ProQuest Dissertations & Theses

Predicting Usage of Dissertations and Theses Across the Globe: An Inferential Analysis of ProQuest Data

Gilia C. Smith, Ph.D.
Director of Partnerships
July 10, 2023
PQDT Global - Connecting Global Scholarship

Collection + Research Tools
- **Graduate students**
  - Find a unique research topic for research
  - Build discipline expertise
  - Track research trends
- **Faculty/Researchers**
  - Source for text and data mining research
  - Stay current on a topic
  - Identify emerging research

Over 5.6 million records and
Over 3.2 million full text

60+ Countries represented in content
4.1K Universities’ content represented
4M Subscribing users
60+ languages represented in dissertation and theses records
Presentation Overview

1. What is our study about?
2. Why did we do it?
3. What did we learn?
4. Discussion: what are the implications?
What is the Study about?

... and WhoDunnit

A cross-sectional, multi-level regression analysis of dissertation and thesis usage in PQDT Global in 2021

Independent expert: Alfred DeMaris
Professor Emeritus Bowling Green University

Performed Statistical modeling and analysis


Clarivate Project lead: Gilia Smith, Ph.D.

Including topical experts from ProQuest content acquisition, operations and TDM teams.
Why Do It?
…and why now

Use rigorous social science methods and our global data cache to investigate how academic researchers use ETDs on the ProQuest Dissertations & Theses (PQDT) Global platform

Two levels of analysis

• **Level 1:** How background characteristics of the ETD itself impact usage

• **Level 2:** How background characteristics of the institution where the ETD was produced impact usage
COMING THIS MONTH!
ProQuest Dissertations & Theses Citation Index on the Web of Science

Research with a platform that now provides the most comprehensive view of how ideas have evolved, from dissertations and theses to patents and books.
Dissertations and theses provide foundational knowledge to researchers

Web of Science users explain why graduate work is important for understanding the background of a topic

“Sometimes theses reveal results that are not yet published in peer-reviewed literature”

“Look to dissertations for more details than can be found in journal articles, especially experimental details.”

“They are often a good resource for detailed materials and methods, and also unpublished results that may nevertheless be interesting and trigger further investigations.”
Embedding discovery of dissertations and theses into existing research workflows

Phase 1: Features released in July 2023

- Discovery of 5.7M+ metadata records for dissertations and theses in a standalone database
- Graduate works from 4K+ universities in 60+ countries around the world
- Add dissertation to Web of Science Author Profile
- For PQDT Global subscribers, links to 3M+ full text documents that provide detailed insights into methodologies, experimental outcomes and unpublished findings found only in the dissertation.

Phase 2: More features coming soon

- Cited reference indexing and linking will be available to further connect dissertations and theses to the Web of Science network
ProQuest Dissertations & Theses Citation Index on the Web of Science platform

Uncover hidden opportunities to advance research

- Identify emerging areas of research and the graduate programs that are engaged in this research.
- Understand the evolution of a topic or discipline by making connections between early career, emerging, and established research.
- Uncover innovative research methodologies not published in peer-reviewed literature.
- Execute more comprehensive, balanced literature reviews by including unpublished scholarship and more diverse perspectives.

Tap into a trove of graduate research only available in PQDT

One study of psychology dissertations found that only 25%* turn into published journal articles. This integration opens discovery of unpublished work not currently available on the Web of Science platform.

Sample Size

N = 3,548,714 observations

• Model built with extremely high levels of precision

• p<.01 = alpha level for significance

• 95% Confidence Levels

Outcome Variables

1. Total Usage: indicates researcher viewed metadata, citation, abstract, print preview or full-text document—or perhaps one or more of the above.

2. Abstract/Citation Usage: indicates the end-user accessed the abstract and/or citation

3. Full-text Usage: indicates the user viewed or downloaded the full-text document
What did we learn?

- **Degree Type** Masters -17%
- **Full-text** +15%
- **Subjects**
  - Social Sci +14%
  - Arts & Humanities +13%
  - Interdisciplinary +10%
  - Professional Fields +8%
  - Earth Sci +1%
  - Math & Physical Sci -5%

Holding all other variables constant...

| Effect                   | Estimate | t Value | Pr > |t| |
|--------------------------|----------|---------|------|---|
| Intercept                | 1.0508   | 61.3    | <.0001 |
| Masters Thesis           | -0.1874  | -80.86  | <.0001 |
| Full text                | 0.1401   | 44.16   | <.0001 |
| Social Sciences          | 0.1311   | 32.86   | <.0001 |
| Arts & Humanities        | 0.1184   | 27.82   | <.0001 |
| Interdisciplinary        | 0.09467  | 14.16   | <.0001 |
| Professional Fields      | 0.07445  | 19.32   | <.0001 |
| Earth Sciences           | 0.01343  | 3.35    | 0.0008 |
| Math & Physical Sciences | -0.05276 | -12.25  | <.0001 |
What did we learn?

- **Degree Age**
- **Language of ETD**
  - Other -9%
  - Chinese +10%
  - German -5%
  - French = English
  - Swedish = English
  - Spanish = English

Holding all other variables constant...

| Effect                        | Estimate | t Value | Pr > |t| |
|-------------------------------|----------|---------|------|---|
| Intercept                    | 1.0508   | 61.3    | <.0001 |
| Degree Age                   | -0.00753 | -26.04  | <.0001 |
| Other Language (x51)         | -0.09186 | -5.49   | <.0001 |
| Chinese                      | 0.1046   | 4.19    | <.0001 |
| German                       | -0.05057 | -2.74   | 0.0061 |
| French                       | -0.02635 | -2.55   | 0.0107 |
| Swedish                      | -0.06975 | -2.42   | 0.0157 |
| Spanish                      | -0.02074 | -1.68   | 0.0938 |
What did we learn?

- Current ProQuest Contributor +23%
- Quantity of ETDS produced -1.3%/1,000
- Region/Country – UK less use than rest of world
- Shanghai Rank -.2%/each drop in rank

### Solution for Fixed Effects

| Effect                  | Estimate | t Value | Pr > |t| |
|-------------------------|----------|---------|------|---|
| Intercept               | 1.0508   | 61.3    | <.0001 |
| Current PQ Contributor  | 0.2091   | 14.99   | <.0001 |
| Rank                    | 0.00043  | 3.69    | 0.0002 |
| CurrentContributor*Rank | -0.00274 | -21.68  | <.0001 |
| United Kingdom          | -0.17    | -11.15  | <.0001 |
| APAC                    | -0.05363 | -2.1    | 0.0359 |
| Africa                  | -0.1167  | -1.55   | 0.1212 |
| Latin America           | 0.1071   | 1.06    | 0.2915 |
| Caribbean & West Indies | -0.2256  | -0.97   | 0.3335 |
| Europe                  | -0.00798 | -0.48   | 0.6335 |
| Middle East             | 0.008486 | 0.15    | 0.8839 |
| Australia/NZ            | -0.00418 | -0.08   | 0.9341 |
Discussion

What implications might this study have for practice at your institution?

Gilia C. Smith, Ph.D.
Director of Partnerships
ProQuest, a part of Clarivate
gilia.smith@clarivate.com
Thanks for Attending our Breakfast!

Questions about the study for Al DeMaris? Email him directly at ademari@bgsu.edu

Gilia C. Smith, Ph.D.
Director of Partnerships
ProQuest, a part of Clarivate
gilia.smith@clarivate.com