

The Relationship Between RGC Funding and Research Productivity Among STEM Scholars in Hong Kong: A Panel Data Analysis (2012–2022)

by

Neel Gautam Bagai

Master of Education (Higher Education)

Faculty of Education

The University of Hong Kong

Supervisor: Professor Hugo Horta

July 2026

Acknowledgements

I would like to express my sincere gratitude to my capstone supervisor, Professor Hugo Horta. Beyond his role as supervisor, Professor Horta taught me during my Master of Education and his continuous guidance, structured feedback, and mentorship were vital to the successful completion of this study.

I am equally grateful to Professor Jisun Jung, who served as both my ethical reviewer and the coordinator of my specialism. Her instruction throughout my Master of Education, alongside her support and leadership in this role, shaped both my research practice and my development as a scholar.

My appreciation extends to the faculty at the University of Hong Kong who shaped my development as a researcher. I am particularly grateful to Dr. Juuso Henrik Nieminen, Professor Clarence Green, Professor Lili Yang, and Professor Jeremy Rappleye for their exceptional instruction and inspiration throughout my degree.

The scale of data collection this project required presented a significant technical challenge. I extend my gratitude to Paul from Elsevier's customer support team, who provided multiple API keys with unlimited calls for the parallel scraping of Scopus data. Without his timely assistance this project would not have been completed on schedule. I am equally grateful to the NamSor team for generously providing a free API key with extended limits.

Finally, a special mention goes to my family. Beyond their unwavering personal support, I owe my parents a massive thank you for graciously footing the electricity bill while I had two computers running simultaneously, twenty-four hours a day for an entire month to collect this data.

Abstract

Hong Kong presents a paradox that demands empirical investigation. The city hosts five universities ranked in the global top 100, receives approximately HK\$1.18 billion annually in competitive research funding from the Research Grants Council (RGC), and sustains some of the highest rates of international collaboration in Asia. Yet its scientific competitiveness has weakened steadily since the early 2000s (Horta, 2018). This study provides the first scholar-level causal evaluation of how RGC funding shapes STEM scholar research productivity in Hong Kong, drawing on a panel of 22,456 scholars and 247,016 scholar-year observations spanning 2012 to 2022. It uses propensity score matching (PSM), two-way panel fixed effects (TWFE), seemingly unrelated regression (SUR), and event-study analysis to cleanly separate the true marginal impact of funding from the inherent selection bias that channels grants toward already productive scholars.

The findings reveal a clear quality-quantity trade-off: grant receipt significantly increases publication volume ($\beta = +1.084$, $p < 0.001$) and short-term field-normalised citation impact ($\beta = +1.663$, $p = 0.006$), whilst simultaneously reducing the share of publications placed in Q1 journals ($\beta = -0.055$, $p < 0.001$) and average journal prestige ($\beta = -0.310$, $p < 0.001$).

This decoupling of citation impact from journal prestige carries important implications for how research quality is measured. It rejects the hypothesis that RGC-funded scholars produce high-volume but scientifically hollow output and instead indicates that funded work retains measurable utility in the citation record regardless of journal tier. Whether this citation premium reflects the intrinsic quality of the research or the expansion of co-authorship networks enabled by grant capital remains an open question for future investigation. Diminishing-returns thresholds of approximately HK\$ 5.22 million for publication volume and HK\$ 3.97 million for short-term citation quality mark the points at which additional cumulative funding begins to erode marginal returns. Funding effects differ substantially across career stage, discipline, and institution type, with directional but statistically inconclusive evidence for gender heterogeneity. Given this, the study argues that continued public investment in competitive grants is justified, whilst portfolio diversification and field-calibrated accountability represent the reforms most likely to maximize returns on Hong Kong's substantial STEM research investment.

Keywords: research funding, research productivity, propensity score matching, bibliometrics, Hong Kong, panel data, diminishing returns, quality-quantity trade-off

Table of Contents

1.0 – Introduction	1
1.1 – The Hong Kong Context	1
1.2 – The RGC Funding Mechanism	1
1.3 – The Research Problem	2
1.4 – Research Questions	2
1.5 – Significance and Preview	2
2.0 – Literature Review	3
2.1 – Theoretical Framework	3
2.1.1 – The Matthew Effect	3
2.1.2 – Principal-Agent Theory and Information Asymmetry	4
2.1.3 – Diminishing Returns in Research Production	4
2.2 – Measurement of Research Productivity	5
2.3 – Empirical Evidence on the Funding-Productivity Relationship	5
2.4 – Methodological Approaches in Prior Research	7
2.5 – Research Gap and Contribution	8
2.6 – Expected Findings	9
2.6.1 – RQ1: The Quality-Quantity Tradeoff	9
2.6.2 – RQ2: Diminishing Returns to Funding Intensity	10
2.6.3 – RQ3: Heterogeneity in Treatment Effects	10
3.0 – Methodology	10
3.1 – Research Design	10
3.2 – Data Sources	11
3.3 – Sample Construction	12
3.4 – Variables	12
3.5 – Propensity Score Matching	17
3.6 – Empirical Strategy	18
3.7 – Limitations	21
3.7.1 – Data Measurement Vulnerabilities	21
3.7.2 – Methodological and Model Constraints	22
3.7.3 – The Boundaries of Interpretation	23

4.0 – Overview of Findings	24
4.1 – Sample and Matching Overview	24
4.2 – The Quality-Quantity Trade-Off	24
4.3 – Diminishing Returns to Funding Intensity	25
4.4 – How Effects Vary Across Scholars	25
4.4.1 – Career Stage	25
4.4.2 – Gender	25
4.4.3 – Discipline	25
4.4.4 – Funding Scheme: ECS vs GRF	25
4.4.5 – Institution Type	25
4.4.6 – Summary	25
4.5 – Broader Behavioral Patterns	26
5.0 – Empirical Results	26
5.1 – Descriptive Statistics and Covariate Balance	26
5.2 – RQ1: TWFE Estimates and Parallel Trends Validation	30
5.3 – RQ2: Quadratic Specification and Optimal Thresholds	32
5.4 – RQ3: Heterogeneity Analysis	34
5.4.1 – Career Stage	34
5.4.2 – Gender	35
5.4.3 – Discipline	36
5.4.4 – Funding Scheme: ECS vs GRF	37
5.4.5 – Institution Type	37
5.4.6 – Summary of Heterogeneity Results	37
5.5 – Robustness and Supplementary Analyses	39
5.5.1 – Alternative Specifications	40
5.5.2 – Estimator Diagnostics	42
5.5.3 – Supplementary Behavioral Analysis	42
6.0 – Discussion	44
6.1 – Theoretical Interpretation	44
6.1.1 – RQ1: The Quality-Quantity Trade-Off	44
6.1.2 – RQ2: Diminishing Returns to Funding	45

6.1.3 – RQ3: Heterogeneity Across Scholar Types	46
6.2 – Policy Implications	46
6.2.1 – Funding Intensity and the F^* Threshold	46
6.2.2 – ECS Scheme Design	47
6.2.3 – Capital Allocation and Institution Type	47
6.3 – Comparison with Prior Literature and Hypotheses.....	48
7.0 – Conclusion	50
7.1 – Summary of Findings and Conclusions	50
7.2 – Contributions to Knowledge	51
7.2.1 – Methodological Contributions	51
7.2.2 – Empirical Contributions.....	52
7.2.3 – Theoretical Contributions	52
7.3 – Future Research Directions	53
7.4 – Policy Recommendations.....	53
References	55

1.0 – Introduction

1.1 – The Hong Kong Context

In an era where scientific output increasingly determines national competitiveness, Hong Kong (HK) faces a troubling paradox. Despite having five of the eight publicly funded universities consistently ranked in the global top 100 (Times Higher Education, 2025) and maintaining one of Asia's highest rates of international research collaboration (Horta, 2018), the city's scientific wealth is in decline. Horta (2018), through a revealed competitive advantage analysis spanning from 1999 to 2016, shows that HK's relative research competitiveness has weakened substantially since the early 2000s, with the city falling behind in key measures it once dominated. Singapore, by contrast, has maintained or strengthened its position over the same period, investing 2.18 percent of GDP in research & development (R&D) against HK's 0.78 percent (Horta, 2018). This trend is still evident, with empirical evidence showing that HK's Gross Expenditure on Research & Development (GERD) stands at just 1.11 percent of GDP in 2023 (Census and Statistics Department, 2024), trailing regional competitors such as Singapore (1.85%), Japan (3.44%) and South Korea (4.96%) (UNESCO Institute for Statistics, 2026). This perspective implies that if underinvestment is the root cause, injecting more capital into the system should theoretically restore HK's research competitiveness. This divergence also raises a fundamental question: is HK's research funding mechanism structured to sustain the territory's scientific standing in an era of intensifying global competition for knowledge production (Sum & Jessop, 2013)?

1.2 – The RGC Funding Mechanism

The Research Grants Council (RGC) is HK's primary government-funded body responsible for allocating competitive research grants to academics across HK's University Grants Committee (UGC)-funded institutions. In 2024/25, the RGC distributed approximately HK\$1.18 billion (~US\$152 million) through its two investigator-initiated funding schemes (University Grants Committee, 2025b), namely the General Research Fund (GRF), which is open to established academics, and the Early Career Scheme (ECS), which targets researchers within three years of their first academic appointment, shielding them from direct competition with established academics in the GRF pool. Both funding schemes subject proposals to rigorous peer review by international referees and discipline-specific panels that assess scientific merit. In the past decade, securing funding from the RGC has become increasingly difficult with application volumes growing substantially. GRF applications rose 10 percent and ECS applications 21 percent in the 2024/25 cycle alone, against a broadly stable success rate of approximately 30 percent (University Grants Committee, 2024). This mirrors international trends observed at other funding bodies such as the National Institutes of Health (NIH), where success rates have declined to approximately 20 percent as application volumes continue to outpace available funding (National Institutes of Health, 2026). The implicit policy assumption underpinning competitive grant allocation is simple: directing public funds to the proposals with the highest scientific merit should yield superior research output. Whether this assumption holds in practice, and whether the relationship between

funding and productivity is linear, diminishing or heterogeneous across scholar characteristics, remains the central empirical question this study addresses.

1.3 – The Research Problem

The central paradox motivating this study is the measurable decline in HK's global scientific standing despite maintaining substantial public investment in research. Two competing theoretical accounts frame this paradox. Horta (2018) points to critical underinvestment in HK's research system relative to competitors, corroborated with HK's 2023 GERD of 1.1 percent (Census and Statistics Department, 2024), the lowest among regional competitors. Whether this underinvestment translates into reduced individual scholar productivity and through what mechanism is precisely the empirical question this study addresses. Tammi (2009), writing about Finnish universities, advances the impoverishment thesis, articulating that the very shift toward competitive, project-based funding generates perverse incentives by diverting researchers from fundamental inquiry toward safe, publishable projects and by imposing administrative burdens that consume effective research time. Between these theories sit several unresolved empirical questions. Does RGC funding primarily increase publication volume, or does it also improve quality as captured by journal prestige and citation impact? Is the marginal return to additional funding constant, or does it diminish beyond some optimal threshold? Do these effects vary systematically across career stages, disciplines, gender and institutional contexts? These raise an important question: can a single competitive funding mechanism simultaneously serve the goals of quantity, quality, and equity, or does it structurally trade one off against another?

1.4 – Research Questions

Three interrelated research questions guide the empirical analysis, each designed to isolate a distinct mechanism through which competitive funding shapes scholarly productivity.

RQ 1: Does RGC funding predict publication in high impact (Q1-ranked) journals, or does it primarily increase publication volume without improving research quality?

RQ 2: What is the relationship between funding intensity (cumulative funding amount) and research productivity, and is there evidence of diminishing marginal returns?

RQ 3: How does the effect of RGC funding vary across scholar characteristics (career stage, gender), funding schemes (ECS vs. GRF), disciplines, and institution types?

1.5 – Significance and Preview

The contribution of this study is three-fold. First, it offers direct policy relevance by providing empirically grounded estimates of the return on public research investment in HK, with implications for both RGC allocation strategy and institutional research management. Second, it advances the methodological toolkit for funding evaluation by integrating propensity score matching (PSM), two-way panel fixed effects (TWFE), seemingly unrelated regression (SUR), and event-study analysis within a single analytical framework, with each layer addressing a valid

complementary threat to causal validity. Third, it provides the first scholar-level panel study of research funding effects in HK, tracking 22,456 STEM scholars across eleven years (2012-2022) and 247,016 scholar-year observations. The remainder of this capstone proceeds as follows: Section 2 reviews the theoretical and empirical literature, Section 3 details the methodology, Section 4 provides an overview of findings, Section 5 presents the full empirical results, Section 6 discusses implications and Section 7 concludes.

2.0 – Literature Review

2.1 – Theoretical Framework

The relationship between competitive research funding and scholarly productivity can be examined through three theoretical lenses, each generating distinct and testable predictions about the mechanisms investigated in this study.

2.1.1 – The Matthew Effect

Merton (1968) coined the term “Matthew Effect” to describe a self-reinforcing cycle of inequality in science. Under this dynamic, established scientists gain disproportionate recognition for the exact same contributions made by their lesser-known peers. Through interviews with Nobel laureates, Merton (1968) demonstrated that scientific recognition is not a passive reward for past work but is a signal that actively generates future resources. Eminent scientists attract visibility, which in turn amplifies access to further resources, talented collaborators, and subsequent recognition. This cumulative advantage extends beyond reputation to dictate the material realities of research production. As Merton observed, the scientific reward system structurally hoards investigative opportunities for the established ‘elite’ scientists. This stratification becomes decisive as modern inquiry shifts from “little science” to “big science,” where producing competitive output strictly requires access to massive, centralized infrastructure.

Bol et al. (2018) provided rigorous empirical validation of this mechanism in the specific domain of research funding. Exploiting a regression discontinuity in scores from the Dutch Research Council (NWO), they demonstrated that researchers who narrowly won early-career grants accumulated approximately twice as much subsequent funding as those who narrowly lost. Over eight years, this head start generated a funding gap of approximately €180,000 between narrow winners and narrow losers. Crucially, the authors found this cumulative advantage was driven by a participation mechanism, with early funding success increasing a scholar’s probability of applying for future grants from 40 percent to 59 percent (Bol et al., 2018, p. 4889). Horta et al. (2018) extended this structural divergence to the actual production of scientific knowledge. Their analysis revealed that securing stable, competitive funding at the earliest career stages triggers a permanent divergence in academic trajectories, strictly elevating a scholar’s research productivity and citation visibility throughout their entire career.

Together, these dynamics confirm that early funding success systematically concentrates both the financial capital and the reputational visibility that drive long-term stratification in modern academic research. Given this empirical consensus, the Matthew Effect operates as a measurable mechanism of advantage rather than a theoretical construct. Initial grant receipt triggers self-reinforcing cumulative advantage loops, providing funded scholars with preferential access to the resources and elite collaborators that systematically accelerate future productivity. This compounding dynamic provides the primary interpretive framework for RQ2's investigation into the effects of funding intensity.

2.1.2 – Principal-Agent Theory and Information Asymmetry

Competitive grant allocation creates a classic principal-agent relationship. The funder (the RGC) cannot directly observe the quality, creativity, or long-term significance of the research it finances. Instead, it relies on observable proxies such as prior publication records, citation counts, institutional affiliation and the quality of the written proposal. The information asymmetry generates well-documented risks. Stephan (2012) warns that when quality is difficult to observe, researchers may rationally prioritize safe, incremental projects that produce output with a higher likelihood of getting published rather than pursue higher-risk, potentially transformative work. Li & Bai (2025), studying NIH-funded principal investigators, found that grant renewal success significantly reduced researchers' probability of switching research focus and decreased research diversity, though it simultaneously increased research novelty.

When researchers face pressure to demonstrate immediate return on investment to secure future funding, they may opt for faster-turnaround, lower-tier journals to guarantee output volume (Stephan, 2012). This is not a failure of individual motivation but a predictable outcome of the incentive architecture. The principal-agent framework predicts a structural quality-quantity tradeoff as a systemic feature of competitive funding, which is the exact distinction that RQ1 investigates using Q1 journal ratios alongside raw publication counts.

2.1.3 – Diminishing Returns in Research Production

Classical production theory predicts that, holding all other inputs constant, the marginal product of any single factor eventually declines. Applied to research production, scholars face fixed constraints in time, cognitive capacity, supervision bandwidth, and laboratory space, all of which restrict the rate at which additional funding converts into outputs. Lee (2021), in a nine-year panel study of 95 Korean universities using SUR, found that whilst research expenditure had significant positive effects on SCI journal articles, patents and licensing revenue, the relationship between remuneration-based productivity incentives and output followed an inverted-U shape, confirming diminishing returns at higher levels. This finding motivates the quadratic specification employed here: by estimating $Y = \beta_1 \text{CumFunding} + \beta_2 \text{CumFunding}^2$, a negative β_2 coefficient would indicate diminishing marginal returns and permit calculation of an optimal funding threshold, $F^* = -\beta_1/(2\beta_2)$, for the HK RGC system.

2.2 – Measurement of Research Productivity

How productivity is measured shapes both the magnitude of estimated funding effects and their interpretation. The choice of metric is therefore an essential design decision.

Raw publication counts remain the most common productivity measure, but they conflate quantity with quality. Hirsch (2005) introduced the h-index to address this very limitation. A scholar has an h-index of h if h of their papers has each received at least h citations. The index has been widely adopted because it combines productivity and impact into a single metric, but it also carries limitations. It favors seniority since it can only increase over time, is field-dependent, and does not consider the number of co-authors on a given paper. The rise of collaborative, multi-authored research in STEM disciplines makes full counting of that paper for all authors involved increasingly problematic. Schreiber (2008) proposed the hm-index as a practical modification that fractionalizes paper counts by the inverse of the author count. Egghe (2006) introduced the g-index as a complementary measure that rewards highly cited papers beyond the h-threshold. A scholar has a g-index of g if their top g papers together have received at least g^2 citations, giving greater weight to the outstandingly cited work than the h-index allows. Leydesdorff & Bornmann (2011) extended the fractional counting approach to journal impact factors, showing that fractional citation counts reduced between-field variance in impact factors by 81 percent. This study runs all productivity analyses in parallel using both standard and fractional metrics (h/hm , g/fg), enabling assessment of whether funding effects are consistent across attribution methods. Citation practices vary systematically across disciplines since a well-cited paper in mathematics may receive fewer raw citations than a mediocre paper in cancer research. Waltman & van Eck (2015) provide a comprehensive review of field-normalization methods, recommending normalization by the mean citation rate of all papers in the same field and year. Bornmann et al. (2020) further argue that field normalization is essential in evaluative bibliometrics whenever comparisons are drawn across disciplines.

The SCImago Journal Rank (SJR) indicator (González-Pereira et al., 2010) offers a standardized measure of journal quality across disciplines via quartile classifications (Q1 – Q4), with Q1 indicating the most prestigious journals. This study uses Q1 ratio as the primary quality-dependent variable, with Q1+Q2 ratio serving as an alternative for sensitivity analysis. Research portfolio diversity is measured using the Herfindahl-Hirschman Index (HHI), a standard concentration measure adapted for bibliometric analysis by Stirling (2007), where lower values indicate greater diversification across research topics.

