rely on asymptotic properties and enables more informative inference about parameters.

The *stan_glm* model was used with weakly informative priors. For the slope, we applied a prior N(0,5), meaning that we had no strong prior expectation about the trend direction, while allowing for a broad range of possible values. For the intercept, a prior N(50,10) was used, since dropout in previous years hovered around 50%, while still allowing flexibility for the model to adapt to the data. For the standard deviation of errors, a weakly informative Half-Cauchy prior was chosen to prevent numerical instability in small samples. This ensured a balance between prior knowledge and empirical evidence.

The results indicate a posterior slope estimate of 0.9, with a standard error of 0.1. The most important finding is that the 95% posterior credible interval is [0.71 ; 1.03]. Since the interval does not include zero, we can conclude with high posterior probability (greater than 95%) that a positive dropout trend exists during 2011–2019. This means that dropout increased on average by approximately one percentage point per year.

The posterior distribution of the slope is clearly concentrated on the positive side, confirming that the increase in dropout after 2015 is the dominant signal in the data. Diagnostic indicators (Rhat = 1, n_eff > 2000) confirm good convergence and stability of the MCMC sampling, and the posterior predictive mean (mean_PPD = 51.0) closely matches the empirical average dropout rate.

A comparison between the bootstrap distribution of Theil–Sen slopes and the Bayesian posterior distribution highlights the key contrast between the two approaches. While the bootstrap produces a wide, symmetric interval that includes zero, the Bayesian analysis yields a narrow, positive interval that excludes zero. This enables probabilistic interpretation: there is a very high probability that the dropout trend is positive.

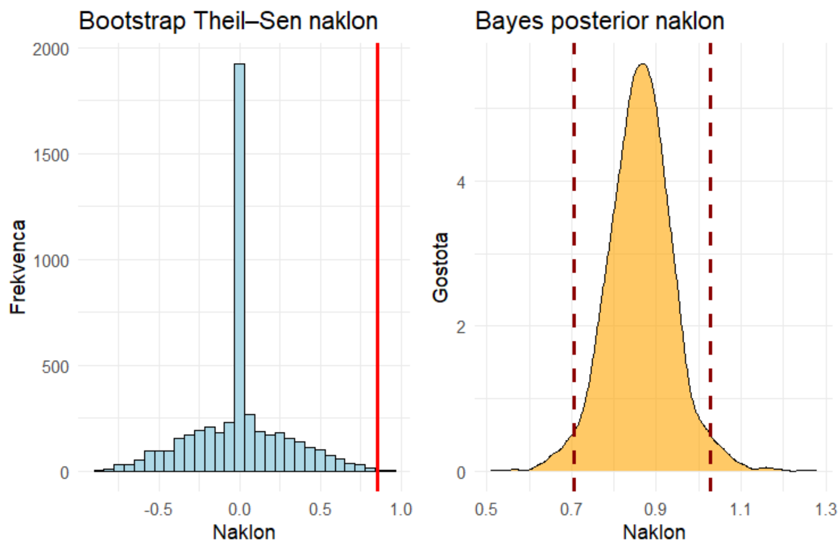

Figure 4: Comparison of bootstrap and Bayesian trend estimates

Despite the advantages of the Bayesian approach, it should be noted that results are to some extent sensitive to the choice of priors. However, in this analysis, we employed weakly informative, empirically grounded priors that do not bias the results but merely stabilise the estimates. Thus, the Bayesian framework surpasses the limitations of classical robust methods and provides more informative and intuitive conclusions when observations are scarce.

4.5 Overview

The analysis of dropout in the Slovenian higher education system during 2011–2019 reveals a multifaceted picture that varies depending on the methodological tool employed. Descriptive statistics and graphical representations indicated that dropout averaged around 51%, with relatively low variability (SD = 2.4 percentage points). The data clearly suggest two phases: a decline between 2011 and 2015, followed by an increase between 2016 and 2019, reaching the highest values in the series. These visual findings point to the possibility of nonlinear dynamics that simple linear approaches struggle to capture.

The Theil–Sen estimator, applied as a robust nonparametric method, produced a slope of –0.0167, implying a negligible decline in dropout over the period. Statistical significance was not confirmed ($p = 1.0$), and the method did not allow for confidence intervals due to the small number of observations. This outcome is consistent with expectations from a robust method, as Theil–Sen averages the entire period and cannot account for the two distinct phases (decline until 2015, increase after 2015). The Hodges–Lehmann estimate of central tendency (50.95%) further confirmed that dropout values were stably centred around the median throughout the period.

