Examining How Region and Individual Competency Affect Team Performance in Artistic Gymnastics Using RStudio

By Yuhan (Jennifer) Xiao

Author Bio

Yuhan Xiao is currently a senior student at Saratoga High School. Through participating in math competitions, she became interested in problem-solving and applications of mathematics. She started to get involved in research in 2021, with her first project centered around the applications of Bayes Theorem in Olympic Gymnastics participation. After learning about statistical measures and RStudio computing and modeling further, she now looks forward to conducting research with a combination of mathematical theory, modeling, and data analysis. Beginning with investigating gymnastics and team sports, she aims to employ quantitative methods to offer insights on and help make decisions under social and economic contexts in the future.

Abstract

Artistic gymnastics has been widely known for the amount of uncertainty involved in competitions, such as falls on the balance beam. However, there might be factors consistently contributing to team competition results. Previous research on gymnastics primarily focused on individual physical fitness, sports medicine, or educational aspects and were mostly qualitative investigations. This study collects NCAA women's artistic gymnastics data from the Road To National database and employs RStudio modeling to analyze how regional background and individual strength play a role in college gymnastics team competitions. Both pre-pandemic and post-pandemic seasons are taken into consideration. Contrary to common assumptions, our results suggest that average team score varies from region to region, and the average of top three scores from each team have a significant impact on its performance, while variance between top scoring individuals cannot predict team performance. These differences can be possibly attributed to local financial situations and Olympic-level gymnasts' participation, offering potential directions for further research on artistic gymnastics and team sports.

Keywords: Artistic Gymnastics, Team Sports, RStudio, Regional Difference, Individual Strength, ANOVA, Linear Regression, NCAA

Introduction

Gymnastics, one of the most watched sports in America, has been known for the uncertainty of athletes' performances and scoring. However, this paper asserts that gymnastics is no different from other competitive sports and predictive models based on fixed independent variables can predict competition outcomes.

For instance, though sports has become more and more accessible to the global population, studies have shown that geographic background and socioeconomic status still make an impact on the performance of a country, narrowing down to athletes, in major competitions (Wang et al.). Despite the strong presence of millions of gymnastics participants in the United States, many gymnasts who come from less privileged schools and regions either cannot receive top-tier coaching or do not have sufficient space and equipment for practice, negatively influencing their gymnastics performance. Therefore, the first goal of this project is to analyze whether regional background plays a role in team performance, using region as a proxy for resources. The second goal of this paper is to examine the specific format of team competitions to determine how extremely competitive or consistent individuals impact a team's performance.

Background & Research Questions

The term Artistic Gymnastics (AG) emerged as early as the 1800s and grew in Europe through the 1880s. The term was used to describe events at competitions and clubs. As the sport evolved during the last century, modern AG now includes athletes performing their own routines on one or more apparatuses, and AG competitions have a variety of individual and team events. Since different branches of AG follow different systems of scoring, in order to eliminate biases and discrepancies in scoring, this project will solely focus on NCAA Gymnastics, more specifically women's AG team performances. There are four apparatuses in women's AG: Vault, Uneven Bars, Balance Beam, Floor Exercise. In a competition, six athletes from each team perform two routines on each apparatus and the lowest score is dropped. The starting value for every gymnast is 9.5; with difficulty bonuses, the maximum potential score is 10.0 (NCAA.org).

The research questions this paper seeks to answer are:

1. How do performance of teams from different regions vary in women's artistic gymnastics competitions? Previous studies suggested wealth as an influential factor in gymnastics training, but they concentrated on relatively small samples of athletes. This study will hopefully fill a gap in the research by investigating regional gymnastics performance differences across the USA.

2. How do the average total score of the top three athletes on each team affect team performance in women's AG? With previous research on gymnastics mostly focusing on athletic skill, sports medicine, or education, it is necessary to examine the relationship between individual strength and team performance using a quantitative approach.

Data

All data is from the Road To Nationals database, created by two NCAA and club gymnastics coaches. In the database, the six regions in the NCAA are designated as: Central (C), North Central (NC), Northeast (NE), South Central (SC), Southeast (SE), and West (W). The specific teams included in a region varied from year to year. Table 1 below displays the distribution of NCAA gymnastics teams by regions for the three athletic seasons of 2019, 2020, and 2022. Due to the pandemic, only about 80% of the teams competed in 2021, so the 2021 season was not analyzed in the project.