2.3 – Empirical Evidence on the Funding-Productivity Relationship

The international evidence in the literature reveals a tension between quantity and quality effects, with the direction of the funding's influence depending heavily on national context, funding structure and analytical method. The most influential study in this domain remains Jacob & Lefgren (2011), who exploited a regression discontinuity in the NIH R01 grant priority scores to estimate the causal effect of funding on productivity. Their findings revealed that a grant worth

approximately US\$1.7 million produced only 0.79 additional publications over five years, a 7 percent relative increase with no statistically significant improvement in quality-adjusted productivity. Boyack & Jordan (2011), analyzing 2.57 million grant-to-article linkages over 30 years, found that PHS-funded articles attracted twice as many citations as unfunded articles from US institutions, a result that is consistent with the cumulative advantage framework.

Corsini & Pezzoni (2023), examining French Agence Nationale de la Recherche (ANR) grants with PSM, identified a 6.93 percent long-term citation boost at t+3 to t+5 years post publication, with no significant short-term effect. The absence of a short-term effect alongside a significant long-term one points towards grant funding enhancing the fundamental quality of research output rather than quickly generating superficial citation gains. Payne & Siow (2003), using an instrumental variables approach, estimated that a US\$1 million increase in federal research funding produced 10 additional publications but had negligible effects on citation quality, in line with the effect on quality finding by Jacob & Lefgren (2011).

The Asian evidence carries relevance for HK given shared institutional features and competitive pressures. Lee (2021), studying 95 Korean universities over nine years using SUR, found that research expenditure had significant positive effects on SCI journal articles, patents and licensing revenue, with cross-equation correlations confirming that these output dimensions are jointly determined. Wang et al. (2025), analyzing 1,769 Chinese communication research articles from 2008 to 2020, found that multi-funded papers significantly outperformed single-funded papers in citation impact. In the HK context specifically, Kwan et al. (2007) evaluated 205 completed Health and Health Services Research Fund projects, finding that 86.5 percent produced publications (mean of 5.4 per project). More recently, Fu et al. (2025), drawing on RGC completion rate records from 2010 to 2020, found that grants not fully completed still generated knowledge of comparable disruptiveness to completed grants, suggesting that rigid accountability frameworks may penalize exploratory research unnecessarily. Horta (2018) provided the definitive macro-level analysis of HK's research trajectory, demonstrating through revealed comparative advantage indices that the territory's scientific standing has declined relative to competitor systems. This study bridges the gap between these macro-level system analyses and micro-level studies by tracking individual scholars across an eleven-year panel.

The literature indicates that gender is also a dimension of heterogeneity in the relationship of funding and productivity. Huang et al. (2020), analyzing the scientific careers of 1.5 million researchers across 83 countries and 13 research fields, documented that female scientists face persistent structural barriers, including shorter career lengths, narrower collaboration networks and lower recognition rates, relative to male counterparts with equivalent output. Nygaard et al. (2022, p. 700) extended this analysis to bibliometric performance indicators, finding that the structural architecture of publication-counting systems structurally disadvantage female scholars by assigning them approximately 10 percent fewer publication points per paper than their male counterparts, independent of how much they publish. Together, these findings motivate the theoretical prediction that resource influxes from competitive grants may generate

disproportionately large productivity responses for female scholars who face a higher marginal return to the collaborative networks, research time, and institutional visibility that grant receipt provides.

The quality-quantity debate in the literature reveals mixed findings. Jacob & Lefgren (2011) observed a volume-only effect with no quality improvement. Corsini & Pezzoni (2023) identified a citation-quality effect with no short-term volume gains. Álvarez-Bornstein & Bordons (2021) examined the funding-quality relationship directly across seven disciplines, finding that funded research appeared in Q1 journals more frequently than unfunded research across all-fields, with collaboration acting as a mediating mechanism. Wang et al. (2019, p. 3), however, found the counterintuitive result that researchers who narrowly failed to win NIH funding produced 21 percent more “hit papers” (top 1% cited) than narrow winners over the subsequent decade, cautioning against any simple causal interpretation of positive funding-productivity associations. This finding appears to contradict the positive citation effects documented here, but the contradiction is methodological. Wang et al. (2019) use a near-miss regression discontinuity design isolating a narrow cohort of near-identical, highly ambitious scholars, and measure top 1% hit papers, a different outcome from the average field-normalized citation impact reported in this study. Furthermore, 12.6 percent of near-misses permanently exited the NIH system, meaning their treated group is a survival-selected elite rather than a representative sample of funded researchers. Ou et al. (2024), analyzing 135,612 PSM-matched papers, found that funded research simultaneously achieves higher academic citation impact but lower societal visibility than unfunded counterparts, suggesting that grant accountability optimizes for inward academic metrics rather than broader research value.

A further dimension largely absent from the quality-quantity debate is the effect of funding on research scope. Li & Bai (2025), studying NIH grant renewals, found that continued funding reduced research diversity and the probability of switching research focus, suggesting that competitive funding may narrow rather than broaden scholarly inquiry over time, quantified through the HHI of research portfolio breadth. This divergence shows that only a few studies simultaneously measure both journal placement prestige and post-publication citation impact. The possibility that funded scholars accept lower-tier journal placements to meet volume requirements whilst simultaneously producing work that ultimately garners high citation impact requires multidimensional measurement to untangle.

2.4 – Methodological Approaches in Prior Research

The challenge of estimating causal effects of research funding is one of selection. Scholars who receive grants compared to those who do not differ systematically in ways that independently predict productivity. Rosenbaum & Rubin (1983) established the theoretical foundations for PSM, demonstrating that conditioning on propensity score is sufficient to remove bias from all observed confounders.

Austin (2009) contributed practical diagnostics, establishing the standardized mean difference (SMD) threshold of 0.1 as the criterion for adequate covariate balance. TWFE models offer a complementary strategy by absorbing time-invariant unobservable heterogeneity, removing all scholar-specific characteristics that do not vary over time (Wooldridge, 2010). When the research question involves multiple dependent variables sharing common unobserved determinants, SUR enables joint estimation of the cross-equation error covariance matrix (Σ) and supports cross-equation hypothesis testing that equation-by-equation estimation cannot provide.

Lee (2021) employed a SUR system in the Korean university context to jointly estimate effects on SCI publications, patents and licensing revenue, finding significant cross-equation error correlations that validated the joint-estimation approach. Event-study specifications provide visual and statistical evidence on the dynamic trajectory of treatment effects, including the ability to test for pre-treatment parallel trends through a joint F-test on pre-treatment coefficients.

2.5 – Research Gap and Contribution

The preceding review of relevant literature identifies five specific gaps that this study directly addresses.

1. No prior study has examined the funding-productivity relationship at the individual scholar level in HK. Horta (2018) demonstrated HK's declining competitiveness at the system level but cannot identify whether this reflects funding allocation, individual scholar behavior, or structural factors unrelated to funding. The macro-level analyses cannot disaggregate the system-level trend into individual scholar trajectories, cannot test heterogeneous treatment effects across career stages or disciplines, and cannot address the selection problem that makes funded and unfunded scholars systematically different before the grant is awarded.
2. The HK specific evidence (Kwan et al., 2007; Fu et al., 2025) relies on cross-sectional or grant-level analyses that cannot track how individual scholars' productivity trajectories evolve before and after receiving funding, making it impossible to separate the funding effect from the pre-existing ability and productivity differentials that positively select scholars into grant receipt.
3. No study has simultaneously examined both quality and quantity dimensions using multiple indicators, including Q1 ratio, short-term and long-term field-normalized citations, h-index, hm-index, g-index and fg-index, within a single analytical framework applied to the same sample. This multi-indicator approach is vital for distinguishing journal prestige effects from citation impact effects, the distinction that is crucial for understanding the quality-quantity trade-off.
4. Whilst Lee (2021) identified diminishing returns in the Korean context, the optimal threshold (F^*) for HK's RGC system is unknown and cannot be extrapolated given different grant size distributions, cost structures and disciplinary differences between the two systems.

5. No prior study has systematically tested how funding effects vary by career stage, gender, funding scheme, discipline, and institution type within a single HK-specific system, despite clear theoretical predictions from both principal-agent theory and the Mertonian cumulative advantage framework that these dimensions should produce heterogeneous effects.

This study addresses all five gaps by constructing a longitudinal scholar-year panel of 22,456 STEM scholars across eleven years (2012-2022), with a five-dataset architecture (D0-D4) comprising an extensive covariate vector per scholar-year. It integrates PSM with TWFE, SUR, and event-study analysis to examine RGC funding effects on multiple productivity dimensions, while testing for diminishing returns and heterogeneous effects across scholar characteristics. Each of the methods employed address a distinct threat to causal validity, with PSM reducing selection bias from observed pre-treatment differences, TWFE eliminating time-invariant unobserved heterogeneity by differencing within each scholar over time, SUR providing the covariance structure for joint hypothesis testing across five outcome equation via estimation of the cross-equation error covariance matrix, and the event-study pre-trend F-test offering transparent, falsifiable evidence on the parallel trends assumption. Combining these methods within a single framework, applied to a HK-specific panel with scholar-level resolution, constitutes the primary methodological contribution of this study.

2.6 – Expected Findings

2.6.1 – RQ1: The Quality-Quantity Tradeoff

H1a (Publication Volume): Grant accountability pressure is expected to produce a significant positive effect on publication volume. The principal-agent framework (Stephan, 2012) predicts that when output is observable, researchers under grant reporting obligations will maximize it, a prediction which is consistent with empirical evidence from Jacob & Lefgren (2011), Lee (2021) and Wang et al. (2025).

H1b (Journal Quality / Q1 ratio): RGC funding is expected to reduce Q1 journal placement. When accountability pressure prioritizes output speed over publication prestige, the rational response is to submit to faster-turnaround, lower-tier journals, which is Tammi's (2009) impoverishment thesis. Although Álvarez-Bornstein & Bordons (2021) found that funded research associated with higher Q1 rates in cross-sectional comparisons, this association fails to distinguish pre-existing productivity differences from the causal effect of the grant. By conditioning on pre-treatment productivity through PSM-TWFE, the Q1 decline is expected to hold.

H1c (Citation Impact): A positive citation effect is expected, but its timing is an open empirical question. Boyack & Jordan (2011) and Wang et al. (2025) found strong positive citation effects, while Jacob & Lefgren (2011) and Corsini & Pezzoni (2023) found no short-term improvement. This study examines both windows separately to address the ambiguity from the literature.

2.6.2 – RQ2: Diminishing Returns to Funding Intensity

H2 (Optimal Funding Threshold): The cumulative funding-productivity relationship is expected to follow an inverted-U, consistent with classical production theory and Lee's (2021) empirical confirmation of diminishing returns across Korean universities. A significantly negative quadratic term ($\beta_2 < 0$) would allow calculation of an HK-specific optimal funding threshold (F^*), which currently doesn't exist in the literature.

2.6.3 – RQ3: Heterogeneity in Treatment Effects

H3a (Career Stage): Funding effects are expected to be strongest for early-career scholars and to decay with seniority. Bol et al. (2018) and Horta et al. (2018) identify that early grant receipt acts as a structural catalyst. Since early-career academics operate with minimal-established resources, a first grant triggers a compounding advantage that permanently separates their productivity from their unfunded peers.

H3b (GRF vs ECS scheme): ECS recipients are expected to experience a sharper quality trade-off than GRF recipients. Whilst both cohorts face identical grant accountability pressure, early-career scholars lack the experience and resources (such as research assistants and collaborators) to deal with it (Stephan, 2012).

H3c (Discipline): The quality-quantity trade-off is expected to be the sharpest in lab-intensive fields with long publication cycles and smallest in fast-turnaround fields such as Engineering and Computer Science, consistent with Álvarez-Bornstein & Bordons (2021).

H3d (Institution Type): Teaching-intensive scholars are expected to show larger productivity gains than their research-intensive scholars. Empirically, grant funding triggers the strongest output spikes among those who lack prior support (Jacob & Lefgren, 2011), given that new capital has the highest marginal return when it helps overcome basic resource bottlenecks, rather than adding to already established infrastructure (Stephan, 2012). This counters the traditional expectations of the Matthew Effect (Merton, 1968; Bol et al., 2018).

H3e (Gender): Female scholars are expected to show a larger positive productivity response to funding than male scholars. Álvarez-Bornstein & Bordons (2021) found that the funding-quality relationship was stronger for female researchers across disciplines and Huang et al. (2020) attributes persistent gender gaps in scientific output to structural disadvantages (weaker network access, heavier service load, constrained research time) that grant resources help to partially offset.

3.0 – Methodology

3.1 – Research Design

This study employs a quasi-experimental panel data design, combining PSM with TWFE, which implements a difference-in-differences (DiD) framework to address the selection bias present in non-random grant allocation. The empirical strategy further incorporates SUR to jointly model

multiple productivity dimensions and event-study analysis to trace the dynamic trajectory of funding effects over time. The panel period runs from 2012 to 2022, with treatment defined as the calendar year corresponding to the start of each funded scholar's first RGC grant. Citation data was collected between January and February 2026, ensuring complete short-term (0-2 year) citation windows for all panel papers and complete long-term (3-5 year) windows for panel papers published up until 2020.

3.2 – Data Sources

Eight data sources underpin the analysis. Individual publication records and citation data for all sample scholars were extracted from the Scopus bibliographic database via per-scholar CSV exports, resulting in 480,369 unique papers and 23.6 million citing-paper records. Grant-level records were extracted from RGC's CERG1 database, covering all STEM-related GRF and ECS grants from academic years 2012/13 through to 2022/23, as well as a lookback window spanning (from 2006/07 to 2011/12) to identify scholars holding grants prior to the study period. Annual SJR quality data (1999-2023) were matched to papers via ISSN and publication year, with each paper assigned the best (lowest-numbered) quartile across their All Science Journal Classification (ASJC) categories.

Institutional data was proxied using a two-year lagged Academic Ranking of World Universities (ARWU) ranking variable ($ARWU_{i,t-2}$). The two-year lag helps prevent reverse causality since a large grant awarded in year t could directly boost an institution's ARWU ranking in the same or following academic year, and using the concurrent ranking would introduce endogeneity into the institutional control. A two-year lag provides sufficient temporal separation to treat the institutional quality proxy as predetermined with respect to the current year's research output. ARWU was selected over alternative rankings such as Times Higher Education and QS due to its reliance on publication and citation data rather than reputational surveys, making it more consistent with the bibliometric framework of this study. All nominal funding amounts were deflated to real 2012/13 HK\$ terms using the Hong Kong Consumer Price Index (CPI) from the Census and Statistics Department. Scholar gender was determined via the NamSor API v2 (an algorithm that can analyze first, last or full name data to provide cultural origin, gender and ethnicity), with limitations of name-based gender classification for East-Asian names discussed in Section 3.7. Linking RGC grant recipients to their Scopus author profiles required a multi-stage disambiguation pipeline, given the high prevalence of shared surnames in the HK academic population. Institutional staff directories for all eight UGC-funded universities were scraped from Scopus and university websites to build a master roster, and grant recipients were matched to this roster using a six-strategy fuzzy name matching algorithm incorporating institution-constrained similarity scoring, with borderline cases routed to manual review and validated against publication records. A university name mapping table linked the diverse affiliation strings appearing in Scopus records to the eight UGC-funded institutions being analyzed in this study. Open-access (OA) status for each publication was retrieved via the Unpaywall API, classifying publications as Gold (freely

available to everyone) or Green (paywalled but manuscript exists in an open repository e.g. arXiv) for the supplementary open-access outcome variable.

3.3 – Sample Construction

The sample contains 22,456 STEM scholars associated with HK’s eight UGC-funded universities. The treatment group consists of 2,545 scholars who received at least one RGC grant (GRF or ECS) during the study period. The control group consists of 19,911 scholars comprising UGC-affiliated STEM scholars with strictly zero RGC funding receipt between 2012 and 2022, classified as STEM based on their primary ASJC subject code in Scopus. This group may nonetheless hold industry grants, teaching development grants, or overseas funding. The estimated effect sizes should therefore be interpreted as the marginal impact of an RGC grant compared to diverse alternative funding sources, rather than a strict comparison against scholars with zero funding. The inclusion criterion requires each scholar to have at least one journal publication indexed in Scopus with a non-null Digital Object Identifier (DOI) and a valid International Standard Serial Number (ISSN) or eISSN, focusing the analysis on peer-reviewed journal output, the standard unit of research productivity measurement in bibliometric literature.

3.4 – Variables

Variables are organized into four categories: dependent variables, independent variables (treatment), control variables and moderating variables for the heterogeneity analysis. Five primary outcomes are estimated within the main SUR and TWFE specifications, with fractional variants of publication volume and Q1 ratios run in parallel as robustness complements.