The bootstrap extension of the Theil–Sen method added an important dimension. The empirical distribution of slopes from 5000 replications showed a symmetric distribution around zero, with the 95% confidence interval (–0.59 ; +0.61) including zero. This confirmed the Theil–Sen outcome: classical robust methods do not detect a statistically significant trend. Within the frequentist framework, we thus remain with an inconclusive result—dropout fluctuated, but no linear trend can be established.

By contrast, the Bayesian regression model produced clearer results. Using weakly informative priors N(0, 5) for slope, N(50, 10) for intercept, estimates were stabilised while the data remained the primary source of information. The posterior slope estimate (0.9) and the 95% credible interval [0.71 ; 1.03], which excludes zero, clearly indicate with high posterior probability that a positive trend exists. The Bayesian framework also enables probabilistic interpretation: there is more than a 95% probability that dropout increased during 2011–2019, at a rate of about one percentage point per year.

The comparison between the bootstrap and Bayesian approaches provides a key methodological insight. While the bootstrap, as a typical representative of the frequentist framework, produces a wide interval and inconclusive results, the Bayesian model—by combining data with prior knowledge—offers more informative conclusions that remain useful even with very small samples. For the analysis of higher education dropout, this means that Bayesian modelling can uncover trends that classical methods fail to detect due to methodological limitations.

Overall, the interpretation is that dropout between 2011 and 2019 followed a two-phase pattern: after an initial decline, it began to rise again after 2015. The Bayesian model confirms that the long-term trend over this period is positive, which has important implications for higher education policy.

5. Discussion

5.1 Comparison with Literature

The empirical results show that the average dropout in Slovenia during 2011–2019 was around 50%, with a two-phase dynamic: a decline in the first part of the period (2011–2015), followed by renewed growth after 2015. Robust classical methods (Theil–Sen estimator and bootstrap) did not confirm a statistically significant linear trend, but the Bayesian approach suggests that the probability of a positive trend after 2015 is high.

This picture is consistent with international research. The HEDOCE project (2013) reported European dropout rates between 15% and 40%, with science and engineering programs typically showing higher values, while health and education recorded lower rates. In this context, Slovenia—with dropout rates around 50%—stands out as a country with above-average risk, a finding also supported by OECD (2023), which reports Portuguese dropout rates averaging 40% but notes that Slovenian levels are even higher.

A comparison with Italian data (Scarpati et al., 2024) also points to similar differences across fields of study, confirming that dropout is a heterogeneous phenomenon. While health sciences register dropout rates as low as 5%, economics and engineering rise to 40%. This suggests that Slovenia's high dropout level cannot be explained solely by national idiosyncrasies, but likely reflects a combination of programmatic and institutional factors, consistent with Tinto's theory of student integration (1993) and subsequent research (Yorke & Longden, 2004).

Our results thus extend existing literature in two ways. First, they confirm that Slovenia belongs to the group of countries with exceptionally high dropout risk, as international comparisons also indicate. Second, they demonstrate that dynamics within the observed period are changing: after an initial decline, dropout began rising again after 2015, pointing to shifts in institutional or structural factors that should be explored in further research.

5.2 Specifics of the Slovenian Case

The key specificity of Slovenia is that dropout in 2011–2019 appears as a highly persistent phenomenon. Despite some year-to-year fluctuations and a temporary decline between 2011 and 2015, dropout rates remained anchored around 50% in the long run, with no signs of sustained improvement. This stability at high levels indicates a structural problem of the higher education system that cannot be explained merely by temporary circumstances or individual cohorts.

The Bayesian approach estimated a positive trend of about +0.9 percentage points per year. From a policy perspective, this is concerning, as it suggests that dropout began rising again in the second half of the period. Although the result must be interpreted cautiously due to the small sample size, it nevertheless provides a clear signal that existing dropout reduction policies have not proven effective.

The persistence of dropout at around 50% and the finding that the trend is not negative but positive place Slovenia in a distinctly worse position compared to the European average. This raises the question of what policies were implemented during this period to reduce dropout and why they did not yield the expected effects. It also highlights the need for systematic monitoring of the effectiveness of existing support mechanisms (tutoring, mentoring, scholarships) and consideration of new interventions that could reduce the risk of premature withdrawal from study.

5.3 Policy Recommendations

Our results, showing high dropout persistence (around 50% throughout the period) and a positive trend (+0.9 percentage points per year after 2015), have direct implications for policy in Slovenian higher education. While dropout is a widespread phenomenon across Europe, such high and persistently stable levels indicate that existing measures in Slovenia have not been sufficient to effectively reduce risk.

Policy recommendations can be structured along three lines:

(i) Strengthening institutional student support mechanisms. International research (HEDOCE, 2013; Yorke & Longden, 2004) consistently confirms the importance of tutoring, mentoring, and early warning systems in reducing dropout. In Slovenia, such measures have been introduced (e.g. tutoring and mentoring programs at universities), but their effectiveness has not been systematically evaluated. A thorough assessment of existing practices is needed, along with targeted improvements to support vulnerable groups.

