

## Is it *Old or New*? Text Mining 90 Years of Scholarly Literature to Understand the Patterns in Student Loan Debt research<sup>1</sup>

Gaurav R. Sinha, University of Illinois at Urbana-Champaign<sup>2</sup>

The outstanding student loan debt in the U.S. has risen to \$1.46 trillion, a rise of \$79 billion in 2018 and a cumulative growth of 1.5 times in 11 years. Although education loans are not considered “bad debt” like credit card debt, concerns are raised that student loan debt can be an impending crisis to the nation’s economy. The primary objective of this study was to examine the nature and extent of scholarly work on student loan debt. Since the literature on the consequences of student loan debt on mental health of individuals is scant, the secondary purpose was to systematically review the literature related to this issue.

Two methods were used. First, since the literature on student loan debt has not been comprehensively reviewed, text mining was used to explore and understand the general overview and topical coverage of scholarly literature available on student loan debt. “All years” search in two academic databases resulted in 2,032 records. Only peer-reviewed unduplicated abstracts (n=988) with complete information were selected for further processing. Selected abstracts were analyzed using topic modeling, a machine learning tool that utilizes Bayesian statistics to extract hidden patterns in the collection of texts. Second, following PRISMA guidelines four databases – EBSCO, Web of Science, PubMed, and Scopus – were searched using variations of the phrase “student loan debt” and “mental health” to synthesize the existing student loan debt literature on mental health of individuals. A two-stage process with explicit inclusion and exclusion criteria was employed to identify relevant peer-reviewed papers.

A 10-topic model was selected for the reasons of parsimony and more accurate clustered representation of the abstracts. Funding and financing higher education (n=133) received the highest priority in the student loan debt literature. Another large set of studies in student loan debt literature came from the medical field (n=126). Repayment schemes, earnings, and effects of student loan debt (n=126) and borrowing, defaults, and employment (n=112) covered another big chunk. Youth, families, and wealth building (n=125) was also a topic of interest among researchers. Five topics received less attention in the literature – political and social (n=40), legal system, hardships, and repayments (n=67), tuition, costs, and future (n=83), programs (n=84), and mental health and wellbeing (n=92). In the systematic review, 46 peer-reviewed papers were included and categorized into mental health issues in two major life trajectories: (a) Before the repayment starts (n=27) (b) After the repayment starts (n=12). Sub-topical coverage in both the topics emphasized focus on stress and financial stress, wellbeing, and self-efficacy.

Both topic modeling and systematic review highlight the growing literature base on student loan debt and its effect on mental health. Across countries and across samples student loan debt is consistently and negatively related to different constructs and dimensions of mental health of individuals during different life trajectories. This finding is amplified for vulnerable populations, where women and racial/ethnic minorities more heavily feel the effects of education debt. Primary research on graduate students and individuals who have started repaying their loans and longitudinal studies with students who are still pursuing their degrees are still lacking and warrants further attention of the researchers.

---

<sup>1</sup> This abstract is a *work-in-progress* document. An elaborated version is prepared as a part of the author’s PhD program requirements.

<sup>2</sup> Gaurav R. Sinha (gsinha2@illinois.edu), PhD student, School of Social Work.