Table 1: Dependent Variables

Variable	Column Name	Description	Source	RQ
Publication volume	pub_count_it	Annual count of Scopus-indexed articles attributed to scholar i in year t	Scopus	RQ1
Publication volume (fractional)	frac_pub_count_it	Annual publication count weighted by reciprocal of co-author count; run in parallel as a robustness check for hyper-authorship inflation	Scopus	RQ1
Q1 journal ratio	q1_ratio_it	Proportion of annual output in SJR Q1-ranked journals, with quartile assigned at publication year	Scopus + SJR	RQ1
Q1 ratio (fractional)	frac_q1_ratio_it	Fractional-weighted Q1 journal ratio; run in parallel as a robustness check	Scopus + SJR	RQ1

Variable	Column Name	Description	Source	RQ
Average SJR score	avg_sjr_score_it	Mean SCImago Journal Rank score across all publications in year t	Scopus + SJR	RQ1
h-index	h_index_it	Annual h-index computed from cumulative citations to year t	Scopus	RQ1
Short-term citations	fn_short_term_cit_it	Field-normalized (FWCI-style) citation count accumulated within two years of publication, normalized by ASJC subject area	Scopus + ASJC	RQ1
Long-term citations	long_term_cit_it	Cumulative citations in the three-to-five-year post-publication window; restricted to publications ≤ 2020 to ensure complete windows	Scopus	RQ1

Secondary/exploratory outcomes (estimated via standalone TWFE; not part of the 5-equation SUR system; reported as supplementary findings):

Variable	Column Name	Description	Source
Research diversity	discipline_hhi_it	HHI across ASJC subject codes; higher values indicate greater specialization	Scopus + ASJC
International co-authorship	intl_collab_prop_it	Proportion of annual publications with at least one non-Hong Kong affiliated co-author	Scopus
Open-access rate	oa_prop_it	Proportion of annual publications with Gold or Green OA status, retrieved via Unpaywall API	Unpaywall

Supplementary bibliometric indices (robustness / multi-indicator consistency):

Variable	Column Name	Description
hm-index	hm_index_it	Fractional h-index variant accounting for co-authorship (Schreiber, 2008)

Variable	Column Name	Description
g-index	g_index_it	g-index rewarding highly cited papers beyond the h-threshold (Egghe, 2006)
fg-index	fg_index_it	Fractional g-index combining Egghe's g with fractional authorship counting
Q1+Q2 ratio	q12_ratio_it	Proportion of annual output in SJR Q1 or Q2 journals; tests whether effects extend to the broader top-half threshold

Table 2: Independent Variables (Treatment)

Variable	Column Name	Description	Source	Used in
Grant receipt	post_treat_it	Binary indicator = 1 for year $t \geq$ first RGC ECS or GRF award; restricted to ECS and GRF competitive schemes	RGC project database	RQ1 main, RQ3
Cumulative funding	cum_funding_M	Cumulative real HKD received from RGC ECS/GRF grants through year t (millions); CPI-deflated to 2012/13 base year	RGC database + CPI	RQ2 main
Cumulative funding ²	cum_funding_M_sq	Quadratic term (cum_funding_M ²); negative β_2 confirms diminishing returns; optimal threshold $F^* = -\beta_1/(2\beta_2)$ via Delta Method (Wooldridge, 2010)	Derived	RQ2 main
Annual funding flow	annual_funding_M	Annual real HKD grant amount in year t (millions); alternative to cumulative specification in Specification E robustness check	RGC database + CPI	Spec E robustness

Table 3: Control Variables

Time-varying regression controls (included in all primary TWFE specifications)

Variable	Column Name	Description	Rationale
Career age	career_age_it	Years elapsed since scholar i 's first Scopus-indexed publication as of year t	Accounts for natural productivity lifecycle independent of grant receipt
Institutional prestige	arwu_score_it	ARWU continuous institutional score, lagged two years (t-2) to prevent reverse causality; unranked institutions imputed with mean ranked HK score	Controls for time-varying institutional resource endowments
ARWU ranking status	arwu_ranked_it	Binary = 1 if institution appears in ARWU rankings in year t (lag-adjusted); flags unranked institutions	Distinguishes ranked vs. unranked institutional environments
Co-author count	avg_author_count_it	Mean co-author count per publication in year t ; trimmed at 99th percentile	Controls for collaborative productivity inflation independent of funding
Affiliation inference	affil_inferred_i	Binary = 1 if HK affiliation was inferred from Scopus metadata rather than directly observed; time-invariant	Controls for differential data quality across scholars

Fixed Effects Structure:

Effect	Specification	Rationale
Scholar fixed effects	One dummy per <i>scopus_id</i>	Absorbs all time-invariant unobserved heterogeneity: ability, discipline, baseline productivity, institutional culture
Year fixed effects	One dummy per calendar year (2012–2022)	Absorbs all aggregate annual shocks: RGC policy changes, global publication trends, COVID-19 disruptions

PSM matching covariates (used in PSM only, not in TWFE regression):

PSM covariate	Column Name	Notes
Career age at baseline	career_age_2012	Career age in 2012 (study baseline), not at first grant
Baseline fractional publications	baseline_frac_pub_count	Pre-treatment fractional publication count
Baseline total citations	baseline_total_cit	Pre-treatment cumulative raw citations
Baseline h-index	baseline_h_index	Pre-treatment cumulative h-index
Baseline field-normalised citations	baseline_field_norm_cit	Pre-treatment FWCI-style citation score
Baseline average SJR score	baseline_avg_sjr_score	Pre-treatment mean journal prestige
Pre-study grant count	pre_study_num_grants	Number of RGC grants received before 2012
Pre-study cumulative funding	pre_study_total_funding_real	Real HKD RGC funding received before 2012 (in 2012/13 HK\$)
Gender *	gender_i dummies	Male / Female / Unknown indicator dummies.

PSM covariate	Column Name	Notes
Institution *	institution_2012 dummies	HK institution affiliation dummies (baseline year)
Discipline *	Stratification variable	Matching performed within ASJC discipline groups, not as a covariate

*Note: * = Additional conditioning variable (not counted in the 8 core covariates)*

Table 4: Moderating Variables (RQ3 – Heterogeneity Analysis)

Variable	Column Name	Description	Interaction Term
Career age at first grant	career_age_at_first_grant	Career age at year of first grant receipt; centered at funded-sample mean; time-invariant	post_treat_it × career_age_centered
Grant scheme	ecs_dummy	Binary = 1 if first grant was ECS; = 0 if GRF; funded subsample only; time-invariant	post_treat_it × ecs_dummy
Discipline	discipline_i	ASJC-based four-category discipline group; time-invariant, absorbed by entity FE; tested via subgroup regressions	Subgroup regressions
Institution type	institution_it	Research-intensive (HKU, HKUST, CUHK) vs. teaching-intensive (PolyU, CityU, HKBU, LU, EdUHK); tested via subgroup regressions	Subgroup regressions
Gender	female_i	Binary = 1 if inferred female via NamSor API v2; 40.7% unclassified; time-invariant	post_treat_it × female_i

3.5 – Propensity Score Matching

Constructing a credible comparison group requires addressing a fundamental baseline inequality. Pre-treatment differences between the treatment (funded) and control (unfunded) groups were substantial. Baseline fractional publication counts averaged 42.37 for the funded group versus just 6.55 for the unfunded group, whilst baseline h-indices averaged 9.08 versus 2.04 respectively,

confirming positive selection bias. Comparing these groups directly would measure pre-existing academic dominance rather than the causal effect of the grant.

This study employs PSM to remove the bias. PSM addresses this fundamental inequality by constructing “statistical twins.” It estimates each scholar’s probability of receiving RGC-funding through a regularized logistic regression (L2 penalty, $C = 1.0$), estimated within each ASJC discipline stratum, conditioning on 8 pre-2012 covariates spanning across three domains. Bibliometric output history was captured by five measures: fractional publication count, field-normalized citation score, h-index, total citation output and average SJR score. Career trajectory was proxied by career age in 2012, defined as years elapsed since the scholar’s earliest Scopus-indexed publication, a measure of research seniority derivable from publication records given the absence of academic rank and year PhD earned data. Funding history was represented by the pre-study RGC grant count covering academic years 2006/07 through to 2011/12 and cumulative real HK\$ funding received in that lookback window (inflated to a 2012/13 base using CPI data). The model additionally included binary gender dummies and a full set of institutional affiliation dummies to condition on systematic differences in funding probability attributable to gender and institutional environment, ensuring matched pairs are comparable on both dimensions. Scholars were assigned a broad discipline group (one of four STEM categories) derived from the most frequent ASJC code across pre-2012 publications which was enforced as an exact-match stratum, preventing cross-field comparisons.

Matching then proceeded by 1:1 nearest-neighbor matching within a 0.10 standard deviation caliper on the estimated propensity score. Post-matching covariate balance is evaluated against Austin’s (2009) threshold of $|SMD| < 0.10$. For the subsequent fixed effects estimations, the primary analyses rely exclusively on the full matched PSM panel. The balanced active panel approach (only including scholars who published in every single panel year) is avoided for analyses as it constitutes a collider variable (Elwert & Winship, 2014), given that continuous publication is a direct downstream consequence of receiving a grant. Restricting the sample to always-active scholars conditions on a post-treatment outcome, introducing endogenous selection bias. This would artificially inflate the estimated treatment effects by selecting only the most successful post-grant researchers. The balanced active panel is therefore reserved purely as a supplementary sensitivity check.

3.6 – Empirical Strategy

To clearly situate the technical procedure, PSM first identifies each funded scholar’s statistical twin from the unfunded pool, matching on every observable characteristic before the grant arrived. TWFE then removes everything about that scholar that does not change over time, leaving only the variation caused by the grant itself. SUR treats the four productivity outcomes as a connected ecosystem rather than isolated metrics. The event study then maps the year-by-year trajectory before and after grant receipt, providing a visual test of whether the two groups were already diverging before treatment.

The analysis employs three complementary model families with the intention of isolating the true treatment effect of the RGC grant. While PSM effectively neutralizes selection bias on observable metrics, it cannot account for unobservable, time-invariant traits such as a scholar's innate intellect, intrinsic research drive, or the pre-existing research culture of their institution. To isolate the causal effect of the grant from these unobservable constants, the primary estimating equation employs a TWFE model. Rather than comparing Scholar A to Scholar B, TWFE isolates the within-scholar variation. This “within-person differencing” mathematically subtracts the scholar's stable, unmeasurable traits out of the equation, leaving only the impact of the grant itself, whilst year fixed effects simultaneously control for the global temporal shocks. However, before applying TWFE, a Hausman-test (Hausman, 1978) is conducted to evaluate whether these unobserved individual specific effects (α_i) are correlated with the regressors, thereby determining whether random-effects or fixed-effects estimation is statistically required. This baseline binary treatment (Model 1) is specified as $Y = \alpha_i + \lambda_t + \beta_1 \text{PostTreat}_{it} + \gamma \text{Controls}_{it} + \varepsilon_{it}$, where α_i denotes scholar fixed effects that absorb unobserved, time-invariant traits of individual scholars, while λ_t denotes year fixed effects and controls for common annual shocks across the dataset. This model is estimated independently for five dependent variables: publication volume (Y_1), Q1 journal ratio (Y_2), short-term field-normalized citations (Y_3), career impact via the h-index (Y_4), and long-term field normalized citations (Y_5).

Since research productivity is multidimensional, scholars experiencing a highly productive year will likely see simultaneous increases across several metrics. Estimating these outcomes via isolated Ordinary Least Squares (OLS) regressions ignores this inherent interconnectivity. SUR is therefore used to mathematically link the outcome equations. Due to all equations in this SUR system sharing identical regressors (the same PostTreat_{it} , Controls_{it} , α_i , and λ_t), Zellner's (1962) theorem dictates that equation-by-equation OLS and SUR will produce numerically identical coefficient estimates. However, because long-term citations require restricting the panel strictly to pre-2020 observations to ensure newer papers have an adequate time window to accumulate citations, including this metric in the SUR system would force the algorithm to drop all 2021 and 2022 observations across the other outcomes to maintain matrix symmetry. Consequently, long-term citations (Y_5) are estimated using the same Model 1 specification on a restricted sample (publication year ≤ 2020), while a 4-equation SUR system is deployed for the remaining outcomes. SUR's methodological contribution in this context is therefore the estimation of the full cross-equation errors covariance matrix (Σ). Defining the variables as outlined above, the system is specified as follows:

$$\begin{aligned} Y_{1,it} &= \alpha_i + \lambda_t + \beta_1 \text{PostTreat}_{it} + \gamma \text{Controls}_{it} + \varepsilon_{1,it} \\ Y_{2,it} &= \alpha_i + \lambda_t + \beta_2 \text{PostTreat}_{it} + \gamma \text{Controls}_{it} + \varepsilon_{2,it} \\ Y_{3,it} &= \alpha_i + \lambda_t + \beta_3 \text{PostTreat}_{it} + \gamma \text{Controls}_{it} + \varepsilon_{3,it} \\ Y_{4,it} &= \alpha_i + \lambda_t + \beta_4 \text{PostTreat}_{it} + \gamma \text{Controls}_{it} + \varepsilon_{4,it} \end{aligned}$$

The off-diagonal elements of Σ capture the degree to which idiosyncratic productivity shocks in one year are simultaneously reflected across multiple dimensions. A Breusch-Pagan (Breusch & Pagan, 1980) test of independence is utilized to evaluate the null hypothesis that all off-diagonal Σ elements are zero. Rejecting this null formally justifies joint estimation and provides the covariance structure required for subsequent cross-equation Wald hypothesis testing.

To validate the foundational assumption of these models, the analysis must prove that funded and unfunded scholars would have followed the exact same trajectory in the absence of the grant. To execute this, an event-study specification replaces the binary treatment variable with a series of relative-time dummies from $k = -5$ to $k = +8$, tracking the scholar's productivity trajectory year by year: $Y = \alpha_i + \lambda_t + \sum_{k=-5, k \neq -1}^8 \beta_k D_{it}^k + \gamma \text{Controls}_{it} + \varepsilon_{it}$.

The period $k = -1$ (the year immediately preceding grant receipt) is omitted as it serves as the baseline reference year, against which all other coefficients are interpreted. In this specification, the k timeline systematically dismantles the treatment effect into specific temporal windows. The negative k -values represent the priors, which are the year strictly preceding the grant receipt. The coefficients on these negative terms must remain statistically indistinguishable from zero and ideally, the pre-trend should prove that the funded and unfunded cohorts were running the exact same race prior to the grant. A joint F-test on these pre-treatment coefficients ($k=-4$ through to $k=-2$) will help provide the formal mathematical proof of parallel trends. The timeline then crosses $k=0$, the exact calendar year the grant is awarded. Finally, the positive k -values map the multi-year impact of the RGC grant across the post grant receipt horizon.

To investigate the optimal threshold of funding intensity (RQ2), Model 2 replaces the binary post treatment indicator with the quadratic continuous treatment to test for diminishing marginal returns: $Y = \alpha_i + \lambda_t + \beta_1 \text{CumFunding}_{it} + \beta_2 \text{CumFunding}_{it}^2 + \gamma \text{Controls}_{it} + \varepsilon_{it}$, where a negative β_2 coefficient empirically confirms diminishing returns. Identifying the precise peak of this inverted U-curve requires calculating the vertex of the function by setting the first derivative to zero, which yields the optimal funding threshold: $F^* = -\beta_1/(2\beta_2)$. Given that this theoretical optimal threshold is derived from a non-linear combination of two estimated coefficients, calculating its standard error requires the Delta Method (Wooldridge, 2010). This first-order Taylor series approximation provides the necessary confidence interval to determine if the identified peak represents a statistically reliable policy threshold rather than statistical noise.

Finally, to safeguard causal validity, six alternative specifications, including alternative sample restrictions, inverse probability of treatment weighting (IPTW), annual funding flow models, and an interaction-weighted event-study (IWES) estimator (Sun & Abraham, 2021) deployed specifically as a diagnostic for heterogeneous treatment timing bias, are specified to confirm the direction and stability of the primary models. Additionally, a leave-one-institution-out jackknife analysis is conducted to guarantee that the estimated treatment effects are structural to the HK system rather than artifacts driven by any single outlier university. All primary analyses are additionally replicated using fractional and alternative impact variants (hm-index, fg-index, g-

index) and a Q1+Q2 ratio as a broader quality threshold, confirming that results are not sensitive to the choice of bibliometric measure.

To address RQ3, the baseline TWFE specification is extended with a series of moderation analyses examining how the productivity response to grant receipt varies across scholar characteristics. Career stage heterogeneity is tested by interacting the treatment indicator with the scholar's career age at first grant, centered at the funded sample mean to determine whether the treatment effect varies with career seniority. Scheme heterogeneity between the ECS and GRF is tested via a binary interaction term on the ECS indicator. Disciplinary and institutional heterogeneity are examined through subgroup regressions across the four ASJC discipline clusters and a research-intensive versus teaching-intensive institution split, with subgroup regressions preferred over interaction terms because the institution type is time-invariant and absorbed by entity fixed effects. Finally, gender heterogeneity is tested by interacting the treatment indicator with a binary female indicator. All moderation models are estimated on the funded subsample with scholar-level clustered standard errors.

3.7 – Limitations

While the integrated empirical strategy addresses the primary threats to causal inference, several structural limitations qualify the findings. The constraints fall into three categories: data measurement boundaries, methodological vulnerabilities, and limits on generalizability.

3.7.1 – Data Measurement Vulnerabilities

Scopus Coverage Bias: The sample is constructed entirely from Scopus-indexed publications. Relying purely on Scopus may disadvantage scholars in applied engineering and practitioner-facing fields as they are underrepresented on Scopus relative to fundamental sciences. Scholars publishing through non-Scopus-indexed modes will appear less productive in the dataset than they are.

The “Contaminated” Control Group: The unfunded cohort is defined strictly by the absence of RGC funding. However, these scholars may hold substantial alternative funding, such as industry contracts, internal university seed funds, or grants from China's National Natural Science Foundation (NSFC). Consequently, the estimated coefficients represent the marginal impact of an RGC grant relative to a baseline of alternative funding, rather than a comparison against a strictly unfunded, zero-resource benchmark.

API Pagination Caps on Hyper-Cited Papers: While the Scopus API reliably extracted complete citation records for most of the dataset, the extraction algorithm encountered pagination limits when processing hyper-cited papers (specifically, two anomaly papers exceeding 20,000 citations). For these specific hyper-cited outliers, extraction capped at approximately 80-88 percent coverage. Furthermore, standard index lagging, where the instantaneous Scopus citation counter marginally outpaces physically extractable document metadata, generated single digit “ghost citation” discrepancies across the broader panel. Given that the API constraints affect only 316 of the

480,369 unique publications (0.066 percent), the structural integrity of the citation metrics remains intact. However, this limitation truncates the extreme right-tail of the citation distribution, marginally understating the absolute maximum impact of the system's most dominant publications.

Staggered Citation Scraping Noise: Citation counts reflect a rolling scrape window spanning from January to February 2026, rather than a uniform, instantaneous snapshot. Given that highly cited papers may have been scraped earlier in the extraction window, they technically had more time to accumulate new citation on Scopus before the dataset was finalized. While this non-random measurement noise is narrow given the five-week window, it structurally correlates with citation volume and may marginally inflate counts for the most productive scholars.

Gender Inference Limitations: Scholar gender was inferred from their names via the NamSor API v2. To prevent misclassification bias, a strict 0.7 confidence threshold was applied. As a result, 40.7 percent of the gender sample is categorized as “Unknown” and is excluded from the relevant interaction models (retaining $N=20,350$). Given this non-random nature of the missing data, it reduces the statistical precision of the gender heterogeneity models. Hence, gender findings must be interpreted as suggestive directional trends, and not definitive conclusions. This missingness rate also reflects a broader methodological limitation of name-based gender inference tools, which were developed primarily on Western-name corpora and perform substantially less reliably on Chinese and East Asian names. Given that Chinese-heritage scholars constitute the majority of HK's STEM scholars, this is a known limitation of current tools.

3.7.2 – Methodological and Model Constraints

Unobservable Confounders & PSM Limits: PSM balances cohorts strictly on observed variables. While baseline journal prestige (SJR score) was successfully incorporated into the primary matching, time-varying unobservable factors such as a sudden upgrade in laboratory infrastructure, may remain systematically different between groups post-matching. As these unobservable factors vary over time, the scholar-fixed effects cannot fully absorb their confounding influence.

Residual Imbalance and TWFE Heterogeneity Bias: Applying Austin's (2009) threshold, only three covariates failed to meet the $SMD < 0.10$ threshold post matching (maximum residual $SMD = 0.2973$). These remaining imbalances are entirely concentrated in grant history, reflecting the structural reality of the Matthew Effect in academia, where funded scholars possess richer grant histories. While the TWFE model mathematically absorbs these baseline level differences, it cannot eliminate heterogeneous treatment effect bias. If scholars with massive pre-treatment funding histories convert new RGC grants into publications at a systematically different rate than historically unfunded scholars, standard TWFE with staggered adoption will produce biased estimates. An IWES estimator (Sun & Abraham, 2021) is therefore deployed as a diagnostic test for this differential timing bias.

Collider Bias in the Balanced Panel: The balanced active panel specification requires scholars to have published continuously in every panel year. Given that post-treatment publication activity

is a direct downstream consequence of receiving a grant, conditioning the sample on this variable selects on the outcome, introducing endogenous collider bias. This mathematically guarantees an upward-biased sub-group average, hence why the balanced active panel is reported as a sensitivity check.

