

## An exploration of gender gap using advanced data science tools: actuarial research community

Mengyu Yu<sup>1</sup> · Mazie Krehbiel<sup>1</sup> · Samantha Thompson<sup>1</sup> · Tatjana Miljkovic<sup>1</sup>

Received: 13 July 2019 / Published online: 14 March 2020

© Akadémiai Kiadó, Budapest, Hungary 2020

### Abstract

This paper explores the role of gender gap in the actuarial research community with advanced data science tools. The web scraping tools were employed to create a database of publications that encompasses six major actuarial journals. This database includes the article names, authors' names, publication year, volume, and the number of citations for the time period 2005–2018. The advanced tools built as part of the R software were used to perform gender classification based on the author's name. Further, we developed a social network analysis by gender in order to analyze the collaborative structure and other forms of interaction within the actuarial research community. A Poisson mixture model was used to identify major clusters with respect to the frequency of citations by gender across the six journals. The analysis showed that women's publishing and citation networks are more isolated and have fewer ties than male networks. The paper contributes to the broader literature on the “Matthew effect” in academia. We hope that our study will improve understanding of the gender gap within the actuarial research community and initiate a discussion that will lead to developing strategies for a more diverse, inclusive, and equitable community.

**Keywords** Gender · Network analysis · Poisson mixture · Web scraping · Actuarial

<sup>1</sup>Tatjana Miljkovic  
[miljkot@miamioh.edu](mailto:miljkot@miamioh.edu)

Mengyu Yu  
[yum4@miamioh.edu](mailto:yum4@miamioh.edu)

Mazie Krehbiel  
[krehbim@miamioh.edu](mailto:krehbim@miamioh.edu)

Samantha Thompson  
[thomp153@miamioh.edu](mailto:thomp153@miamioh.edu)

Miami University, Oxford,  
OH, USA

Full article available at:  
<https://doi.org/10.1007/s11192-020-03412-w>