(ii) Focusing on fields with high dropout rates. Our results confirm that dropout is a structured phenomenon, more pronounced in certain disciplines. Italian research (Scarpati et al., 2024) reaches similar conclusions. Policymakers should therefore develop field-specific measures—for example, preparatory training programs for students in technical and natural science fields, where risk is highest.

(iii) Addressing socio-economic factors. Literature (Heineck & Kucel, 2015; Thomas, 2002) shows that students from lower socio-economic backgrounds face greater dropout risk. In Slovenia, forms of support such as state scholarships and subsidised accommodation have been available, but questions remain as to whether they were sufficiently accessible and effective. Consideration should be given to targeted scholarship schemes for vulnerable student groups and greater flexibility in study programs to facilitate the combination of study and work.

All of this suggests that without additional measures, the high persistence of dropout in Slovenia is likely to continue. Based on our findings, it would be advisable for policymakers to develop a national dropout reduction strategy that brings together universities, the ministry, and other stakeholders, while ensuring regular monitoring of the effectiveness of adopted measures.

5.4 Limitations and Reliability of Results

When interpreting the findings, it is necessary to acknowledge the limitations of both the data and the methodological approach. The analysis is based on nine consecutive student cohorts, which constitutes a relatively small sample. Such a limitation affects the reliability of the results, as classical regression methods are not stable with small datasets, and their assumptions of normally distributed errors and homoskedasticity are unlikely to hold. For this reason, instead of standard linear regression, we applied the Theil–Sen estimator, which is more robust and reduces the influence of outliers.

Nevertheless, even the Theil–Sen method does not allow for classical confidence interval derivation with only nine observations. We addressed this by using the bootstrap method, which enabled an empirical assessment of uncertainty. However, bootstrap cannot overcome the limitations of the underlying sample, as it always relies on resampling the observed data. Consequently, the resulting confidence intervals remain wide and include the possibility of a zero trend.

The Bayesian approach provided more stable estimates and allowed for probabilistic interpretation of the slope. Yet, posterior distributions are still fundamentally derived from the same limited informational base. The results should therefore be understood as indicative and cautionary, rather than as definitive evidence of the existence or direction of a trend.

Despite these limitations, the combination of robust and Bayesian methods offers a relatively reliable insight into the dynamics of dropout, as the approaches employed are particularly suitable in contexts where the number of observations is small and anomalies are possible. The results do not provide final confirmation but rather highlight that high dropout rates in Slovenia are persistently present, and that the detected positive trend is a serious indication warranting further investigation. Future research should extend the analysis to a longer period or employ more disaggregated data, for example, by field of study, level, or institution. Such refinements would allow for more precise identification of underlying causes and facilitate more effective policy design.

References

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics, 7*(1), 1–26. https://doi.org/10.1214/aos/1176344552

Heineck, G., & Kucel, A. (2015). Fields of study and graduates' overqualification: A comparative analysis of EU countries. *IZA Discussion Paper Series, No. 9148.* Bonn: IZA Institute of Labour Economics.

HEDOCE. (2013). *Higher Education Drop-out and Completion in Europe.* Final report. European Commission, Directorate-General for Education and Culture.

Johnes, G., & Johnes, J. (2004). *International Handbook on the Economics of Education.* Cheltenham: Edward Elgar.

Martins, P. (2025). The labor market consequences of dropping out of university in Europe. *European Economic Review, 163,* 104215.

OECD. (2021). *Education Policy Outlook: Shaping Responsive and Resilient Education in a Changing World.* Paris: OECD Publishing.

OECD. (2023). *Education at a Glance 2023: OECD Indicators.* Paris: OECD Publishing. https://doi.org/10.1787/69096873-en

Scarpati, C., Bison, I., & Mezzanzanica, M. (2024). Understanding university dropout: Evidence from Italy. *Higher Education, 88*(2), 315–339. https://doi.org/10.1007/s10734-023-01073-5

Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association, 63*(324), 1379–1389. https://doi.org/10.1080/01621459.1968.10480934

Theil, H. (1950). A rank-invariant method of linear and polynomial regression analysis. *Nederlands Akonomisch Instituut.*

Thomas, L. (2002). Student retention in higher education: The role of institutional habitus. *Journal of Education Policy, 17*(4), 423–442. https://doi.org/10.1080/02680930210140257

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed.). Chicago: University of Chicago Press.

Yorke, M., & Longden, B. (2004). *Retention and student success in higher education.* Maidenhead: Open University Press.