Incomplete Citation Window: To avoid the bias of an incomplete time window for recently published papers, long-term citation models restrict the panel to scholar-years up to 2020 to guarantee the standard 3-to-5-year citation accumulation window. While necessary to prevent bias, dropping 2021 to 2022 data systemically underestimates the long-term citation impact of later grant cohorts.

The 8-Institution Cluster Problem: Standard errors are clustered at the scholar level. However, clustering at the institutional level is mathematically unreliable in this context because HK possesses only eight-UGC funded universities. As demonstrated by Cameron & Miller (2015), standard errors clustered across too few groups are systemically downward-biased, generating overconfidence in statistical significance. The leave-one-institution-out jackknife analysis is deployed to partially mitigate this vulnerability, but the small number of institutions fundamentally limits the precision of the estimates.

Editorial Pipeline Lags: The annual panel structure cannot perfectly model the 12-to-24-month peer review and editorial lag inherent in STEM publishing. Consequently, the temporal precision of the event study is inherently blurred and the immediate productivity effects observed at $k=0$ and $k=1$ partially capture the printing of pre-funded research rather than new, grant-funded research.

3.7.3 – The Boundaries of Interpretation

The Average Treatment Effect on the Treated (ATT): Achieving rigorous covariate balance required dropping 38.7 percent of funded scholars who fell outside the common support region. These dropped individuals represent the extreme upper tail of academic productivity such as the superstar scholars who were too dominant to have any unfunded statistical counterpart. Since these elite researchers typically rely on existing institutional or international backing, their output is largely insulated from RGC funding decisions. Therefore, the ATT estimated here applies specifically to the broad middle tier of scholars within the common support region and captures the demographic whose publication strategies are most responsive to grant incentives. While the results do not extrapolate the top percentile of the system, they offer a highly precise, policy-relevant measure of how grant receipt shapes the scholars for whom the funding actually matters.

Structural Instability in the Post-2019 Period: The panel encompasses the 2019-2020 social unrest in HK as well as the COVID-19 pandemic from 2020-2022, all of which materially altered HK's institutional environment, academic labor market and international collaboration landscape. Year fixed effects absorb the aggregate-level impact of these shocks but cannot test whether grant-productivity treatment effect is structurally stable across the pre- and post-2019 cohorts. Scholars who received their first RGC grant after 2019 may have operated under qualitatively different

institutional constraints than those who received grants earlier in the panel. Estimates from the later grant cohorts should therefore be interpreted with additional caution, and explicit sub-period heterogeneity (splitting the panel at 2019) is a noteworthy topic for future work.

Underpowered Sub-Samples: The Early Career Scheme (ECS) cohort comprises only 285 recipients (11.2 percent of the funded sample). Models running on this subset will be statistically underpowered, yielding wide confidence intervals. Scheme-level comparisons between GRF and ECS must therefore be interpreted as exploratory hypotheses rather than confirmed empirical facts.

Statistical Fragility of the Threshold Estimator: The optimal funding threshold (F^*) is derived mathematically from the ratio of two estimated coefficients (the Delta method). Since this is a non-linear combination, the resulting standard errors can inflate substantially, particularly for outcome variables with high idiosyncratic variance, such as publication volume. Any subsequent policy recommendations regarding absolute financial caps must therefore be contingent on the confidence intervals remaining strictly positive and structurally narrow.

Geographic External Validity: The causal dynamics identified in this study are structurally bound to HK's highly compact, internationally oriented, English-medium, and UGC-funded academic dynamic. Extrapolating the F^* diminishing returns threshold or the specific quality-quantity behavioral trade-offs to vastly different incentive environments such as Thailand or the United States, requires explicit geographic replication.

4.0 – Overview of Findings

The findings reported across Sections 4 and 5 converge on a consistent and counterintuitive verdict. HK's competitive grant system produces the opposite of its intended quality outcomes for the scholars most responsive to its incentives. Across all primary outcomes, heterogeneity subgroups and robustness specifications, grant receipt increases research quantity while degrading journal prestige.

4.1 – Sample and Matching Overview

The study analyzed a comprehensive, decade-long panel of STEM scholars across HK's publicly funded universities. Initial baseline comparisons revealed strong positive selection bias, with scholars who received RGC grants being significantly more productive than their unfunded peers long before the funding was awarded. To ensure a fair assessment of the grant's true impact, statistical matching techniques were successfully implemented to construct a balanced comparison between funded and unfunded scholars.

4.2 – The Quality-Quantity Trade-Off

RGC grant receipt generates a consistent quality-quantity trade-off since funded scholars publish more frequently and accumulate stronger short-term citation impact but channel this output toward lower-prestige publications. The theoretical implications of this citation-prestige paradox are explored in Section 6.

4.3 – Diminishing Returns to Funding Intensity

The funding-productivity relationship follows an inverted-U curve, with marginal returns to additional grants eventually turning negative beyond an optimal threshold (F^*). For the typical RGC-funded scholar, cumulative career funding sits well below these thresholds, meaning most scholars remain in the increasing-returns phase where every new grant still generates more output. Only serial grant accumulators, who may be veteran researchers stacking multiple GRF grants over a decade, approach or exceed the ceiling.

4.4 – How Effects Vary Across Scholars

4.4.1 – Career Stage

The volume boost from grant receipt is concentrated in early-career scholars and dissipates with seniority, but the shift toward lower-prestige journals is uniform across all career stages. The structural pressure to meet RGC output requirements alters journal submission behavior universally, regardless of academic seniority.

4.4.2 – Gender

The gender analysis is limited by 40.7 percent of the sample being unclassifiable due to Chinese-heritage naming conventions. Within the classifiable subsample, both male and female scholars experience the same quality-quantity trade-off, targeting lower-prestige, faster turnaround journals under identical grant accountability pressure.

4.4.3 – Discipline

Discipline stratified regressions reveal that the behavioral response to RGC funding depends heavily on the natural speed of the research field. Engineering and CS scholars convert funding into publication volume with minimal quality sacrifice, while lab-intensive Life and Physical Science scholars face the sharpest prestige drop due to longer experiment lead times forcing faster journal choices.

4.4.4 – Funding Scheme: ECS vs GRF

ECS recipients experience a sharper quality trade-off than GRF recipients, suggesting the scheme's accountability architecture is misaligned with the developmental stage of its intended recipients, placing an output pressure on scholars at the career stage when research time is scarcest.

4.4.5 – Institution Type

Teaching-intensive scholars show the largest publication gains and smallest quality sacrifice, while research-intensive scholars show the opposite, suggesting the Matthew Effect directs capital to precisely where marginal returns are lowest.

4.4.6 – Summary

Across every subgroup tested, the Q1 ratio declines. Career stage, discipline, institution type, and gender all shape publication volume, but none escape the quality trade-off. The accountability architecture operates structurally, not selectively.

4.5 – Broader Behavioral Patterns

Beyond the primary outcomes, grant receipt broadens research portfolios and increases international co-authorship. The topic diversification may not be independent of the quality trade-off since spreading across more fields can thin domain expertise, pushing scholars toward lower-quality journals when outside of their core specialization. The causal pathway cannot be confirmed here but the co-occurrence warrants future work.

5.0 – Empirical Results

5.1 – Descriptive Statistics and Covariate Balance

The full panel dataset is made up of 22,456 STEM scholars affiliated with HK's eight UGC-funded universities, generating 247,016 scholar-year observations spanning 2012 to 2022. Of these scholars, 2,545 received at least one RGC grant (GRF or ECS) during the panel period, whilst the remaining 19,911 held strictly zero RGC funding. Pre-treatment differences between the funded and unfunded groups were substantial and systematic. Baseline fractional counts averaged 42.37 for the funded group versus 6.55 for the unfunded group, baseline h-indices averaged 9.08 versus 2.04, and baseline total citations averaged 2,346.57 versus 303.06, confirming strong positive selection into grant receipt on all observable productivity dimensions.

Of the full pool, 8,422 scholars met the completeness criteria for PSM eligibility (2,122 funded and 6,300 unfunded). Following 1:1 nearest neighbor PSM within each broad ASJC discipline stratum using 8 pre-2012 covariates, the matching produced 1,560 funded-control pairs (3,120 scholars with 34,320 scholar-year observations across 2012-2022). Post-matching, the mean absolute standardized mean difference (SMD) fell from 0.3301 to 0.0519, indicating substantially improved covariate balance. Applying Austin's (2009) threshold of $|SMD| < 0.10$, three covariates remain above this criterion after matching: pre-study grant count (0.160), had pre-study grants (0.297) and cumulative pre-study funding (0.142). These residual imbalances are concentrated entirely in pre-study grant history variables and reflect the inherent structural reality of the Matthew Effect in academia where funded scholars are disproportionately likely to possess prior grant experience. TWFE differencing absorbs these stable baseline level differences within each scholar, and the IWES estimator is deployed as a diagnostic check for heterogeneous timing bias.

Post-matching descriptive statistics confirm close alignment between matched funded and unfunded groups. Mean career age in 2012 is 11.94 (funded) versus 11.72 (unfunded). Baseline fractional publication counts average 10.54 versus 11.60, baseline Q1 proportions average 0.708 versus 0.717, and baseline average SJR scores average 1.915 versus 2.065. These small residual differences in publication-based covariates are well within the range absorbed by scholar-level

fixed effects. The average post-treatment observation window for funded scholars is approximately 6 to 7 years, sufficient to detect both short-term publication volume responses and the short-term citation accumulation windows (0-2 years post-publication) that define the primary citation quality metric.

Table 5: Descriptive Statistics – Full Sample and PSM Matched Groups [Mean (SD)]

Variable	Full Funded	Full Unfunded	PSM Funded	PSM Unfunded
Career age in 2012	10.66 (10.10)	0.33 (11.51)	11.94 (8.36)	11.72 (9.85)
Baseline pub count (fractional)	42.37 (67.82)	6.55 (27.04)	37.16 (47.44)	40.62 (77.74)
Baseline Q1 proportion	0.700 (0.240)	0.588 (0.360)	0.708 (0.244)	0.717 (0.279)
Baseline h-index	9.08 (9.92)	2.04 (4.81)	9.04 (8.43)	9.35 (10.50)
Baseline total citations	2,347 (4,766)	303 (1,795)	2,009 (3,431)	2,200 (5,319)
Baseline avg SJR score	2.04 (1.67)	1.55 (1.22)	1.92 (1.27)	2.07 (1.84)
Pre-study grant count	0.579 (1.306)	0.017 (0.262)	0.179 (0.650)	0.078 (0.624)
Had pre-study grants	21.8%	0.6%	9.8%	2.6%

Note: Full funded N = 2,545; full unfunded N = 19,911; PSM matched N = 1,560 per group. SD in parentheses. The low mean career age for the full unfunded group reflects inclusion of scholars who first published in or near 2012, for which the PSM-matched subsample shows close balance.

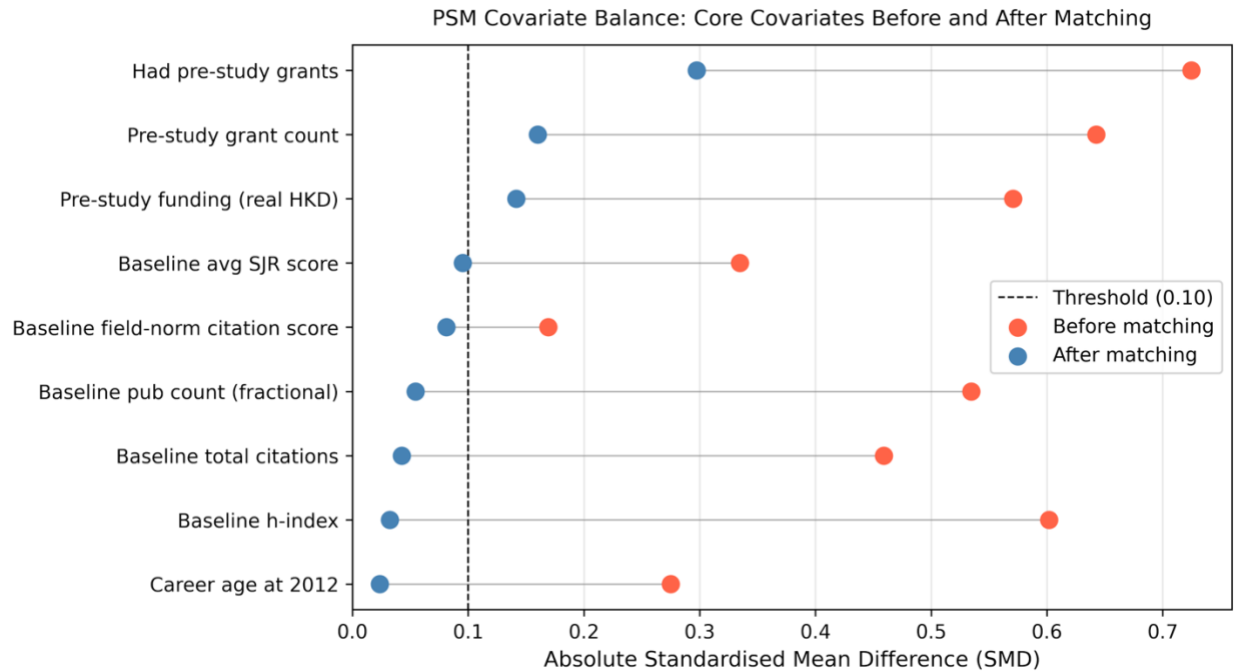


Figure 1: PSM Covariate Balance Before and After Matching

Note: Dashed line indicates the $|SMD| < 0.10$ threshold (Austin, 2009). Three covariates related to pre-study grant history exceed the threshold post-matching, consistent with the structural Matthew Effect in academic funding.

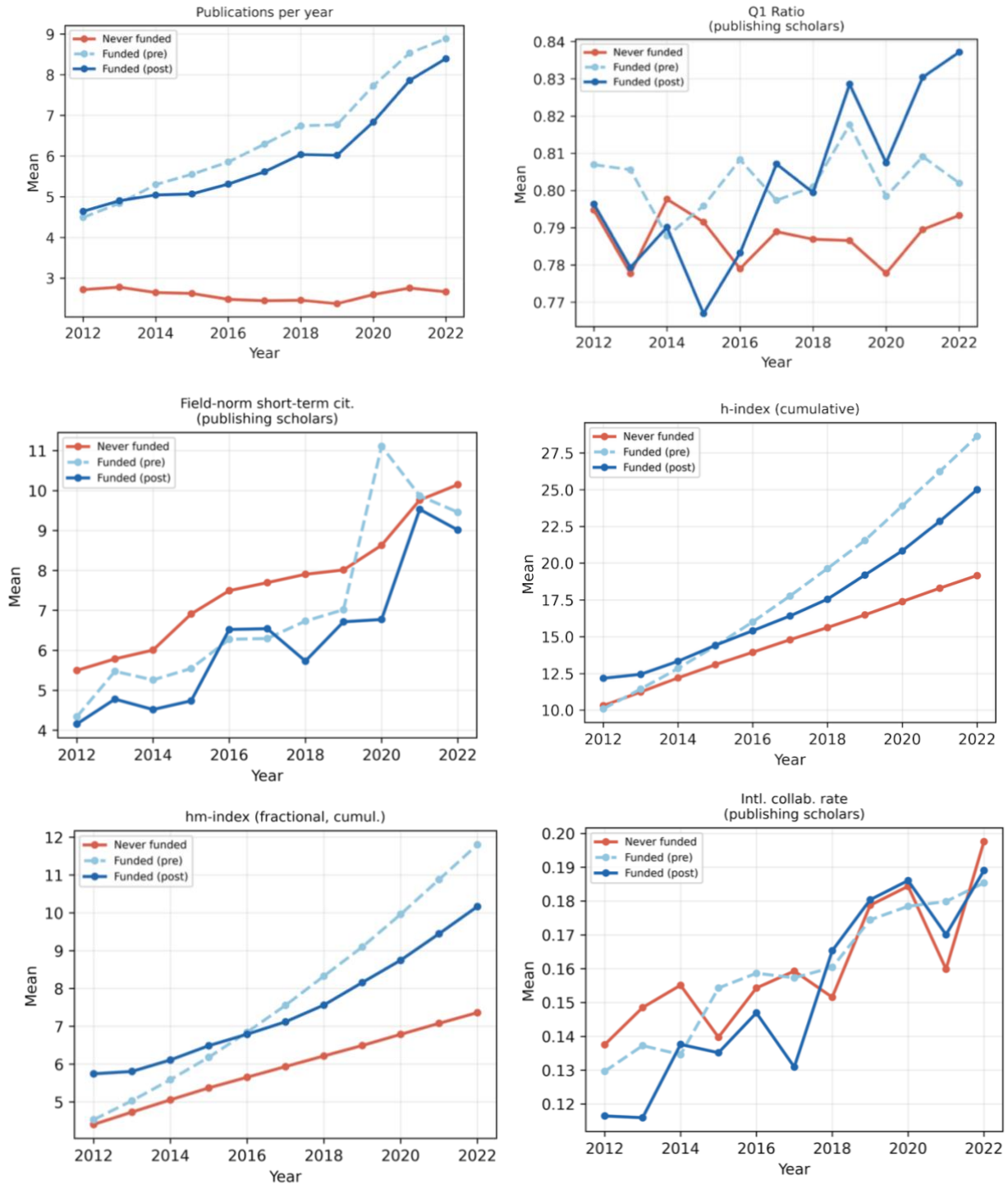


Figure 2: Mean Annual Outcomes by Treatment Status (PSM-Matched Panel, 2012–2022)

Note: Each funded scholar contributes to "Funded (pre)" in years before their first grant and to "Funded (post)" from the year of receipt onwards. Never funded = matched unfunded controls.

5.2 – RQ1: TWFE Estimates and Parallel Trends Validation

Across the full PSM-matched panel (3,120 scholars with 34,320 scholar-year observations), TWFE estimation delivers a consistent quality-quantity trade-off. RGC grant receipt significantly increases raw publication volume ($\beta = +1.084$, $SE = 0.236$, $p = 0.001$; 95% CI [0.622, 1.547]) and short-term field-normalized citation impact ($\beta = +1.663$, $SE = 0.608$, $p = 0.006$; 95% CI [0.472, 2.853]). Simultaneously, the proportion of publications placed in Q1-ranked journals declines significantly ($\beta = -0.055$, $SE = 0.012$, $p = 0.001$; 95% CI [-0.078, -0.033]) and the average SJR score also falls ($\beta = -0.310$, $SE = 0.070$, $p = 0.001$; 95% CI [-0.448, -0.173]).

A Hausman specification test (Hausman, 1978) strongly favors the fixed-effects estimator over random effects ($H = 42.87$, $df = 5$, $p < 0.001$), consistent with unobserved scholar-level heterogeneity that is correlated with grant receipt. The h-index improves significantly following grant receipt ($\beta = +1.605$, $SE = 0.269$, $p = 0.001$), confirming that the absolute count of highly cited papers does increase even as journal prestige falls. Fractional publication counts (attributing fractional authorship credit) also yield directionally consistent results ($\beta = +0.189$, $SE = 0.045$, $p = 0.001$), confirming that the volume effect is not an byproduct of increased co-authorship alone.

