

Modelling the Academic Publishing System: A Data-Driven Agent-Based Approach

**A dissertation submitted in partial fulfillment of the
Requirements for the degree of
DOCTOR OF PHILOSOPHY
in
Information Technology
by
Xin Gu**

**The University of Newcastle
Australia**

2017

Declaration

I hereby certify that this thesis is submitted in the form of a series of published papers of which I am a joint author. I have included as part of the thesis a written statement from each co-author; and endorsed by the Faculty Assistant Dean (Research Training), attesting to my contribution to the joint publications.

Acknowledgements

First of all, my sincere appreciation goes to my principal supervisor, Dr. Karen Blackmore. Thank you for giving me the opportunity to start this PhD work. You gave me the liberty to follow my own interest while at the same time you always guided and supported me throughout my entire PhD journey. It is always a pleasure to work with you. I am grateful for your detailed feedback and suggestions for improvement throughout all my journal and conference papers. In addition, the support from two co-supervisors, Dr. Keith Nesbitt and Dr. David Cornforth, is also highly appreciated. And to my friends at UoN, you have all made my PhD journey more enjoyable.

From the bottom of my heart, I would like to thank my parents, Mr. Jinqun Gu and Ms. He Huang. I thank you both for your unconditional love and endless support. I especially thank my father for guiding me step by step in my life. And I thank my mother for teaching me to be happy and optimistic. I hope you are proud of me.

Most importantly, my deep appreciation goes to Mr. Stephen Bancroft, the other half in my life. Without you, there would be no thesis. Your 'brain-mining' concept and never-give-up spirit will always guide me through my career. Indeed, you help me fulfil all my dreams. I couldn't find any words to express my gratitude for your love. Thank you for making everything possible for us.

A very special thank you to my baby-to-be, Alice Mei-Ling Bancroft, my little treasure. You are my inspiration. It has been an amazing and special journey to share this eight-month PhD time with you. You are so sweet and patient that you put up with me working on the thesis day and night for so long. Thank you for encouraging me to strive for excellence. My 'text-book' baby, you will always lighten up my life.

Table of Contents

Declaration.....	2
Acknowledgements.....	3
Table of Contents.....	4
Abstract.....	6
List of Publications	7
List of Abbreviations	9
Glossary	10
1 Introduction	14
1.1 Background	14
1.2 Academic Publishing System.....	15
1.3 Research Questions and Objectives	16
1.3.1 Empirical Analysis.....	17
1.3.2 Agent-Based Modelling and Simulation (ABMS)	18
1.4 Contribution of the Thesis.....	18
1.4.1 Contributing Publications	18
1.4.2 Supporting Publication.....	22
1.5 Summary	23
2 Literature Review	25
2.1 Agent-based Modelling and Simulation (ABMS).....	25
2.2 Advantages of ABMS	26
2.3 Limitations of ABMS.....	26
2.4 ABMS Applications in an Academic Publishing Context	27
2.5 ABM Development	28
2.5.1 Conceptual Model Development.....	28
2.5.2 Empirical Data Analysis	29
2.5.3 ABM Implementation	29

2.5.4	ABM Validation.....	30
2.5.5	Simulation Experiments	31
2.6	Academic Publishing System	31
2.7	Models of Academic Science	33
2.8	Summary	34
3	Conclusion.....	36
3.1	Research Conclusions.....	36
3.2	Contributions	37
3.2.1	Contribution to Knowledge on ABMS Applications in the Higher Education Domain...37	
3.2.2	Contribution to Knowledge on Understanding Validation Requirements for an ABAPM 37	
3.2.3	Contributions to Knowledge on Recent Trends in Academic Journal Growth	38
3.2.4	Contribution to Knowledge on Characterisations of Academic Journals.....	38
3.2.5	Contribution of a Consolidated Dataset of Academic Journals.....	38
3.2.6	Contribution to Knowledge on Quantitative Relationships among Journal Characteristics 39	
3.2.7	Contribution to Knowledge on Scholar Typology of Publishing Patterns	40
3.2.8	Contribution of a Consolidated Dataset of Scholars with Publishing Pattern	41
3.2.9	Contribution to Knowledge on Quantitative Study of Australian Academic Science.....	41
3.2.10	Contribution to Knowledge of a Conceptual Model of the Academic Publishing System 42	
3.2.11	Contribution of an ABAPM.....	44
3.3	Limitations.....	45
3.4	Future Work	46
4	References for Thesis	49
5	References for Publications.....	52

Abstract

Academic publication is used to disseminate and advance knowledge. It also provides a measure of individual research performance for career advancement and personal prestige. Furthermore, it is one measure of university performance and productivity for global rankings and government policy purposes. This research is motivated by the importance of academic publication to these various stakeholders, and aims to address the following research question:

How can the academic publishing system be explained by modelling the strategic behaviours of scholars and academic journals?