Table 1: Distribution of NCAA Gymnastics Team by Region

Year \ Region	C	NC	NE	SC	SE	W
2019	13	17	19	13	9	11
2020	13	17	19	12	9	11
2022	13	17	20	12	9	10

Methods

This project consists of two sections; one investigates regional differences in gymnastics team performance, while the other studies the impact of the

top three scoring individuals on each team. In order to include a broader population and eliminate differences between the elite and college scoring systems, only NCAA gymnastics was taken into consideration. Further, team performance was evaluated using the average team score from the targeting season(s) to minimize discrepancies caused by judging bias and athlete conditions.

For the first part of the project, we studied the relationship between the categorical independent variable, regions, and dependent variable, team average score, employing a one-way analysis of variance (ANOVA). Each region's mean team score was compared to every other region's mean team score using the TukeyHSD function in RStudio. We also constructed several plots to compare the range and 95% confidence interval for each region's team average scores. The modeling process was repeated for the 2019, 2020, 2022 seasons.

For the second part of the project, data from the NCAA database was used to model the relationship between the top three scoring individuals from each team and the corresponding team's average score. Multiple linear regression was used in order to take into consideration the variance between the top three scores and number of attendance. ggplots were created to visualize the strength of the linear regression model based on each of the two independent variables. The model only focused on the 2022 season since top individuals' performance fluctuated little from year to year.

Results

In order to study the performance of gymnastics teams across different regions, we first performed one-way ANOVA to compare the effect of regions on team score for the 2019, 2020, 2022 seasons. Respectively, 82, 81, and 81 teams participated in NCAA gymnastics in 2019, 2020, and 2022. The one-way ANOVA model in RStudio revealed that there was a statistically significant difference between mean team scores between at least two NCAA regions at a 0.01 significance level for all three seasons. Table 2 includes the specific F-value and p-value from all three years.

Table 2: One-way ANOVA Statistics

Year	F-value	p-value
2019	3.944	0.00311
2020	4.607	0.00102
2022	4.459	0.00131

We further conducted Tukey's HSD Test to investigate every single pair of regions. In 2019, Tukey's HSD Test reported significant differences in mean team scores between North Central-Central (p=0.0337), South Central-North Central (p=0.0388), and West-North Central (p=0.0460) at a 0.05 significance level. In 2020, Tukey's HSD Test revealed significant differences in mean team scores between North Central-Central (p=0.0361), South Central-North Central (p=0.0308), and Southeast-North Central (p=0.0432) at a 0.05 significance level. In 2022, Tukey's HSD Test revealed significant differences in mean team scores between Northeast-Central (p=0.0498) and South Central-Northeast (p=0.0472) at a 0.05 significance level. Table 3 below collects the p-value from all three years' Tukey's HSD Test.

Table 3: Tukey's HSD Test p-value by Year

Regions Compared	2019	2020	2022
NC-C	0.0336535	0.0361226	0.0630993
NE-C	0.2312146	0.0915545	0.0498278
SC-C	1.0000000	0.9999962	0.9999994
SE-C	0.9999998	0.9998784	1.0000000
W-C	1.0000000	0.9999995	1.0000000
NE-NC	0.9246678	0.9968817	1.0000000
SC-NC	0.0388036	0.0307666	0.0592711
SE-NC	0.0649913	0.0432435	0.1257375
W-NC	0.0459586	0.0655222	0.1050475
SC-NE	0.2555576	0.0779405	0.0471544
SE-NE	0.3143723	0.0993613	0.1086421
W-NE	0.2683729	0.1497310	0.0887639
SE-SC	0.9999971	0.9999947	0.9999998
W-SC	0.9999996	0.9999601	0.9999996
W-SE	1.0000000	0.9996174	1.0000000

In order to display the results, box plots of team average scores by region were graphed for the three seasons. For instance, in Fig 1, the x-axis is labeled as "Region", the y-axis is labeled by "Average Score", and each of the six regions is color coded. The bolded, center line within each box represents the region's mean team average score in 2019. The upper and lower boundaries of the box mark the 75th and 25th percentile team score of the region. Since the boxes' mean values and interquartile ranges differ, it is confirmed visually that team performances differ by region. Fig 2 and Fig 3 confirm similar results for the 2020 and 2022 season.