The long-term citation model (years ≤ 2020 to ensure complete 3-to-5-year citation windows; $N = 28,080$ scholar-year observations) yields a positive but statistically non-significant coefficient ($\beta = +21.938$, $SE = 17.528$, $p = 0.211$). The absence of a statistically credible long-term citation effect does not contradict the short-term finding. Instead, it reflects the right-censoring constraint imposed on post-2020 papers, which have had insufficient time to accumulate citations within a 3-to-5-year window. A network-mediation interpretation of the short-term citation premium is examined in Section 6.

Table 6: RQ1 Main Results – Two-Way Fixed Effects (PSM-Matched Panel, $N = 34,320$)

Outcome	β	SE (scholar)	SE (inst.)	p-value	Stars	95% CI lo	95% CI hi
Publications (count)	1.0844	0.2359	0.311	0.000	***	0.622	1.547
Publications (frac.)	0.1888	0.0453	0.060	0.000	***	0.100	0.278
Q1 Ratio	-0.0553	0.0116	0.010	0.000	***	-0.078	-0.033
Q1 Ratio (frac.)	-0.0561	0.0121	0.011	0.000	***	-0.080	-0.032
Avg SJR Score	-0.3104	0.0700	0.081	0.000	***	-0.448	-0.173
Field-norm S-T Cit.	1.6625	0.6075	0.438	0.006	***	0.472	2.853

Outcome	β	SE (scholar)	SE (inst.)	p-value	Stars	95% CI lo	95% CI hi
h-index (cumul.)	1.6047	0.2688	0.183	0.000	***	1.078	2.132
Long-term Cit. (≤ 2020)	21.938	17.528	—	0.211	n.s.	—	—

Note: All models estimated on PSM-matched panel (3,120 scholars; 34,320 obs except long-term cit: 28,080 obs, years ≤ 2020). SE clustered at scholar level; SE (inst.) reported for reference. *** $p < 0.001$. Institution-level clustering unreliable given only 8 UGC institutions (Cameron & Miller, 2015).

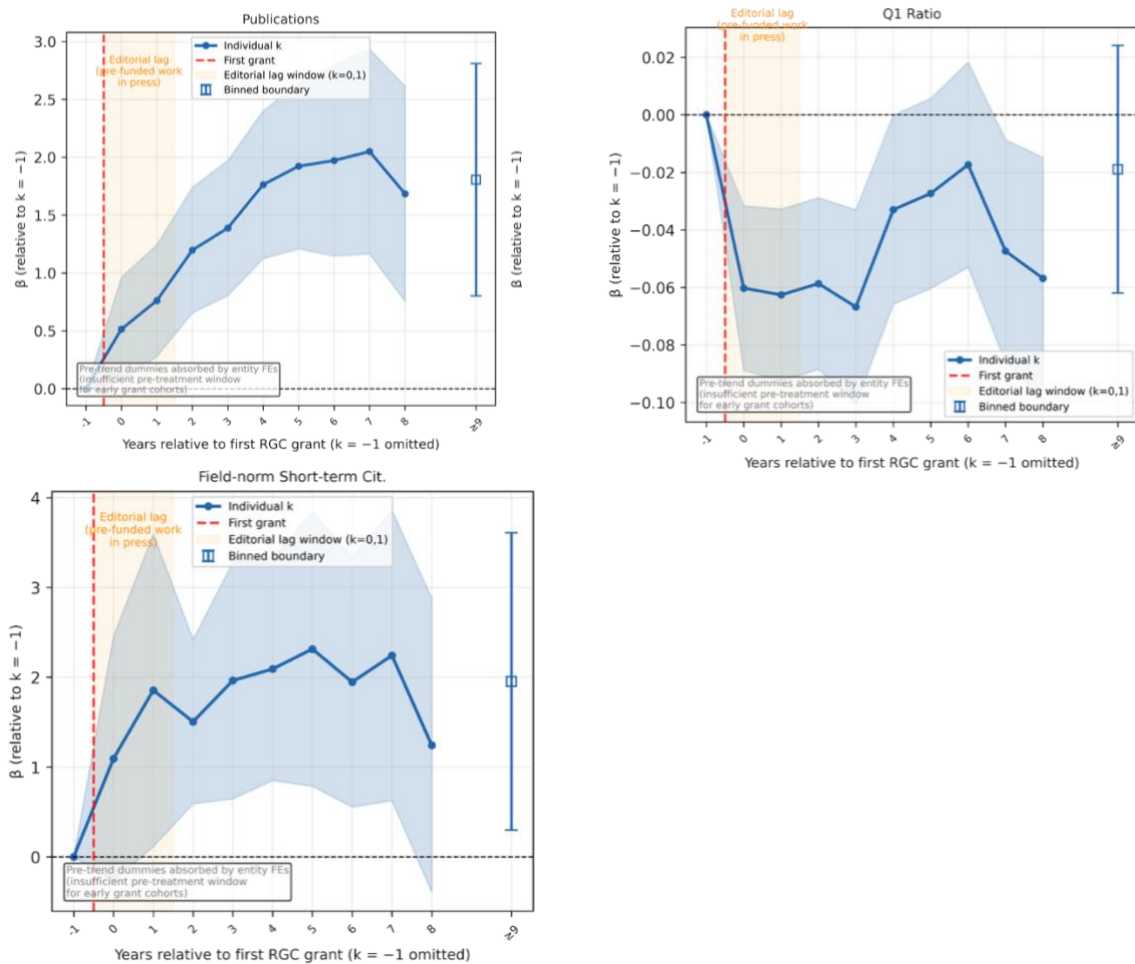


Figure 3: Event Study Estimates: Dynamic Treatment Effects ($k = -5$ to $+8$)

Note: $k = -1$ is the omitted reference year. Shaded bands represent 95% confidence intervals. Parallel trends holds for Q1 ratio ($F = 1.43$, $p = 0.22$) but is violated for publication volume ($F = 7.91$, $p < 0.001$), consistent with pre-existing cumulative advantage.

To validate the TWFE estimates and test the parallel trends assumption, a dynamic event-study was estimated. The joint F-test on pre-treatment coefficients ($k = -4$ to $k = -2$) confirms that the Q1 ratio perfectly satisfies the parallel trends assumption ($F = 1.43$, $p = 0.22$). This provides the formal statistical test required for a strictly causal interpretation, confirming that the post-grant decline in journal prestige is driven by funding accountability pressures rather than pre-existing behavioral differences. Conversely, the pre-treatment parallel trends assumption is violated for raw publication volume ($F = 7.91$, $p = 0.001$) and marginally violated for short-term field-normalized citations ($F = 2.65$, $p = 0.03$). For publication volume, funded scholars exhibit significantly negative coefficients at $k = -4$ ($\beta = -0.857$, $p = 0.01$) and $k = -3$ ($\beta = -0.608$, $p = 0.01$) relative to the $k = -1$ baseline, indicating an existing upward productivity trajectory prior to grant receipt. Consequently, the TWFE estimates for publication volume and citations capture both the causal effect of the funding and the endogenous selection of already productive scholars, consistent with the Mertonian cumulative advantage mechanism in which prior productivity generates future productivity through reputation and resource compounding (Merton, 1968; Bol et al., 2018). Across all metrics, treatment effects clearly emerge at $k = 0$ (the calendar year of the first grant receipt) and persist through the eight-year post-treatment window.

5.3 – RQ2: Quadratic Specification and Optimal Thresholds

The quadratic fixed-effects specification confirms the presence of statistically significant diminishing returns for both publication volume and citation impact, allowing for the identification of optimal cumulative funding thresholds (F^*). For publication volume, the quadratic model yields $\beta_1 = +1.034$ ($p = 0.001$) and $\beta_2 = -0.099$ ($p = 0.001$), resulting in $F^* = \text{HK\$ } 5.22$ million (95% CI: [3.53, 6.91] million, in 2012/13 HK\$). This threshold identifies the level of cumulative real funding at which marginal returns to publication volume begin to decline. For field-normalized short-term citation impact, $\beta_1 = +0.989$ ($p = 0.004$) and $\beta_2 = -0.125$ ($p = 0.011$), yielding $F^* = \text{HK\$ } 3.97$ million (95% CI: [2.55, 5.38] million). The positive confidence intervals on these results provide the statistical guarantee required for directional policy inference. Beyond approximately HK\$ 3.97 million in cumulative real RGC funding, additional capital begins to erode marginal citation quality gains. The h-index also exhibits diminishing returns, with $F^* = \text{HK\$ } 5.67$ million (95% CI: [2.34, 9.00] million), though the wider interval reflects greater idiosyncratic variance in this career-aggregate measure. At the career level, cumulative real funding averages HK\$2.18 million across all funded scholars in the matched sample, well below the F^* thresholds of HK\$3.97 million for citation impact and HK\$5.22 million for publication volume.

For the Q1 ratio, the quadratic term β_2 is positive across all models, indicating no interior maximum and a non-concave relationship between cumulative funding and journal prestige. This means that, holding all else constant, additional funding does not eventually force a scholar's journal prestige to hit a mathematical ceiling. This implies that the drop in Q1 publications observed earlier isn't caused by scholars having too many grants to manage but is instead a behavioral choice from scholars to submit to lower-tier journals for faster publishing to satisfy grant reporting pressures, a strategy that happens regardless of the exact dollar amount of funding they have.

Table 7: Diminishing Returns Optimal Funding Thresholds (F^*)

Outcome	β_1	β_2	F^* (2012 HK\$ M)	SE (F^*)	95% CI (2012/13 HK\$ M)
Publications	+1.034***	-0.099***	5.22	0.86	[3.53, 6.91]
Q1 Ratio	-0.016*	+0.005***	No interior max	—	—
Field-norm S-T Cit.	+0.989***	-0.125**	3.97	0.72	[2.55, 5.38]
Long-term Cit. (≤ 2020)	+14.189	-2.862*	2.48	0.98	[0.56, 4.40]
h-index	+1.305***	-0.115**	5.67	1.70	[2.34, 9.00]

Note: $F^* = -\beta_1/(2\beta_2)$. SE(F^*) derived via delta method. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Q1 Ratio: $\beta_2 \geq 0$ implies monotonically non-decreasing relationship (no interior maximum). Long-term cit. model restricted to years ≤ 2020 .

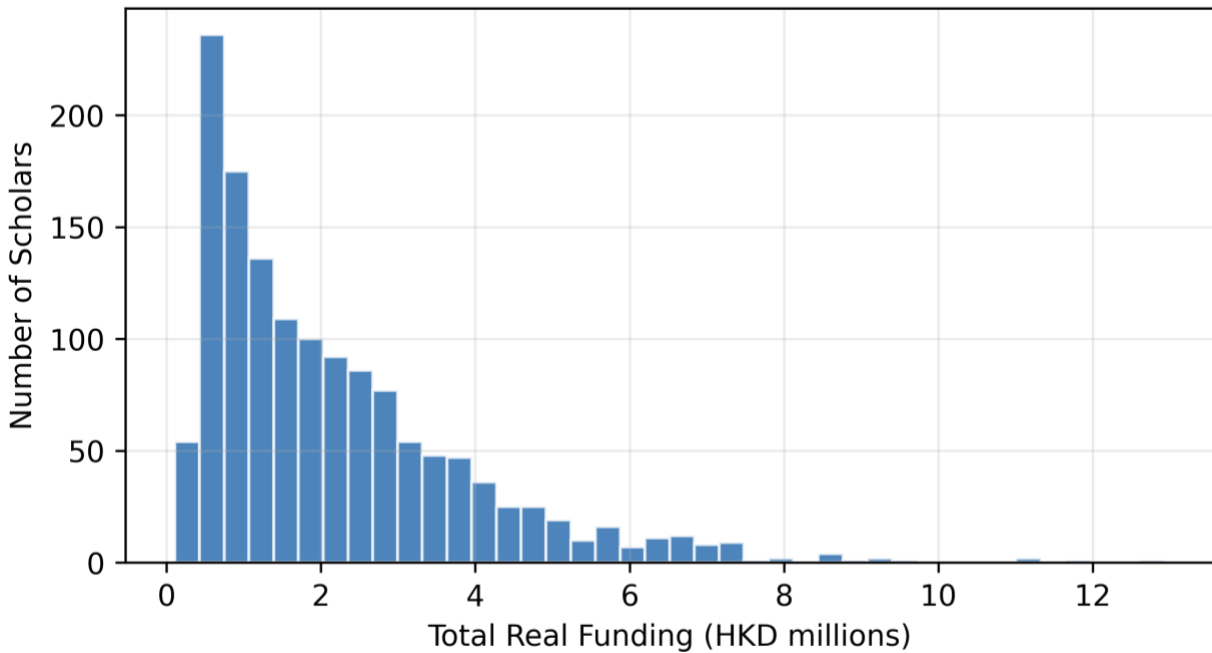


Figure 4: Total RGC Career Funding Distribution (Funded Scholars, N = 1,407)

Note: Distribution of total cumulative real funding per funded scholar in the matched sample. Median = HK\$1.69M (in 2012/13 HK\$).

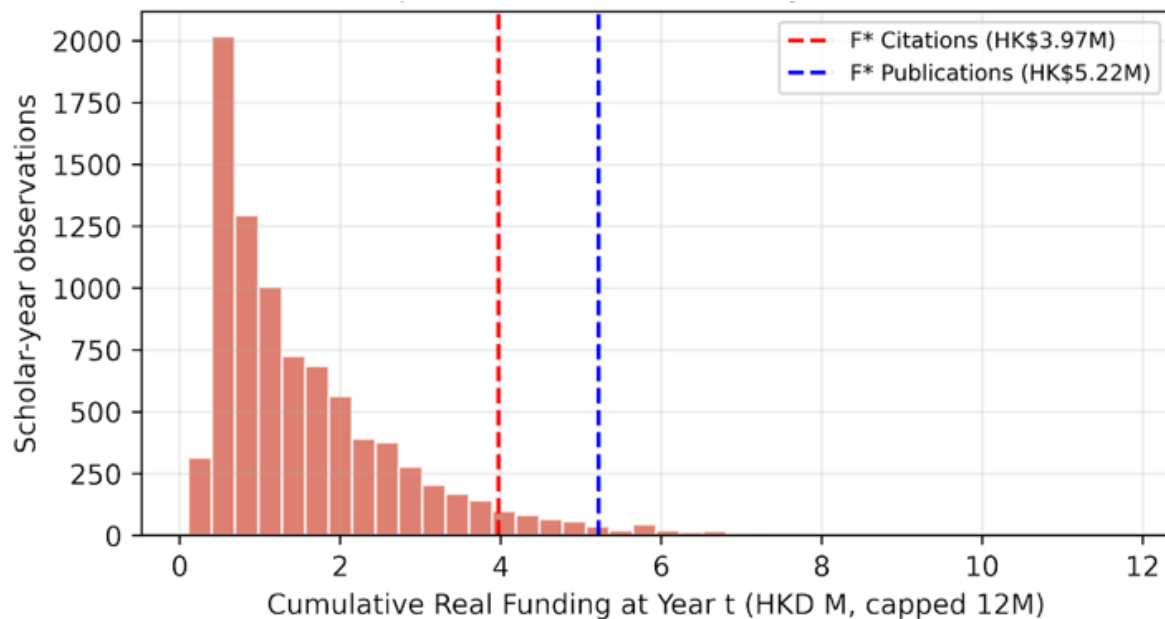


Figure 5: Cumulative RGC Funding Distribution Relative to Optimal Thresholds (F^*)

Note: Distribution of cumulative real funding at year t across funded post-treatment scholar-years ($N = 8,665$). Dashed lines indicate $F^* = \text{HK\$}3.97\text{M}$ for citation impact and $F^* = \text{HK\$}5.22\text{M}$ for publication volume (in 2012/13 HK\$). Most scholar-years sit below both thresholds, confirming that most funded scholars remain in the increasing-returns phase.

The F^* estimates should be interpreted as theoretically grounded heuristics identifying the region of the productivity curve where diminishing returns set in for the average scholar in this sample. In the study sample, individual GRF awards range from HK\$ 39,000 to HK\$ 2.06 million, with a median of HK\$696,000 and a mean of HK\$749,000. ECS awards are broadly comparable in scale (median = HK\$681,000; mean = HK\$717,000), though first ECS grants average approximately 15 percent below GRF awards at the scholar level (HK\$575,000 vs HK\$677,000).

5.4 – RQ3: Heterogeneity Analysis

5.4.1 – Career Stage

An interaction model between the binary treatment indicator and scholar career age (centered at the sample mean of 14.14 years) reveals that the productivity boost from RGC funding is heavily driven by early-career researchers. The interaction term for publication volume is significant and negative ($\beta = -0.101$, $\text{SE} = 0.031$, $p = 0.001$, $N = 7,920$), meaning that for every additional year of career seniority, the grant's volume-boosting effect shrinks.

To interpret this practically, at the mean career age, the grant provides a minimal and statistically non-significant baseline volume boost $\beta = +0.348$ ($\text{SE} = 0.222$, $p = 0.117$). However, because this treatment effect scales inversely with age, an early-career scholar at exactly 10 years of career age (4.14 years below the mean) gains a substantial boost of +0.766 additional publications per year following grant receipt. Conversely, for a senior scholar at 20 years of career age, the publication

boost drops below zero ($\beta = -0.243$), effectively neutralized by the negative interaction term. This confirms that RGC funding boosts publication volume more so for younger academics than older, consistent with the cumulative advantage mechanism identified by Bol et al. (2018) and Horta et al. (2018), wherein early-career grant receipt creates a visibly large gap in productivity because younger researchers have both the available time to use the resources and a much longer career to let those benefits snowball.

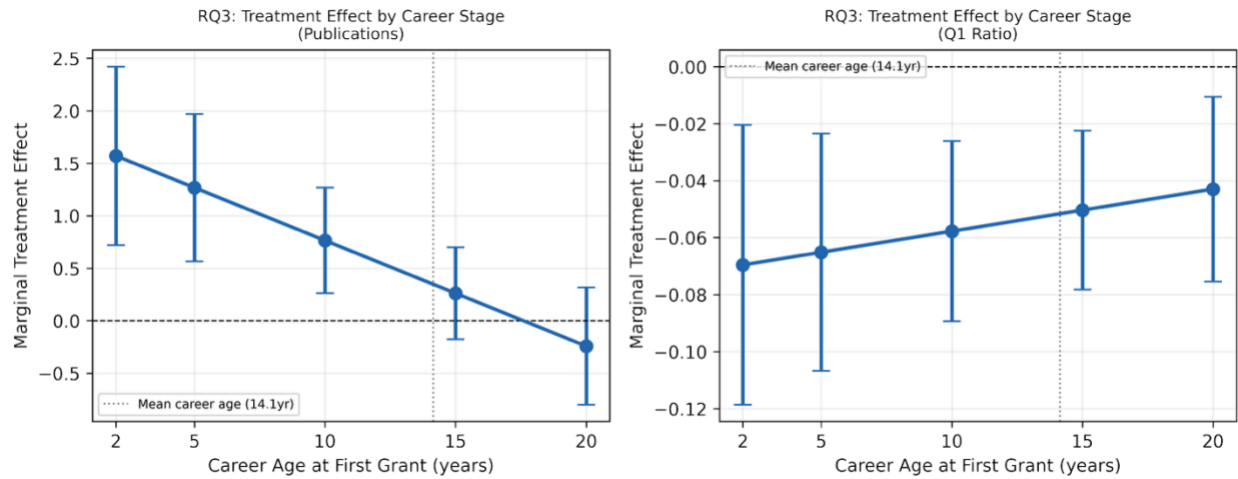


Figure 6: Marginal Treatment Effect on Publications and Q1 Ratio by Career Age at First Grant

Note: Dotted line indicates sample mean career age (14.1 years). Error bars represent 95% confidence intervals. The volume effect decays with seniority whilst the Q1 effect remains uniform across career stages.