This research uses existing literature to define the strategic publishing behaviours of scholars and the strategic approaches of academic journals associated with publication. These investigations are then used to develop a conceptual model of the academic publishing system. The conceptual model is then operationalised using an agent-based modelling and simulation approach. Finally, real-world data is used to parameterise and validate this model.

This research question is addressed through the empirical analysis of bibliometric data, and the implementation of the agent-based model. The analysis aims to fully understand the quantitative status of the academic publishing system, in order to calibrate and validate the agent-based model. Academic journals are characterised using 13 attributes that are captured across bibliographic data sources. The quantitative relationships among these journal characteristics are analysed over four impact-oriented journal quartiles to further characterise journal types. Cluster analysis of data collected for 11,427 scholars, and their 284,128 journal publications, identifies six types of scholars: *singleton*, *small-team low performer*, *small-team high performer*, *big-team strategist*, *free-style follower*, and *life-time warrior*. The empirical analysis identifies new, previously undescribed, growth patterns of academic journals and scholars, and contributes a characterisation of scholar types that provides a robust and holistic conceptualisation of academic publishing behaviours.

A methodological approach of coupling the agent simulation with empirical research is introduced. It uses empirical data to set the parameters of the conceptual agent-based model that best reflects the scholars' strategic publishing-decision processes. The simulation results validate independent historical data at both macro- (growth patterns of scholars and their publications, as well as the performance measures of scholars) and micro-levels (number of scholars and their publications for each calendar year). Thus, the thesis contributes a new typology of academic publishing strategy, and empirical journal characterisation. This informs a new understanding of the individual and collective behaviours that lead to a variability in research outcomes between scholar types. Importantly, it provides a framework that can be used to optimise the distribution of scholar types within teams and institutions, with the aim of maximising research outputs.

List of Publications

Contributing Publications

	Publication Details	Journal Ranking
1	Gu, X., & Blackmore, K. (2015). A Systematic Review of Agent-Based Modelling and Simulation Applications in the Higher Education Domain. <i>Higher Education Research & Development</i> , 34(5), 883-898.	ERA 2010: A SJR h-Index: 22 Citation: 5 Altmetrics: 2
2	Gu, X., Blackmore, K., Cornforth, D., & Nesbitt, K. (2015). Modelling Academics as Agents: An Implementation of an Agent-Based Strategic Publication Model. <i>Journal of Artificial Societies and Social Simulation</i> , 18(2), 10.	ERA 2010: C SJR h-Index: 34 Citation: 3 Altmetrics: Null
3	Gu, X., & Blackmore, K. (2016). Recent Trends in Academic Journal Growth. <i>Scientometrics</i> , 108(2), 693-716.	ERA 2010: A SJR h-Index: 78 Citation: 4 Altmetrics: 7
4	Gu, X., & Blackmore, K. (2017a). Characterisation of Academic Journals in the Digital Age. <i>Scientometrics</i> , 110(3) 1333-1350.	ERA 2010: A SJR h-Index: 78 Citation: Null Altmetrics: 25
5	Gu, X., & Blackmore, K. (2017b). Toward a Broader Understanding of Journal Impact: Measuring Relationships Between Journal Characteristics and Scholarly Impact. <i>World Academy of Science, Engineering and Technology, International Journal of Social, Behavioral Educational, Economic, Business and Industrial Engineering</i> , 11(10), 2144-2149.	Conference Journal No Ranking

- 6 Gu, X., & Blackmore, K. (forthcoming). Developing a Scholar Typology from Publication Patterns in Academic Science: A Cluster Analysis Approach. *Journal of the Association for Information Science and Technology*. (Under Review) ERA 2010: A*
SJR h-Index: 100

- 7 Gu, X., & Blackmore, K. (2017c). Quantitative Study on Australian Academic Science. *Scientometrics*. (2017): 1-27. ERA 2010: A
SJR h-Index: 78

- 8 Gu, X., & Blackmore, K. (forthcoming). Insight into Scholarly Performance: An Agent-Based Model of the Academic Publishing System. *Journal of Artificial Societies and Social Simulation*. (Under Review) ERA 2010: C
SJR h-Index: 34

Supporting Publications

	Publication Details	Journal Ranking
S1	Gu, X., & Blackmore, K. (2014). <i>The publishing game: An analysis of "Game" related academic publishing patterns</i> . Paper presented at Interactive Entertainment 2014 , University of Newcastle, Australia.	Conference

List of Abbreviations

Abbreviation	Full Term
ABM	Agent-Based Model
ABMS	Agent-Based Modelling and Simulation
ABAPM	Agent-Based Academic Publishing Model
Cabells	Cabell's Periodical Directory
GS	Google Scholar Metrics
JCR	Journal Citation Reports
SJR	SCImago Journal & Country Rank
Ulrichs	Ulrich's Periodicals Directory
WoS	Web of Science