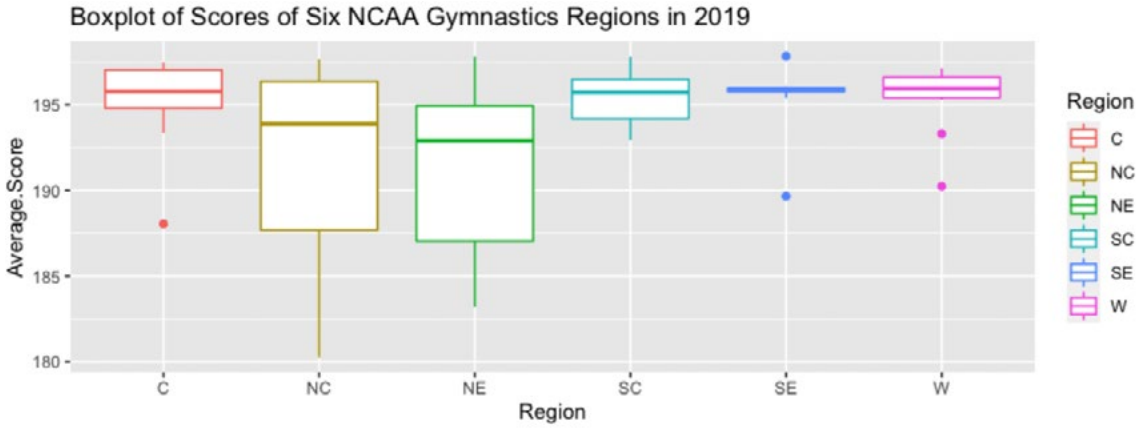

Fig 1. Boxplot of Team Average Scores by Region in 2019

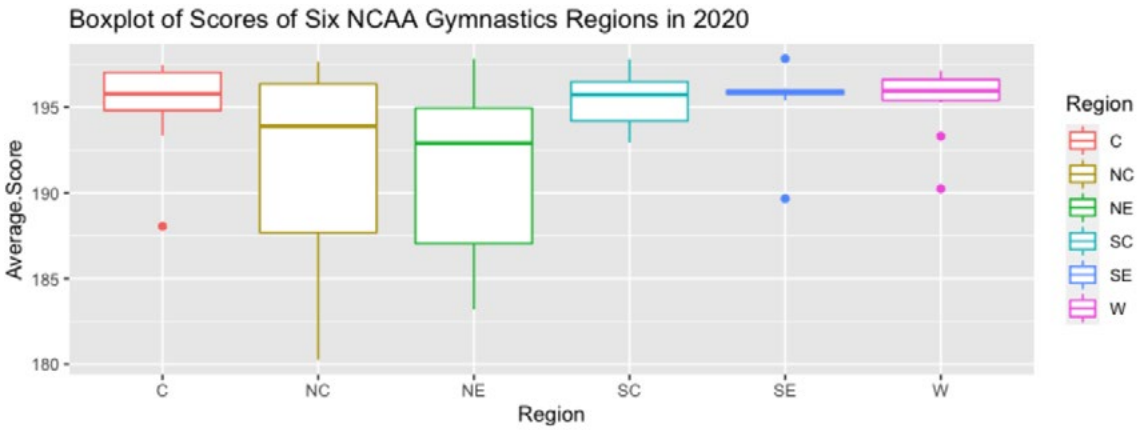

Fig 2. Box Plot of Team Average Scores by Region in 2020

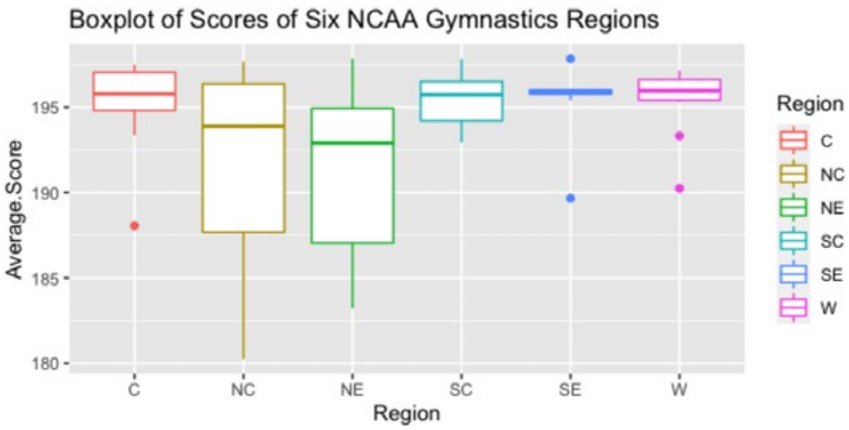

Fig 3. Box Plot of Team Average Scores by Region in 2022

Fig 4, 5, and 6 shown below are 95% family-wise confidence levels of average team scores by region constructed for the 2019, 2020, and 2022 seasons. Each pair of regions being compared is labeled on the y-axis, while x-axis is labeled with values from negative through positive. Each line segment represents the 95% confidence interval of the difference between two regions' mean team score. If the value zero is not within the range of the 95% confidence interval, there is a significant difference between the two regions compared. There are respectively three, three, and two intervals that do not cover zero in Fig 4, Fig 5, and Fig 6, which are consistent with Tukey's HSD Test results.