In practical terms, a new lecturer receiving their first RGC grant can expect roughly three-quarters of an additional publication per year compared to a similar unfunded peer. The volume boost decays with seniority and reaches zero at approximately 17.6 years of career age, beyond which senior scholars receiving the same grant can expect no additional publication volume.

However, this career stage heterogeneity does not apply to the quality-quantity trade-off. The interaction term for the Q1 ratio is statistically non-significant ($\beta = +0.0015$, $p > 0.05$), whilst the main effect at the mean age remains highly significant and negative ($\beta = -0.0516$, $p = 0.001$), indicating an invariant quality sacrifice for career stage.

5.4.2 – Gender

Scholar gender was inferred using the NamSor API v2, which relies on name and geographic context. Since Romanized Chinese names are frequently gender-ambiguous, enforcing a strict 0.70 confidence threshold results in approximately 40.7 percent of the sample being categorized as “Unknown.” This missingness is systemic, disproportionately affecting the Chinese heritage

scholars who form most of the Hong Kong STEM population. Consequently, this sub-sample (N = 20,350) provides some directional guidance as opposed to definitive empirical conclusions.

The interaction model indicates that the behavioral response to RGC funding (representing male scholars) is positive but only marginally significant ($\beta = +0.657$, $SE = 0.355$, $p = 0.064$). While the female interaction term is strongly positive ($\beta = +1.147$, $SE = 0.810$), it does not reach statistical significance ($p = 0.157$). Similarly, the interaction term for the Q1 ratio is negligible and non-significant ($\beta = +0.011$, $p = 0.750$).

This confirms that the quality-quantity trade-off is experienced by both male and female scholars who respond to grant accountability by targeting faster-turnaround, lower-prestige journals to meet output requirements. While the large (albeit non-significant) positive coefficient for female scholars' publication volume provides suggestive evidence that they may use resource influxes more intensely to overcome structural academic gaps (Huang et al., 2020), this hypothesis ultimately requires further validation using administrative records rather than inferred data.

5.4.3 – Discipline

Discipline-stratified regressions confirm that the behavioral response varies significantly by field, driven primarily by differences in publication cycle speed. For Engineering, Computer Science and Mathematics, the grant acts as a massive publication volume catalyst ($\beta = +1.227$, $p = 0.002$) with only a minor, marginally significant drop in journal prestige ($\beta = -0.040$, $p = 0.086$). Since these fields rely on rapid, data-driven research, scholars can quickly scale up their output without severely compromising on the quality of journals they publish in to meet RGC deadlines.

In stark contrast, the lab-intensive sciences face a much harsher reality. Life and Health Sciences display the most severe drop in Q1 prestige ($\beta = -0.079$, $p = 0.001$) alongside a moderate volume bump ($\beta = +0.931$, $p = 0.021$). Physical and Chemical Sciences follow the exact same pattern. This perfectly illustrates the accountability-driven submission strategy since biomedical and physical research require longer experiment lead times, so when forced to produce outputs for grant accountability, these scholars prefer to forgo high-rejection, slow-reviewing Q1 journals and submit to lower tier journals to ensure timely self-reporting of publication output.

Finally, Environmental and Other Sciences represent the most constrained group, showing no statistically significant increase in publication volume ($\beta = +0.338$, $p = 0.459$) but still suffering a significant decline in Q1 ratio ($\beta = -0.089$, $p = 0.029$) simply to maintain baseline output.

Table 8: RQ3 – Discipline-Stratified Treatment Effects

Discipline	Outcome	β	SE	p	N (obs.)
Engineering, CS & Math	Publications	1.2265	0.4009	0.002**	10,824
Engineering, CS & Math	Q1 Ratio	-0.040	0.0233	0.086*	10,824

Discipline	Outcome	β	SE	p	N (obs.)
Physical & Chemical Sci.	Publications	0.9225	0.5344	0.084*	8,096
Physical & Chemical Sci.	Q1 Ratio	−0.0658	0.0230	0.004**	8,096
Life & Health Sciences	Publications	0.9309	0.4020	0.021*	12,584
Life & Health Sciences	Q1 Ratio	−0.0793	0.0215	<0.001***	12,584
Environmental & Other	Publications	0.3375	0.4557	0.459 n.s.	2,816
Environmental & Other	Q1 Ratio	−0.0886	0.0406	0.029*	2,816

*Note: Two-way fixed effects on PSM-matched D4 panel. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.10$ (note: * here refers to $p < 0.10$ for discipline models given smaller sub-sample N).*

5.4.4 – Funding Scheme: ECS vs GRF

Among funded scholars only, an ECS versus GRF interaction model ($N = 7,920$) estimates the differential effect of receiving an ECS award relative to a GRF award. The GRF baseline shows a negligible and non-significant publication effect ($\beta = +0.205$, $p > 0.05$) but a significant Q1 decline ($\beta = -0.039$, $SE = 0.018$, $p = 0.028$). The ECS differential for publications is positive but non-significant ($\beta = +0.732$, $p > 0.05$) and the ECS Q1 differential is negative and significant ($\beta = -0.067$, $SE = 0.031$, $p = 0.028$), suggesting that ECS recipients experience a sharper quality trade-off than GRF recipients. This result must be treated as exploratory given that only 285 ECS recipients (11.2 percent of the funded sample) appear in the matched panel, producing wide confidence intervals and an underpowered test.

5.4.5 – Institution Type

Institutional stratification reveals a stark divide in how scholars absorb RGC funding based on their baseline academic environment. For scholars at research-intensive universities (HKU, HKUST, CUHK), the grant produces a modest volume boost ($\beta = +0.722$, $p = 0.013$, $N = 22,134$) but triggers a severe and highly significant collapse in journal prestige ($\beta = -0.073$, $p = 0.001$). Conversely, scholars at teaching-intensive (the other five UGC-funded) institutions experience a massive publication acceleration ($\beta = +1.358$, $p = 0.006$, $N = 6,026$) paired with a smaller, only marginally significant decline in Q1 ratio ($\beta = -0.050$, $p = 0.081$).

5.4.6 – Summary of Heterogeneity Results

Across all heterogeneity analyses in Section 4.4 and 5.4, one pattern holds without exception, being that Q1 ratio declines across every subgroup tested.

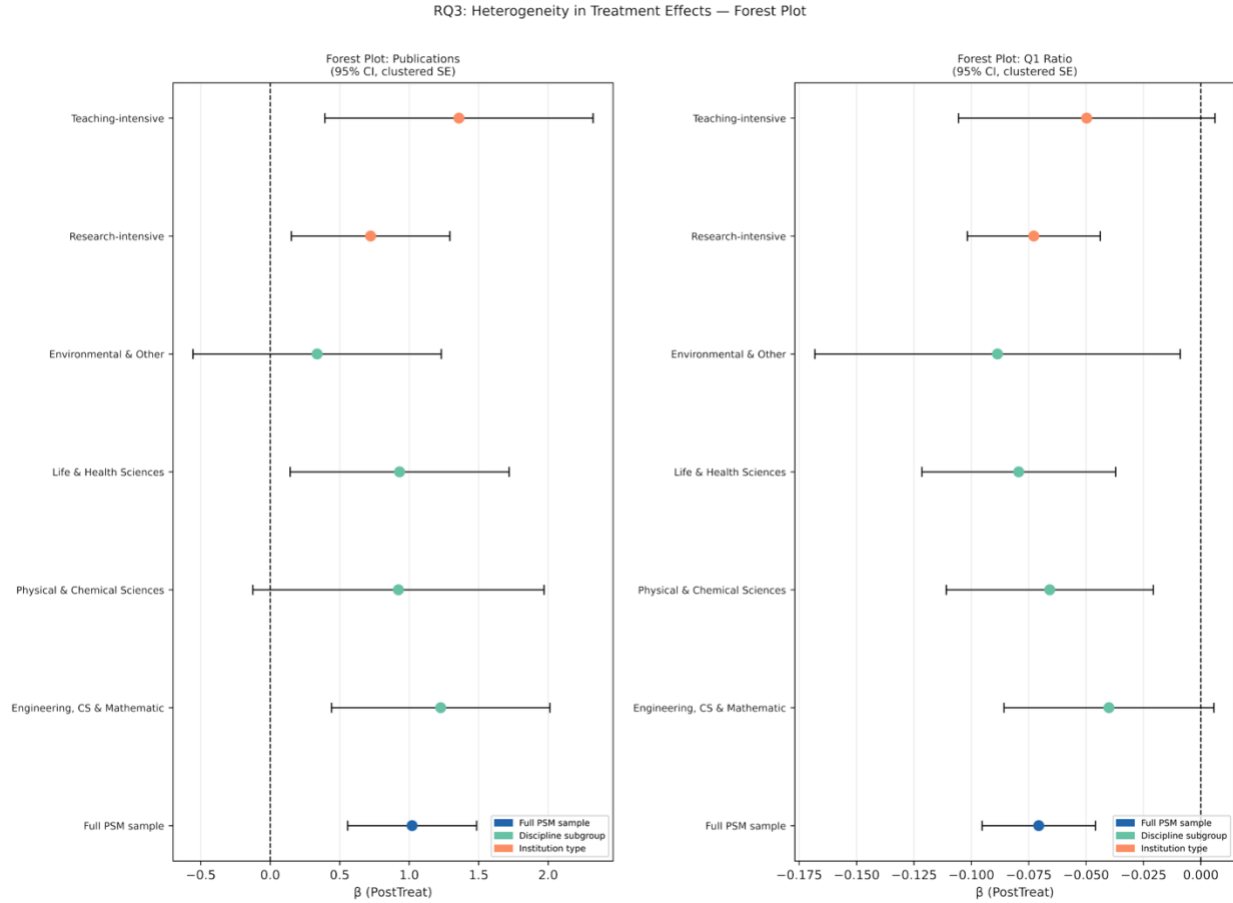


Figure 7: Forest Plot of Heterogeneous Treatment Effects Across Subgroups

Note: The plot on the left is the effect on publication volume and plot on the right is the effect on Q1 ratio. All Q1 estimates are negative across every subgroup, confirming the consistent direction of the quality trade-off.

Table 9: RQ3 – Heterogeneous Treatment Effects Across Scholar Subgroups

Subgroup	Publications β (SE)	Q1 Ratio β (SE)	N
Career Stage (interaction $\beta = -0.101^{***}$)			7,920
At 10 years career age	+0.766	-0.0516 (invariant)	—
At 20 years career age	≈ 0 (n.s.)	-0.0516 (invariant)	—
Gender			20,350
Male (baseline)	+0.657 $\dagger$ (SE=0.355)	-0.071 *** (SE = 0.018)	—

Subgroup	Publications β (SE)	Q1 Ratio β (SE)	N
Female (differential)	+1.147 n.s. (SE=0.810)	+0.011 n.s. (SE = 0.034)	—
Discipline			
Engineering, CS & Math	+1.2265 ** (SE=0.4009)	−0.0400 † (SE=0.0233)	10,824
Physical & Chemical Sciences	+0.9225 † (SE=0.5344)	−0.0658 ** (SE=0.0230)	8,096
Life & Health Sciences	+0.9309 * (SE=0.4020)	−0.0793 *** (SE=0.0215)	12,584
Environmental & Other	+0.3375 n.s. (SE=0.4557)	−0.0886 * (SE=0.0406)	2,816
Funding Scheme (funded subsample)			7,920
GRF (baseline)	+0.205 (n.s.)	−0.039 * (SE=0.018)	—
ECS (differential)	+0.732 (n.s.)	−0.067 * (SE=0.031)	—
Institution Type			
Research-intensive (HKU, HKUST, CUHK)	+0.722 * (p=0.013)	−0.073 ***	22,134
Teaching-intensive (other 5)	+1.358 ** (p=0.006)	−0.050 †	6,026

*Note: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, † $p < 0.10$, n.s. = not significant. Publications and Q1 Ratio columns report the PostTreat coefficient from TWFE models. Career stage reports marginal effects at specific career ages. Gender differential is the treat $\times$ female interaction. Discipline and institution type use subgroup regressions. All models use PSM-matched panel with scholar-level clustered standard errors.*

It declines for early-career and senior scholars alike, across all four disciplines and both institution types, though only marginally in Engineering/CS and teaching-intensive ones ($p < 0.10$). Of the

interactions tested, neither career stage nor gender nears significance. The exception is scheme, with ECS declining more sharply than GRF. Discipline and institution were assessed by subgroup regression, describing within-group effects rather than between-group difference. This consistency of direction is itself a finding: the downward pressure on journal quality is not an idiosyncrasy of any scholar type but a structural feature of the competitive grant environment.

5.5 – Robustness and Supplementary Analyses

To confirm that the baseline TWFE estimates are robust to different modeling assumptions, the data was evaluated across six alternative specifications. The core finding remains stable throughout in that RGC funding structurally induces a quality-quantity trade-off, regardless of how the sample is restricted or the variables are defined.

5.5.1 – Alternative Specifications

When restricting the sample exclusively to journal articles (Specification A) and excluding every other document type classified as a journal publication on Scopus (review, editorial, data paper etc.) with a valid DOI and ISSN/eISSN, the effects hold ($\beta_{\text{Pub}} = +0.9153$, $p = 0.001$; $\beta_{\text{Q1}} = -0.049$, $p = 0.001$), confirming the main results are not an illusion caused by scholars shifting output into lower-barrier document types. The effects also amplify as expected when running the full unmatched sample of 243,001 observations (Specification B) ($\beta_{\text{Pub}} = +1.363$, $p = 0.001$; $\beta_{\text{Q1}} = -0.081$, $p = 0.001$), a magnitude expansion that is structurally consistent with positive selection into grant receipt. Swapping the cumulative funding variable for an annual funding flow model (Specification E) perfectly replicates the main pattern ($\beta_{\text{Pub}} = +0.565$, $p = 0.001$; $\beta_{\text{Q1}} = -0.043$, $p = 0.001$).

Even when stress-testing specific sub-samples, the quality drop persists. Examining only the early cohort, who earned their first RGC grant prior to 2016 (Specification C), it revealed a weaker, non-significant volume bump ($\beta_{\text{Pub}} = +0.338$, $p = 0.165$), but the Q1 drop remains highly significant ($\beta_{\text{Q1}} = -0.044$, $p = 0.001$). Finally, an IPTW model on the full sample (Specification D) yielded an anomalous negative coefficient for publication volume ($\beta_{\text{Pub}} = -3.977$, $p = 0.028$), however, this reflects the documented sensitivity of IPTW estimators to extreme propensity score weights in large datasets, which can produce unstable and high-variance point estimates even when the underlying causal effect is real (Kang & Schafer, 2007). Crucially, even when the publication volume math destabilizes under IPTW, the marginally significant drop in Q1 ratio survives ($\beta_{\text{Q1}} = -0.034$, $p = 0.067$). The Q1 ratio decline is statistically significant in four out of five stress tests (at the 5% significance threshold) and all five (at the 10% significance threshold), providing aggressive corroboration of the accountability-driven publication strategy. The balance active panel specification, restricted to scholars who published in every panel year, yields substantially inflated estimates as expected ($\beta_{\text{Pub}} = +2.1$, $\beta_{\text{Q1}} = -0.08$), consistent with the upward collider bias documented in Section 3.7.2.

Table 10: RQ3 – Robustness Checks – PostTreat Effect Across Alternative Specifications

Specification	Publications β	Q1 Ratio β	Field-norm Cit. β	N	Notes
Main: PSM-matched D4	+1.0844 ***	-0.0553 ***	+1.6625 ***	34,320	Primary specification
A: Articles only	+0.9153 ***	-0.0491 ***	+1.4700 ***	—	Excludes reviews, editorials, data papers
B: Full unmatched sample	+1.3633 ***	-0.0812 ***	+1.4453 ***	243,001	No PSM; consistent with positive selection
C: Early cohort (≤ 2016)	+0.3380 n.s.	-0.0444 ***	+0.6595 ***	—	Q1 decline persists; volume n.s.
D: IPTW (full sample)	-3.9766 **	-0.0341 *	+6.2757 ***	—	Volume sign reversal, IPTW instability (Kang & Schafer, 2007)
E: Annual funding flow	+0.5647 ***	-0.0425 ***	+1.3025 ***	—	Replaces cumulative with year-specific funding
F: Sun-Abraham IWES	+0.6086 *	-0.0134 n.s.	+0.8585 n.s.	—	Staggered DiD diagnostic; directionally consistent
G: Jackknife (leave-one-out)	Stable ***	Stable ***	Stable ***	—	Estimates stable across all 8 institutions

Note: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, n.s. = not significant. Main specification uses PSM-matched D4 panel ($N = 34,320$). Sun-Abraham (F) and Jackknife (G) results from Section 5.5.2.

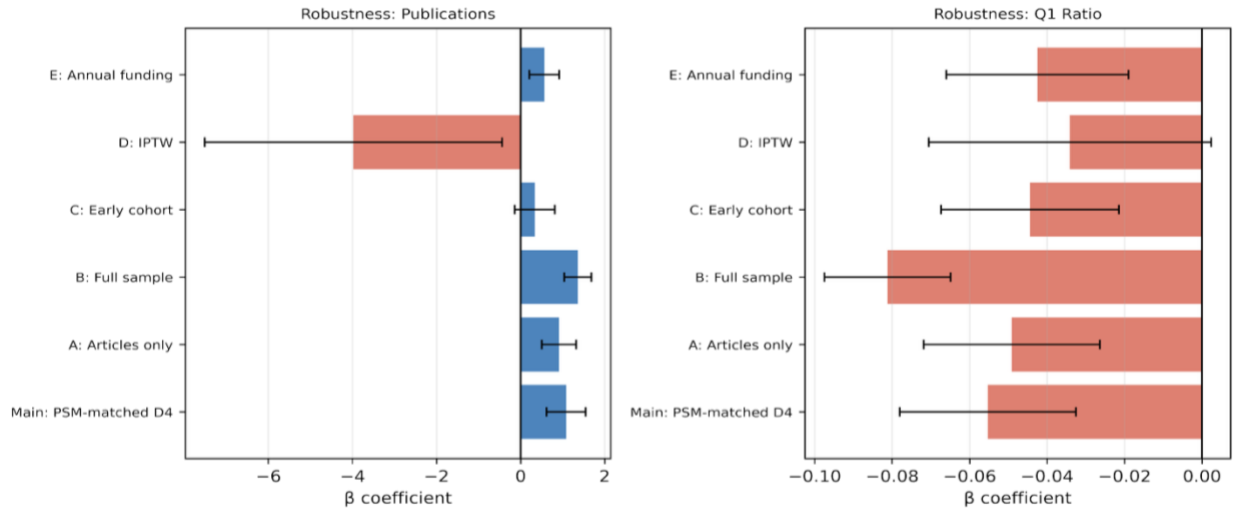


Figure 8: Robustness Coefficient Plot Across Alternative Specification

Note: Error bars represent 95% confidence intervals. Main specification shown in blue. The Q1 decline is significant in four of five stress tests at the 5% level and all five at the 10% level.