Glossary

Term (word or phrase)	Explanation in plain language
Agent-based modelling and simulation (ABMS)	Agent-based modelling and simulation (ABMS) is a computational approach to modelling systems comprises of interacting, autonomous, decision-making ‘agents’ or individuals.
Agent	Within an ABMS, agents can be representations of any type of self-governing individuals or entity with behaviours described using simple rules.
Careerist	It is a scholar type in the hypothetical conceptual ‘strategic publication model’ from Mö lders, Fink, and Weyer (2011). This type of scholars focus publishing in higher ranked journals.
Orthodox scientist	It is a scholar type in the hypothetical conceptual ‘strategic publication model’ from Mö lders et al. (2011). This type of scholars focus publishing in their interests.
Mass producer	It is a scholar type in the hypothetical conceptual ‘strategic publication model’ from Mö lders et al. (2011). This type of scholars focus publishing more papers without considering the quality of the journals.
Acceptance rate (AR)	It is a descriptive attribute of scholarly journals. Acceptance rate presents the publishing percentage of all received manuscripts from a scholarly journal.
Impact factor (IF)	It is a descriptive attribute of scholarly journals. Impact factor presents the quality of the journals. The calculation of the impact factor is based on the citations the journal received. Different bibliographical databases have their own way to calculate impact factor.

Topical fitting factor (TFF)	It is a descriptive attribute of scholarly journals. Topical fitting factor presents the topics of the journals. Some journals have very wide topical fittings , but some have narrow topical fittings for their publishing papers.
Leading university	In this research, top #200 universities ranked in Global Best University SQ ranking are categorised in the leading university group.
Middle-tier university	In this research, top #201 - #400 universities ranked in Global Best University SQ ranking are categorised in the middle-tier university group.
Non-comprehensive university	In this research, top #401-#600 universities ranked in Global Best University SQ ranking are categorised in the non-comprehensive university group.
Effective scholars	The <i>Scopus</i> scholars having at least one journal article, and with an <i>h</i> -index value above zero are referred as effective scholars.
Non-effective scholars	The <i>Scopus</i> scholars having no publications in journals and / or have null <i>h</i> -index values, due to no citations of their articles are referred as non-effective scholars.
Q1	Journals in SCImago are analysed based on <i>SJR Best Quartile</i> , which is based on the scholarly impact distribution that a journal occupies in its subject categories. Q1 journals occupies the top 25% of the scholarly impact distribution.
Q2	Q2 occupies the middle-high position between top 50% and top 25% in the <i>SJR Best Quartile</i> ranking.
Q3	Q3 occupies the middle-low position between top 75% and top 50% in the <i>SJR Best Quartile</i> ranking.

Q4	Q2 occupies the bottom position occupying the bottom 25% of the scholarly impact distribution.
Singleton	Scholars focusing on publishing individually are referred as singletons. 8% of scholars are classified in this group.
Small-team low performer	Scholars, who prefer to work in a small team of two or three scholars, and the quality and quantity of their research outputs are both comparatively slow, are defined as small-team low performers. 16% of scholars belong to this group.
Small-team high performer	Scholars, who prefer working in a small team of 2-3 scholars with research outputs found predominantly in Q1 journals with high impact, are referred as small-team high performers. 17% scholars belong to this group.
Big-team strategist	Scholars, who tend to work in a team with more than six authors with majority of the work published in Q1 journals, are referred as big-team strategists. 22% of scholars are classified in this group.
Free-style follower	Scholars, who show no preference for any particular research team size, instead equally likely work in all size teams, are referred as free-style followers. 21% scholars are classified in this group.
Life-time warrior	Scholars, who has the long publishing history (average 23.49 years) with the large number of publications (average 80.81), are referred as life-time warriors. 17% of the entire scholar population belong to this group.
Emerging scholar	Scholars with publishing age less than 4 years are referred as emerging scholars.
Early career scholar	Scholars with publishing age greater than or equivalent to 4 but less than 10 years are referred as early career scholars.

Mid-career scholar	Scholars with publishing age greater than or equivalent to 10 but less than 20 are referred as mid-career scholars.
Late mid-career scholar	Scholars with publishing age greater than or equivalent to 20 but less than 30 are referred as late mid-career scholars.
Senior scholar	Scholars with publishing age greater than or equivalent to 30 but less than 40 are referred as senior scholars.
Life-time scholar	Scholars with publishing age greater than or equivalent to 40 are referred as life-time scholars.
Emerging phase (LCP1)	All publications are categorized into six groups based on different life cycle phase (LCP). LCP 1 contains papers published when authors had less than four years publishing age.
Early career phase (LCP2)	LCP 2 contains papers published when authors had greater than or equal to four but less than 10 years publishing age.
Mid-career phase (LCP3)	LCP 3 contains papers published when authors had greater than or equal to 10 but less than 20 years publishing age.
Late mid-career phase (LCP4)	LCP 4 contains papers published when authors had greater than or equal to 20 but less than 30 years publishing age.
Senior phase (LCP5)	LCP 5 contains papers published when authors had greater than or equal to 30 but less than 40 years publishing age.
Life-time phase (LCP6)	LCP 6 contains papers published when authors had greater or equal to 40 years publishing age.