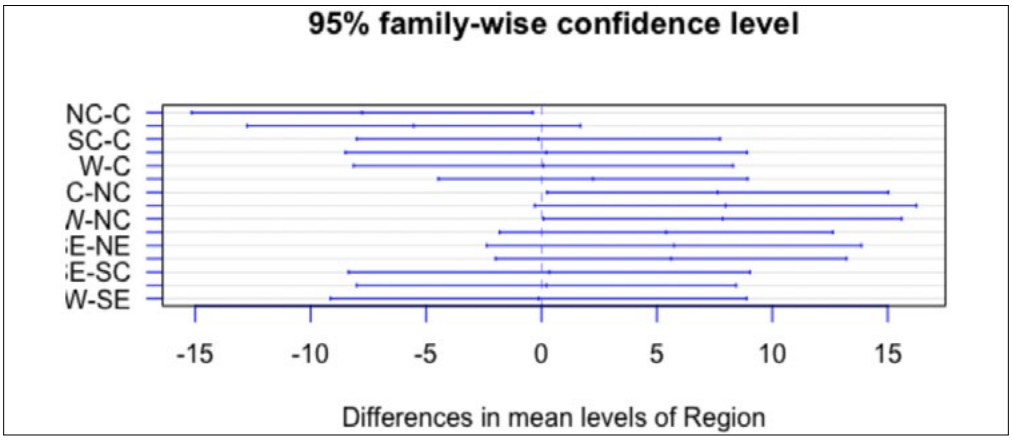

Fig 4. Plot of 95% Family-wise Confidence Level in 2019

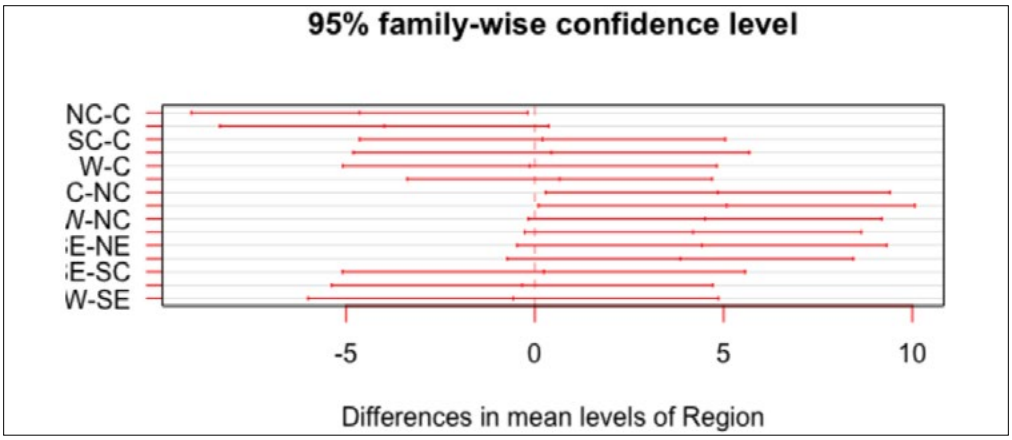

Fig 5. Plot of 95% Family-wise Confidence Level in 2020

Fig 6. Plot of 95% Family-wise Confidence Level in 2022

To examine whether the top three total scores' average and their variance could predict team performance, a multiple linear regression model was created for the 2022 season. Overall, the fitted regression model was of moderate strength and statistically significant at a 0.01 significance level. The average of top three scores was significantly predictive of the team performance, but the variance did not significantly predict the team score. Table 4 below includes the slope coefficient and p-value for both variables from the multiple linear regression model.