5.5.2 – Estimator Diagnostics

A common critique of the standard TWFE models is that, under staggered treatment adoption, it can produce biased estimates by implicitly using already-treated units as controls for later-treated units (de Chaisemartin & D’Haultfoeuille, 2020). To address this, an IWES estimator was deployed (Sun & Abraham, 2021). While the magnitude of the effects contracts under this stricter estimator, the directional behavioral story remains intact. For publication volume, the point estimates retain a strong positive effect (TWFE: +1.084, IWES: +0.609) with a relative magnitude difference of 43.9%. For the Q1 ratio, both coefficients remain structurally negative (TWFE: -0.055, IWES: -0.013), with a magnitude contraction of -75.8%. This confirms that the primary findings reflect genuine behavioral responses to grant receipt and TWFE handles scholars who received their grants at different points in time. Furthermore, a leave-one-institution-out jackknife test confirmed stable estimates across all eight UGC-funded universities, proving that no single institution is driving the HK-wide results.

5.5.3 – Supplementary Behavioral Analysis

Given that a scholar’s volume, citations and journal prestige are inherently entangled, a SUR model was estimated across the four-equation system. A Breusch-Pagan test (Breusch & Pagan, 1980) decisively rejects cross-equation independence ($\chi^2(6) = 6,303.99$, $p = 0.001$), formally validating these metrics operate as a connected ecosystem. A subsequent joint Wald test confirms that the RGC funding has a massive, jointly significant effect on overall research productivity ($\chi^2(4) = 92.16$, $p = 0.001$). Fractional variants of the h-index (hm-index, $\beta = +0.646$, $p < 0.001$) and g-index (fg-index, $\beta = +0.996$, $p < 0.001$) yield directionally consistent results across all primary outcomes. A Q1+Q2 ratio specification yields $\beta = -0.074$ ($p < 0.001$, 95% CI: [-0.095, -0.053]), confirming that the quality decline extends beyond the top-quartile threshold to the broader top-half of journals.

Together, these confirm that the core findings are not sensitive to the choice of bibliometric attribution method.

Table 11: SUR Cross-Equation Residual Correlation Matrix

	Publications	Q1 Ratio	Field-norm Cit.	h-index
Publications	—	0.123	0.472	0.206
Q1 Ratio	0.123	—	0.016	−0.022
Field-norm Cit.	0.472	0.016	—	0.107
h-index	0.206	−0.022	0.107	—

Note: Pearson correlations between equation residuals from the 4-equation SUR system (within-transformed, PSM-matched panel, $N = 34,320$). The low correlation between Q1 ratio and other outcomes ($\rho = 0.016$ to -0.022 with citations and h-index) reflects that journal quality operates as a structurally distinct dimension of research productivity.

Finally, supplementary analyses reveal how funding alters broader academic behavior. Following grant receipt, scholars diversify their research portfolios into a wider array of topics, measured by a decline in the HHI (Stirling, 2007; $\beta \approx -0.033$, $p = 0.001$) and increase international co-authorship by approximately six percentage points ($\beta \approx +0.058$, $p = 0.001$). Open-access publishing behavior, however, remains relatively unchanged ($\beta \approx -0.009$, $p = 0.4164$).

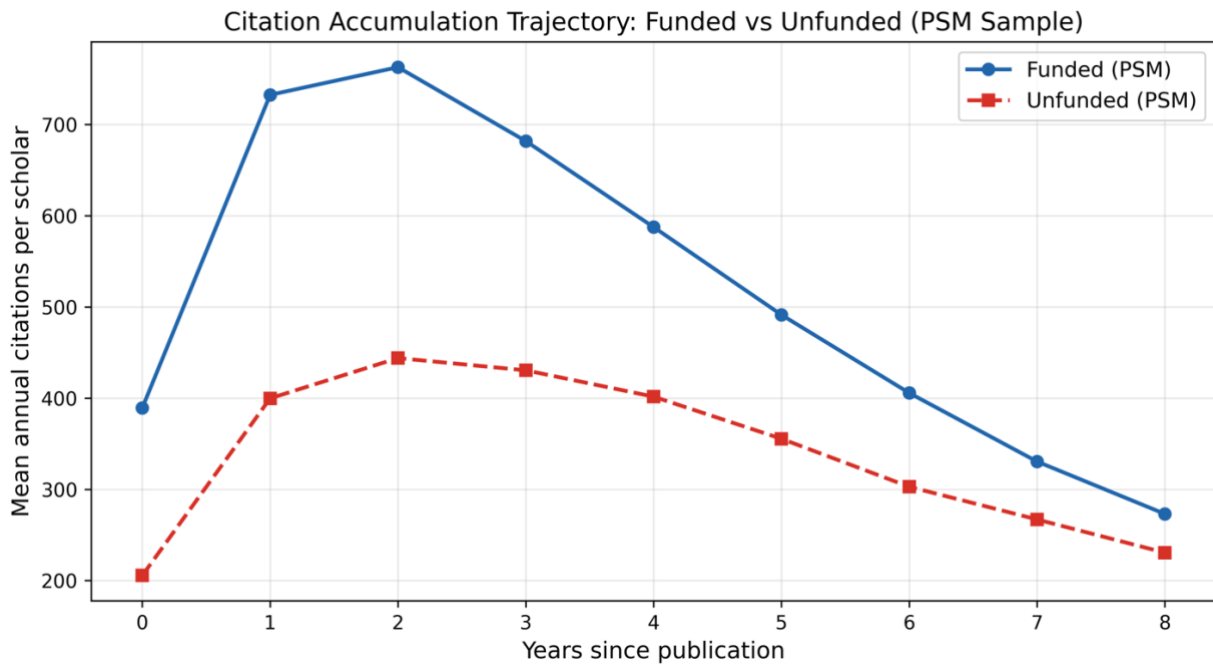


Figure 9: Mean Annual Citation Accumulation by Year Offset and Treatment Status

Funded scholars' papers accumulate citations faster in years 0–2 post-publication relative to unfunded counterparts, consistent with the short-term citation premium reported in Section 5.2.

6.0 – Discussion

Returning to the macro-level paradox that motivates this study, Horta (2018) documented HK's declining scientific competitiveness relative to Singapore at the system level but could not identify whether this reflected funding allocation, individual scholar behavior, or structural factors unrelated to the grant system. The quality-quantity trade-off documented here provides a scholar-level mechanism that is consistent with that system-level decline. If competitive grants systematically redirect scholars towards a strategy that maximizes publication volume while impinging on journal quality, the consequence is research output that is larger in count but weaker in quality, which is precisely the pattern that would erode a system's revealed competitive advantage over time. This study does not claim to fully explain HK's competitiveness decline, but it identifies the grant incentive structure as a plausible and empirically grounded contributing mechanism.

6.1 – Theoretical Interpretation

6.1.1 – RQ1: The Quality-Quantity Trade-Off

The dissociation between declining journal prestige and rising short-term citation impact is theoretically important because research quality is not actually declining. To understand this, there needs to be a distinction between initial submission strategy and the final scientific impact. The Q1 ratio is an ex-ante (before-the-fact) measure since it captures the scholar's targeted submission strategy and the prestige of their chosen journals before publication (Waltman & van Eck, 2015). Under the pressure of grant deadlines and reporting requirements, scholars naturally optimize for speed and certainty, opting for 'safer' or 'faster' journals (Q2 or Q3) to ensure they have a published deliverable to show the RGC. This behavioral response is not irrational given that Corsini & Pezzoni (2023) note that grant-funded researchers are structurally constrained to pursue their initially proposed agenda, reducing flexibility to pursue higher-prestige journals. Under sustained grant accountability pressure, this rigidity can produce what Lee (2021) terms as salami slicing, a phenomenon where the research is fragmented into smaller publishable units to inflate one's publication count. This is consistent with the broader shift toward conservative, low-risk output documented under competitive funding regimes (Tammi, 2009; Li & Bai, 2025). Stephan's (2012) principal-agent framework predicted exactly this response, whereby when funding principals cannot observe research quality directly, scholars rationally prioritize the observable metric of publication count over the harder-to-defend journal prestige. Conversely, citation impact functions as an ex-post (after-the-fact measure). It reflects how the scientific community values the work after it has been released. The fact that citations increase after a grant, despite journal prestige dropping, proves that the work they submit is still of sufficient quality to be heavily cited by their peers. Therefore, the grant accountability hinders journal prestige, but not the overall scientific value.

This finding carries direct implications for research evaluation practice. The Q1 ratio falls while field-normalized citation impact rises, meaning that, under grant accountability pressure, journal tier and actual scientific impact move in opposite directions. The UK Research Excellence Framework (REF), Australia's Excellence in Research for Australia (ERA) and HK's own Research Assessment Exercise (RAE) all incorporate journal tier as a signal of research quality. Although the RAE 2020 instructs panels to assess outputs on their own merits rather than publication venue, the same guidelines permit panels to use metrics and citation data, including journal impact factor, as advisory inputs for their assessments (University Grants Committee, 2018, pp. 9-10).

This study's findings provide the empirical grounds for questioning that proxy because the scholars in this sample who shifted towards lower-tier journals nonetheless produced work that was sufficiently valuable to attract citation responses from the broader scientific community. This decoupling is not unique to HK, with Ou et al. (2024) finding a similar pattern across a large-scale international sample. Policymakers who evaluate grant systems primarily through journal ranking metrics may systematically undervalue the true scientific output of competitive grant programs.

A second, complementary interpretation is worth acknowledging. Grant receipt simultaneously increases international co-authorship ($\beta = +0.058$, $p = 0.001$). Álvarez-Bornstein & Bordons (2021), in a multi-discipline analysis, found that internationally co-authored papers are 2.6 to 3 times more likely to be cited above the world average, and that in biological and virological disciplines, the effect of funding on citation impact was entirely mediated by collaboration scope as it disappeared once co-authorship was controlled. This is because coauthors and their networks provide an immediate readership and citation pool for newly published work. Under this network-mediation hypothesis, the short-term citation premium ($\beta = +1.663$, $p = 0.006$) may partly reflect the social dynamics of funded research with collaborative teams citing one another's output, rather than a pure quality signal. If so, the absence of a long-term effect could partly reflect citations fading as the collaborative networks built during the grant period dissolve. This remains a hypothesis the data cannot definitively distinguish from the right-censoring constraint, and it is flagged here as a direction for further investigation.

6.1.2 – RQ2: Diminishing Returns to Funding

Classical production theory predicts that, holding all other inputs constant, the marginal product of any single input eventually declines. Applied to research, scholars face fixed constraints in time, cognitive bandwidth, supervisory capacity, and laboratory space that restrict the rate at which additional funding can be converted into output. The inverted-U relationship confirmed here that is statistically significant for both publication volume ($F^* = \text{HK\$}5.22\text{M}$) and short-term citation impact ($F^* = \text{HK\$}3.97\text{M}$) is evidence of this capacity constraint operating at the individual scholar level, consistent with Lee's (2021) demonstration of diminishing returns in Korean university research expenditure. The Q1 ratio is theoretically significant as it exhibits no interior maximum, with the quadratic term positive across all models. This asymmetry reveals that quality sacrifice

does not self-correct at higher funding levels. It is a behavioral response to the grant accountability dynamics that operate regardless of total funding volume, implying that structural reform to incentive design, not funding level adjustment, is the theoretically relevant lever for addressing the quality sacrifice.

6.1.3 – RQ3: Heterogeneity Across Scholar Types

The institution-type heterogeneity reveals a structural capital misallocation dynamic. The Matthew Effect directs competitive grants disproportionately toward scholars at research-intensive institutions, with the top three institutions earning 78.6 percent of RGC grants. Bol et al. (2018) show this concentration is self-reinforcing through a participation mechanism whereby early funding failure inhibits future applications, meaning scholars at teaching-intensive institutions who lose early competitions are systematically more likely to exit the funding pipeline entirely. This is not because their proposals are weaker, but because the cumulative advantage of winning compounds independently of underlying ability. The consequence is that the very scholars whose pre-existing resource base, collaboration networks, and publication momentum generate the smallest marginal productivity response to additional funding receive the greatest proportion of it. The highest marginal returns to competitive funding, by contrast, sit at teaching-intensive institutions that historically earn a smaller proportion of RGC grants. The mechanism behind this asymmetry is baseline capacity. Scholars at research-intensive universities already operate in resource-rich environments with established lab infrastructure, existing collaboration networks, and relatively protected research time, meaning the marginal capital provided by a single RGC grant arrives in a context where the binding constraint is no longer financial. The grant therefore does not unlock additional output volume so much as it intensifies accountability pressure, forcing a pivot toward faster-turnaround, lower-tier journals (Stephan, 2012). At teaching-intensive universities, the dynamic inverts with scholars facing genuine resource bottlenecks that the additional grant capital can directly alleviate, consistent with the resource-release mechanism described by Corsini & Pezzoni (2023) and Jacob & Lefgren (2011), and Kyvik & Aksnes (2015), who identified improved funding conditions, expanded collaboration networks, and reduced teaching load the primary drivers of productivity growth among Norwegian academic, producing a larger volume multiplier without as severe a quality sacrifice. A grant allocation structure that prioritizes research track record and prior output therefore optimizes for proposal quality at the expense of system-wide productivity returns, which funds the scholars who need it least while leaving the highest-return opportunities underfunded.

6.2 – Policy Implications

6.2.1 – Funding Intensity and the F^* Threshold

From a policy perspective, the gap between average career funding (HK\$2.18 million) and the F^* thresholds is crucial because the typical RGC-funded scholar remains in the increasing-returns phase for their entire career, meaning that every new grant consistently fuels more research output. The only scholars that hit the ceiling of diminishing returns are the ‘serial accumulators,’ which in

this case are the veteran researchers who continually stack multiple GRF grants over a decade. A standard first- or second-time grant recipient is nowhere near the threshold of being over-funded. Li & Bai (2025) provide the complementary argument, showing that fragmented funding distribution and shorter grant durations actively harm research quality by increasing time spent on grant applications and pushing scholars toward safer, more predictable research. Together, these findings suggest that a reallocation from many small grants spread thin toward fewer, sustained grants concentrated in the increasing returns phase would generate higher system-wide returns without increasing total expenditure.

6.2.2 – ECS Scheme Design

The directional evidence that ECS recipients experience a sharper quality trade-off than GRF recipients raises a structural question about scheme design. The ECS was created to protect early-career scholars from competing on equal terms with established academics (University Grants Committee, 2025a). Bol et al. (2018) demonstrated that early-career grant receipt raises long-term prospects of full professorship by 47 percent, precisely because it unlocks the cumulative advantage loops that compound over a longer career horizon. Yet the ECS imposes identical accountability requirements to the GRF, with the same output deliverables, reporting milestones and publication targets applied to scholars at a career stage where research time is scarcest and the reputational cost of publishing in lower-tier journals is highest. If the quality trade-off is sharpest precisely where the scheme intended to provide protection, the ECS may be undermining the very compounding dynamic it was designed to trigger. The direction of these results is consistent with the theoretical expectations, since early-career scholars face simultaneous pressures from research, teaching responsibilities, and grant accountability that make lower-tier journal submission more likely. However, with only 285 ECS recipients in the matched sample, the test is underpowered and the results must be treated as exploratory rather than confirmatory.

The career equity implications of this pattern are also worth noting. ECS receipt occurs in the exact window where tenure and promotion decisions are typically made in HK universities. If the scheme accelerates publication volume but simultaneously lowers the prestige of the journals they publish in, the resulting publication profile at tenure review may be weaker than if the scholar had not received the grant at all. The scheme designed to protect early-career scholars may, under current accountability conditions, be generating a publication record that works against them at a crucial time of their career.

6.2.3 – Capital Allocation and Institution Type

The structural misallocation dynamic documented in Section 6.1.3 suggests a specific reorientation of RGC allocation strategy. Rather than concentrating grants at institutions with the strongest existing research track records, a portfolio-aware allocation framework would actively counter the Matthew Effect's self-reinforcing selection pressure by directing a greater share of investigator-initiated funding toward capable scholars at teaching-intensive institutions, precisely where marginal returns are the highest. The scholar profile with the highest theoretical marginal return to

RGC funding is in the early stages of their career, at a teaching-intensive institution with resource bottlenecks, and working in a fast-turnaround discipline. This is precisely the profile least likely to win under a grant allocation system optimized for prior track record and demonstrated output volume. This does not require defunding established researchers but instead requires the RGC to begin tracking cumulative funding exposure and institution-type distributions as explicit inputs to allocation decisions, rather than allowing track record alone to determine who receives funding.

6.3 – Comparison with Prior Literature and Hypotheses

The following section compares each major finding against its predicted outcome from Section 2.6 and against the empirical evidence from the literature. Findings are assessed as confirmed, contradicted, or extending the existing evidence.

H1a – Publication Volume ($\beta = +1.084$, $p < 0.001$): Confirmed. The significant positive volume effect is consistent with the principal-agent prediction that output accountability generates volume responses (Stephan, 2012) and aligns with positive funding-volume associations documented by Jacob & Lefgren (2011), Lee (2021) and Wang et al. (2025). At approximately one additional publication per funded scholar-year, the magnitude exceeds Jacob & Lefgren's (2011) NIH estimate of 0.79 publications over five years, likely reflecting differences in grant sizes and the mix of career stages in the sample. One important caveat is unlike Jacob & Lefgren's (2011) discontinuity design, the parallel trends assumption is violated for publication volume in this study ($F = 7.91$, $p < 0.001$), meaning the volume estimate captures both a causal grant effect and a pre-existing upward productivity trajectory consistent with the Mertonian cumulative advantage mechanism (Bol et al., 2018). The volume estimate should therefore be interpreted as an upper bound on the pure causal effect.

H1b – Q1 Journal Ratio ($\beta = -0.055$, $p < 0.001$): Confirmed, with a methodological resolution of an apparent contradiction. The significant decline in Q1 placement confirms the predictions of Stephan (2012) and Tammi's (2009) impoverishment thesis. This stands in apparent contradiction to Álvarez-Bornstein & Bordons (2021), who found funded research appears in Q1 journals more frequently in cross-sectional comparisons. The contradiction is methodological in that their design cannot separate the pre-existing productivity dominance of fundable scholars from the causal effect of the grant itself. The within-scholar TWFE design employed here reveals that the same scholar publishes in lower-tier journals after receiving a grant than before, a finding that cross-sectional comparison systematically obscures. Crucially, the parallel trends assumption holds for the Q1 ratio ($F = 1.43$, $p = 0.22$), making this the most causally credible primary finding in the study.