Table 4: Multiple Linear Regression Statistics

	Slope Coefficient	Standard Error	t-statistics	p-value
Intercept	1.789×10^{-2}	1.511	118.432	$<2 \times 10^{-16}$
Average MVP Score	3.878×10^{-2}	3.972×10^{-3}	9.762	3.62×10^{-15}
Variance	7.966×10^{-5}	8.126×10^{-5}	0.980	0.33

In order to visualize to what degree the actual scores fit into the multiple regression model, two ggplot were constructed for the linear model on each of the two independent variables, displayed in Fig 7 and Fig 8. Both figures' y-axis are labeled with Team Score; Fig 7's x-variable is Average MVP Score, while Fig 8's x-variable is Variance between Top 3 MVP Scores. The blue straight line in each graph is the linear regression model, or linear function, for predicting the team score based on the x-variable. The gray zone is the 95% confidence level for the linear model by default. The black dots represent actual data of participating college teams; the x-coordinate is the team's top three average score (Fig 7) or the variance between top three scores (Fig 8), while the y-coordinate is the team's actual average score during the season. In Fig 7, the black dots are moderately skewed, with about half of the points concentrated around the line, confirming the moderate strength of the model. In contrast, in Fig 8, the black dots are basically randomly scattered, suggesting the model is of weak to no strength and variance between top individual scores do not affect team scores.

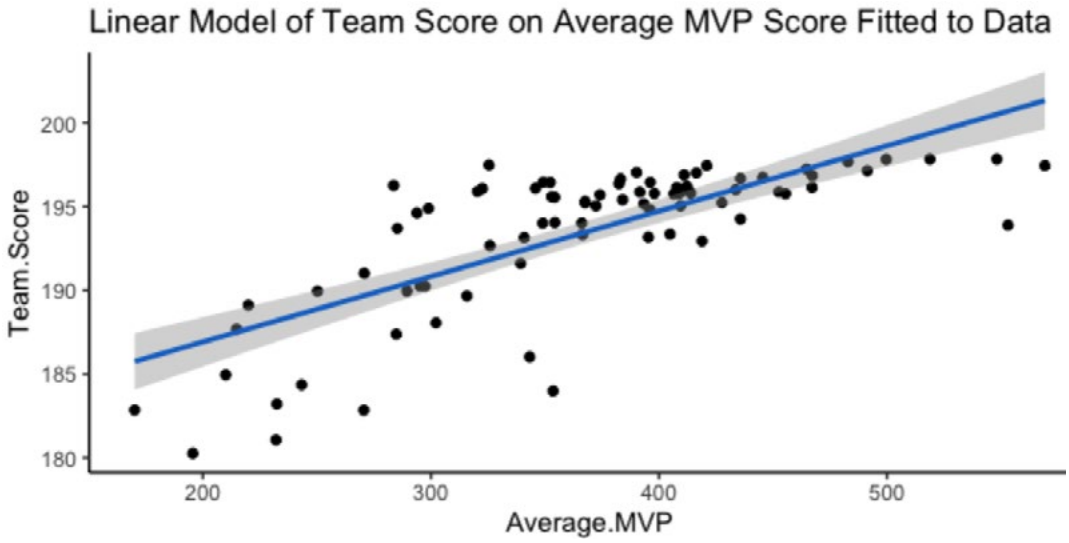

Fig 7. Linear Model of Team Score on Average MVP Score

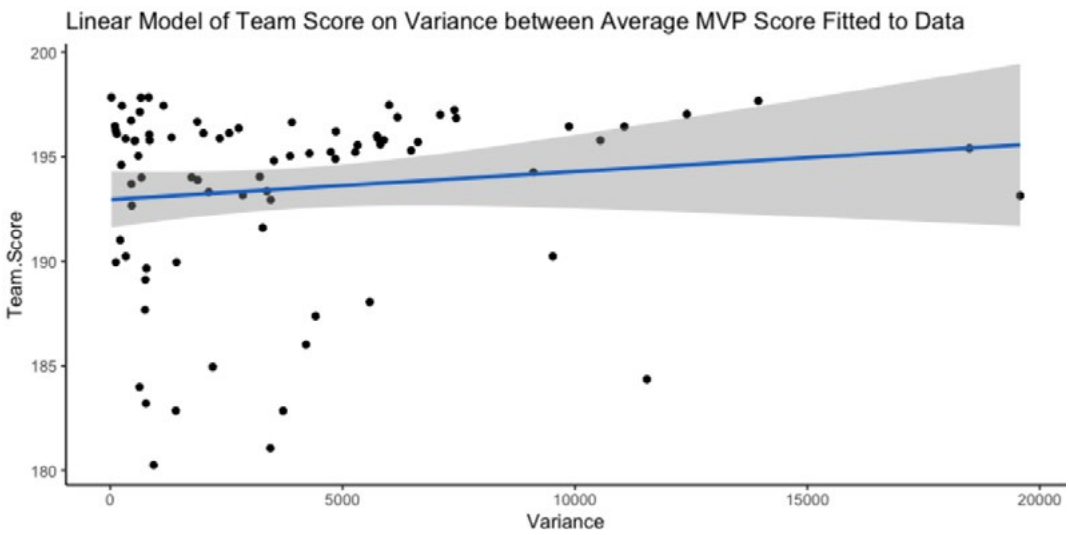

Fig 8. Linear Model of Team Score on Variance between Top 3 Average MVP Score

Discussion & Conclusion

This study found significant responses to the research questions of how differently teams perform due to their regions and top three scoring athletes. In response to the first research question, for all three seasons (2019, 2020, 2022), the one-way ANOVA model revealed average team scores were significantly different between at least two pairs of regions. In 2019, 2022, and 2022, there were 3 pairs, 3 pairs, and 2 pairs of regions that significantly differed in terms of mean team scores. Speaking of the multiple linear regression model for the second question, the average of top three scores in each team could significantly predict the team score. This suggests that top performing individuals were influential on a team's performance, especially when Olympians were competing for their schools. In contrast, the variance between the top three individuals did not affect the team performance noticeably.

It is possible that mean team scores of regions in NCAA women's artistic gymnastics varied from each other due to budget and financial situations. Before the pandemic, the North Central region's mean team score was significantly different with over half of the other teams. Multiple campuses of the University of Wisconsin (UW), included in the North Central region, might have caused this scenario. As Division III schools, UW campuses are relatively weak in terms of school budget and size, thereby differentiating the North Central region from the others. After the pandemic, Tukey's HSD test showed statistically significant differences between only two pairs of schools, the Northeast region included in both. Similar to the North Central region, the Northeast region has several Division III schools, such as Springfield College, that skew the region's mean team score lower. The finding in this project is thereby important, since it confirms that poorly-funded schools face financial barriers or limitations in recruiting, training, and eventually formal gymnastics competitions.

The research methods still have room for improvement. Specifically, to see the association between MVPs' and the entire team's performance, we could have considered each apparatus separately. However, this could not be achieved because not every athlete competed in all four apparatuses. We also could have taken into consideration the conditions of athletes at each game, but there is limited information about every athlete's injuries or physical fitness.

Overall, however, this study fills the expected gap in gymnastics-related research by taking into consideration a large-sized sample and using statistical

approaches. Current understanding of gymnastics concentrates on the physical aspect of athletes and research projects often study small samples. This project can lead to future looking at socioeconomic and geographical factors in gymnastics, analyzing which regions have limited resources, and considering regional differences in gymnastics through a more humane lens. The association between top individual scores and team performance might also be applicable to other sports, offering a new perspective to research on team sports, team games, and other activities involving teams.

Works Cited

1. Fredericks, M., & Fredericks, J. (2013). Road to nationals - NCAA gymnastics rankings. Retrieved August 15, 2022, from https://roadtonationals.com/results/

2. Gymnastics rules of the game. NCAA.org. (n.d.). Retrieved August 15, 2022, from https://www.ncaa.org/sports/2013/12/5/gymnastics-rules-of-the-game.aspx

3. Shasha, W., Abbasi, B. N., & Sohail, A. (2022). Assessment of Olympic performance in relation to economic, demographic, geographic, and social factors: quantile and Tobit approaches. Economic Research-Ekonomska Istraživanja, 1-25.