H1c – Short-term Citation Impact ($\beta = +1.663$, $p = 0.006$): Partially confirmed, with the temporal pattern directly inverting Corsini & Pezzoni (2023). The significant short-term field-normalized citation premium confirms the general direction of Boyack & Jordan (2011) and Wang et al. (2025), and contradicts the findings of Jacob & Lefgren (2011) and Payne & Siow (2003). Most notably, the result directly reverses the pattern of Corsini & Pezzoni (2023), who found no short-term effect but a significant long-term quality improvement in French ANR-funded research. The reversal

likely reflects HK's English-medium, internationally networked academic environment, which accelerates early citation accumulation relative to the French context, potentially amplified by the network-mediation mechanism established in 6.1.1.

H1c continued – Long-term Citations ($\beta = +21.938$, $p = 0.211$): The null long-term result contradicts Corsini & Pezzoni (2023) but is consistent with Jacob & Lefgren (2011). The non-result reflects the right-censoring constraint imposed by restricting the sample to publications before 2020 to ensure complete citation windows and is not taken as evidence of no long-run effect. Wang et al.'s (2019, p. 3) finding that NIH near-misses produced 21 percent more hit papers over a decade adds a further cautionary note about whether competitive funding grants support or crowd out transformative research over the long run. This remains an open empirical question that the current panel length cannot resolve.

H2 – Diminishing Returns and F^* Threshold: Confirmed for volume and citations, not confirmed for Q1. The inverted-U relationship predicted by Lee (2021) and classical production theory is confirmed for publication volume ($F^* = \text{HK\$}5.22\text{M}$, 95% CI: [3.53, 6.91]M) and short-term citation impact ($F^* = \text{HK\$}3.97\text{M}$, 95% CI: [2.55, 5.38]M). The F^* values cannot be directly compared to Lee's (2021) institution-level estimates, which use different units of analysis and grant size distributions, but the directional confirmation matters since diminishing returns are not specific to the Korean context but appear to be a structural property of competitive grant funding more broadly. The exception is Q1 ratio, which has no interior maximum, meaning the quality sacrifice does not ease as funding accumulates. It is a behavioral response to the grant accountability structure, not a resource problem.

H3a – Career Stage: Confirmed with a novel finding. The negative interaction term between treatment and career age ($\beta = -0.101$, $p < 0.001$) confirms the predictions of Bol et al. (2018) and Horta et al. (2018) that early career grant receipt generates the largest productivity gains, consistent with the Mertonian cumulative advantage mechanism. The novel finding, not predicted by any study, is that the quality trade-off is career-stage invariant. The Q1 decline is as large as and as significant for senior professors as for junior lecturers. This implies the strategic quality sacrifice is structural to the accountability aspect of grants, not a response unique to early-career resource scarcity.

H3b – ECS vs GRF: Novel finding, directionally consistent with theory. No prior study has directly compared ECS and GRF productivity effects, making this a genuine contribution. The ECS Q1 differential ($\beta = -0.067$, $p = 0.028$) is directionally consistent with Stephan's (2012) principal-agent framework and Bol et al.'s (2018) compounding advantage argument, but the small ECS sample ($N = 285$) limits inferential power. The result should be treated as an exploratory hypothesis requiring a dedicated larger-sample study for confirmation.

H3c – Discipline Heterogeneity: Confirmed and extended. The prediction that lab-intensive fields face the sharpest quality trade-offs is confirmed. Life and Health Sciences face the most severe Q1 decline ($\beta = -0.079$, $p < 0.001$) while Engineering, CS and Mathematics have the smallest decline

($\beta = -0.040$, $p = 0.086$), broadly consistent with Álvarez-Bornstein & Bordons' (2021) finding that discipline moderates funding-quality relationships. The specific mechanism, that severity of the quality trade-off increases with the length of the research-to-publication pipeline rather than with field norms about prestige, is a more precise claim than prior literature makes and represents an extension of the existing evidence.

H3d – Institution Type: Confirmed but contradicts the Matthew Effect prediction. The prediction that teaching-intensive scholars show larger gains is confirmed since they show the largest volume gains ($\beta = +1.358$, $p = 0.006$) and smallest quality sacrifice ($\beta = -0.05$, $p = 0.081$) relative to research-intensive scholars ($\beta = +0.722$, $p = 0.013$ for volume) and ($\beta = -0.073$, $p = 0.001$ for Q1 ratio).

However, this directly contradicts the conventional Matthew Effect prediction (Merton, 1968; Bol et al., 2018) that prestige amplifies funding returns. No prior study has directly tested this institutional dichotomy in a competitive grant context, and the result is one of the most policy-relevant novel contributions of the paper.

H3e – Gender: Not confirmed statistically. The prediction of Huang et al. (2020) and Álvarez-Bornstein & Bordons (2021) that female scholars show a larger positive productivity response is directionally consistent with the data ($\beta = +1.147$) but does not reach statistical significance ($p = 0.183$). With 40.7 percent of the sample unclassifiable due to Chinese-heritage naming conventions, the gender model is severely underpowered and cannot provide confirmatory evidence either way.

H – Research Diversification ($\beta = -0.033$, $p = 0.001$) and International Collaboration ($\beta = +0.058$, $p = 0.001$): The HHI decline indicating broadened research portfolios directly contradicts Li & Bai (2025), who found that NIH grant renewal reduced research diversity. The distinction is theoretically important as first-time grant receipt may unlock intellectual exploration that was previously resource-constrained, while renewal grants lock scholars into the trajectory that justified renewal. If so, this study and Li & Bai (2025) are documenting opposite ends of the same funding career arc rather than contradictory findings. The international co-authorship increase is consistent with Álvarez-Bornstein & Bordons' (2021) collaboration mediation mechanism and provides the most direct empirical support for the network-mediation hypothesis raised in Section 6.1.1.

7.0 – Conclusion

7.1 – Summary of Findings and Conclusions

This study set out to provide the first scholar-level causal evaluation of how RGC funding shapes the research productivity of STEM scholars in HK. Drawn on an eleven-year panel of 22,456 scholars and integrating PSM with TWFE, SUR and event-study analysis, the study isolates the marginal impact of competitive grants from the selection bias that channels funding toward already-productive researchers.

The central finding is a structural quality-quantity trade-off. RGC grant receipt significantly increases raw publication volume and short-term field-normalized citation impact, whilst simultaneously reducing the share of publications placed in top-tier Q1 journals and average journal prestige. These findings are robust across all alternative specifications, estimator diagnostics, and heterogeneity subgroups. The parallel trends assumption holds for the Q1 ratio, providing the formal statistical foundation for a strictly causal interpretation of the journal quality decline. Pre-trend violations for publication volume and short-term citations are theoretically informative rather than disqualifying, consistent with the Mertonian cumulative advantage mechanism, and the volume estimates should be read as upper bounds on the pure causal effect rather than unreliable. The study further identifies diminishing-returns thresholds of HK\$5.22 million for publication volume and HK\$3.97 million for citation quality, above which additional cumulative funding begins to erode marginal returns. The heterogeneity analysis reveals that whilst volume effects are concentrated in early-career scholars and teaching-intensive institutions, the quality trade-off is structurally uniform across all career stages, disciplines, institution types, and gender groups, identifying it as a property of grant accountability rather than the behavior of any scholar type.

These findings provide a plausible micro-level mechanism for the system-level phenomenon documented by Horta (2018), namely, the steady decline of HK's scientific competitiveness despite maintained public investment. A funding mechanism that systematically steers its most productive scholars toward lower-prestige journals to satisfy reporting deliverables will, over successive cohorts, generate exactly the kind of aggregate quality erosion captured in revealed comparative advantage indices. The concentration of RGC funding at research-intensive institutions, where marginal returns are the lowest, offers a structural explanation for why increased expenditure has not reversed this decline. These mechanisms are consistent with Horta's (2018) macro-level observations but confirming them as sufficient causal accounts requires direct comparative analysis against equivalent regional funding systems.

7.2 – Contributions to Knowledge

7.2.1 – Methodological Contributions

This study advances the methodological toolkit for funding evaluation by integrating PSM, TWFE, SUR and event-study analysis within a single cohesive framework, with each layer addressing a distinct threat to causal validity. The transparent reporting of the event-study pre-trend F-test, including honest disclosure of its violation for publication volume alongside its satisfaction for the Q1 ratio, models best practice in staggered DiD research, demonstrating that pre-trend violations can be theoretically informative rather than merely methodological failures (Sun & Abraham, 2021; de Chaisemartin & D'Haultfoeuille, 2020). The five-dataset architecture spanning 480,369 unique papers and 23.6 million citing-paper records also constitutes a data contribution that enables a level of covariate richness unavailable to prior HK-focused studies.

7.2.2 – Empirical Contributions

This study makes five novel empirical contributions.

1. This is the first causal estimate of career stage heterogeneity in RGC funding effects within HK's STEM system. Volume gains are concentrated in early-career scholars, but the quality trade-off is career-stage-invariant, meaning the strategic publication shift is structural rather than an early-career survival response.
2. The institution-type analysis provides the first empirical demonstration within a competitive grant context that the Matthew Effect's concentration of resources at high-prestige institutions can be misaligned with system-level efficiency. Teaching-intensive scholars show both larger volume gains and smaller quality sacrifices than their research-intensive peers.
3. The ECS-versus-GRF comparison produces the first bibliometric evidence that HK's early career funding architecture may operate against its stated purpose. ECS recipients experience a significantly sharper quality trade-off despite being the intended beneficiaries of the scheme's protection.
4. The research diversification finding directly contradicts Li & Bai (2025), who found that NIH grant renewal narrows research scope. First-time grant receipt in HK broadens it, establishing a distinction between the scope-broadening effects of initial grants and the trajectory-narrowing effects of renewal.
5. Causal estimates for the broader behavioral outcomes, including a six-percentage-point increase in international co-authorship and a non-significant effect on open-access behavior, characterize how competitive funding reshapes the academic repertoire beyond raw productivity metrics.

7.2.3 – Theoretical Contributions

Beyond confirmatory findings, the study makes three broader theoretical contributions.

1. It contributes to the research evaluation literature by providing causal evidence that, under competitive grant accountability pressure, journal prestige and citation-demonstrated scientific impact are systematically decoupled. Evaluation frameworks including the UK REF, Australia's ERA, and HK's RAE rely on journal tier as a quality proxy (Waltman & van Eck, 2015; Bornmann et al., 2020; University Grants Committee, 2018). This study demonstrates that doing so will structurally undervalue the scientific output generated during the most highly funded periods of a scholar's career.
2. It provides the first causal, quantitative validation of Tammi's (2009) impoverishment thesis at the individual scholar level. The consistent direction of the Q1 decline across every subgroup, with no detectable variation by career stage or gender, confirms that impoverishment dynamics are embedded in the accountability architecture itself rather than being a behavioral response.

3. Synthesizing the research diversification with Li & Bai (2025) generates a new theoretical distinction in that the funding-diversity relationship is trajectory-dependent rather than directionally fixed. Initial grants enable intellectual exploration and renewal grants lock scholars into their productive agenda. These dynamics describe different phases of the same funding career arc rather than being contradictory findings.

7.3 – Future Research Directions

Three priorities follow directly from the limitations of this study.

1. A dedicated ECS panel study exploiting a regression discontinuity in ECS funding cutoff scores, using administrative records that scheme administrators already hold, would move the ECS-versus-GRF comparison from exploratory to confirmatory and test whether institutional co-investment requirements moderate the quality trade-offs currently imposed on early-career scholars.
2. A comparative replication across small open research economies such as Singapore or South Korea's National Research Foundation (NRF) and Taiwan's National Science and Technology Council (NSTC), which represent the most directly comparable systems, would help determine whether F^* thresholds and institution-type reversal identified here are general structure properties of competitive grant mechanisms or idiosyncrasies of the HK context.
3. A co-authorship graph analysis linking grant receipt to collaboration network growth and subsequently to citation accumulation would test the network-mediation hypothesis raised in Section 6.1.1, specifically, whether the short-term citation premium reflects intrinsic quality improvement or acquired social capital, a distinction with direct implications for how citation metrics should be interpreted in funded versus unfunded research evaluation contexts.

7.4 – Policy Recommendations

The paradox that opened this study is now answerable at the scholar level. HK's competitive research funding mechanism is not broken, but its accountability architecture systematically redirects the effort it finances. Competitive public funding for research is not merely an input to the production function of science but is an institutional signal that shapes the incentive environment within which scholars choose what to study, how to write it up, and where to submit it. The evidence suggests that the RGC's current mechanisms are broadly effective at stimulating output but adds pressure on scholars to publish in lower-tier journals. The scholar most likely to convert additional investment into genuine productivity gains is early-career, based at a resource-constrained teaching-intensive institution, and working in a fast-turnaround discipline. This is precisely the profile that current selection criteria systematically pass over in favor of established researchers with strong prior track records. Targeted reforms in accountability design, including field-calibrated performance metrics, progressive allocation adjustments for scholars approaching cumulative funding thresholds, and enhanced early-career support, are the most evidence-backed

levers for maximizing the return on HK's substantial investment in competitive research funding whilst preserving the scientific quality that ultimately determines its long-run research competitiveness.

References

- Álvarez-Bornstein, B., & Bordons, M. (2021). Is funding related to higher research impact? Exploring its relationship and the mediating role of collaboration in several disciplines. *Journal of Informetrics*, 15(1), 101102. <https://doi.org/10.1016/j.joi.2020.101102>
- Austin, P. C. (2009). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine*, 28(25), 3083–3107. <https://doi.org/10.1002/sim.3697>
- Bol, T., de Vaan, M., & van de Rijt, A. (2018). The Matthew effect in science funding. *Proceedings of the National Academy of Sciences*, 115(19), 4887–4890. <https://doi.org/10.1073/pnas.1719557115>
- Bornmann, L., Haunschild, R., & Mutz, R. (2020). Should citations be field-normalized in evaluative bibliometrics? An empirical analysis based on propensity score matching. *Journal of Informetrics*, 14(4), 101098. <https://doi.org/10.1016/j.joi.2020.101098>
- Boyack, K. W., & Jordan, P. (2011). Metrics associated with NIH funding: A high-level view. *Journal of the American Medical Informatics Association*, 18(4), 423–431. <https://doi.org/10.1136/amiajnl-2011-000213>
- Breusch, T. S., & Pagan, A. R. (1980). The Lagrange Multiplier Test and Its Application to Model Specification in Econometrics. *Review of Economic Studies*, 47, 239–253. <https://doi.org/10.2307/2297111>
- Cameron, A. C., & Miller, D. L. (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources*, 50(2), 317–372. <https://doi.org/10.3368/jhr.50.2.317>
- Census and Statistics Department. (2024). *Statistics by Subject: Health*. Census and Statistics Department. <https://www.censtatd.gov.hk/en/scode580.html>

- Corsini, A., & Pezzoni, M. (2023). Does grant funding foster research impact? Evidence from France. *Journal of Informetrics*, 17(4), 101448. <https://doi.org/10.1016/j.joi.2023.101448>
- de Chaisemartin, C., & D'Haultfoeuille, X. (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review*, 110(9), 2964–2996. <https://doi.org/10.1257/aer.20181169>
- Egghe, L. (2006). Theory and practise of the g-index. *Scientometrics*, 69(1), 131–152. <https://doi.org/10.1007/s11192-006-0144-7>
- Elwert, F., & Winship, C. (2014). Endogenous selection bias: The problem of conditioning on a collider variable. *Annual Review of Sociology*, 40, 31–53. <https://doi.org/10.1146/annurev-soc-071913-043455>
- Fu, J., Liu, X., Zhang, C., & Li, J. (2025). Unfinished grants, unending progress: The impact of unfinished research grants on scientific innovation. *Journal of Informetrics*, 19(4). <https://doi.org/10.1016/j.joi.2025.101734>
- González-Pereira, B., Guerrero-Bote, V. P., & Moya-Anegón, F. (2010). A new approach to the metric of journals' scientific prestige: The SJR indicator. *Journal of Informetrics*, 4(3), 379–391. <https://doi.org/10.1016/j.joi.2010.03.002>
- Hausman, J. A. (1978). Specification Tests in Econometrics. *Econometrica*, 46(6), 1251–1271. JSTOR. <https://doi.org/10.2307/1913827>
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102(46), 16569–16572. <https://doi.org/10.1073/pnas.0507655102>
- Horta, H. (2018). The declining scientific wealth of Hong Kong and Singapore. *Scientometrics*, 117(1), 427–447. <https://doi.org/10.1007/s11192-018-2845-0>

- Horta, H., Cattaneo, M., & Meoli, M. (2018). PhD funding as a determinant of PhD and career research performance. *Studies in Higher Education*, 43(3), 542–570.
<https://doi.org/10.1080/03075079.2016.1185406>
- Huang, J., Gates, A. J., Sinatra, R., & Barabási, A.-L. (2020). Historical comparison of gender inequality in scientific careers across countries and disciplines. *Proceedings of the National Academy of Sciences*, 117(9), 4609–4616.
<https://doi.org/10.1073/pnas.1914221117>
- Jacob, B. A., & Lefgren, L. (2011). The impact of research grant funding on scientific productivity. *Special Issue: The Role of Firms in Tax Systems*, 95(9), 1168–1177.
<https://doi.org/10.1016/j.jpubeco.2011.05.005>
- Kang, J. D. Y., & Schafer, J. L. (2007). Demystifying Double Robustness: A Comparison of Alternative Strategies for Estimating a Population Mean from Incomplete Data. *Statistical Science*, 22(4), 523–539. JSTOR.
- Kwan, P., Johnston, J., Fung, A. Y., Chong, D. S., Collins, R. A., & Lo, S. V. (2007). A systematic evaluation of payback of publicly funded health and health services research in Hong Kong. *BMC Health Services Research*, 7(1), 121. <https://doi.org/10.1186/1472-6963-7-121>
- Kyvik, S., & Aksnes, D. W. (2015). Explaining the increase in publication productivity among academic staff: A generational perspective. *Studies in Higher Education*, 40(8), 1438–1453. <https://doi.org/10.1080/03075079.2015.1060711>
- Lee, Y.-H. (2021). Determinants of research productivity in Korean Universities: The role of research funding. *The Journal of Technology Transfer*, 46(5), 1462–1486.
<https://doi.org/10.1007/s10961-020-09817-2>

- Leydesdorff, L., & Bornmann, L. (2011). How fractional counting of citations affects the impact factor: Normalization in terms of differences in citation potentials among fields of science. *Journal of the American Society for Information Science and Technology*, 62(2), 217–229. <https://doi.org/10.1002/asi.21450>
- Li, B., & Bai, A. (2025). The influence of grant renewal on research content: Evidence from NIH-funded PIs. *Scientometrics*, 130(5), 2617–2638. <https://doi.org/10.1007/s11192-025-05295-1>
- Merton, R. K. (1968). The Matthew Effect in Science. *Science*, 159(3810), 56–63. <https://doi.org/10.1126/science.159.3810.56>
- National Institutes of Health. (2026). *Success Rates | NIH Budget and Spending Data (RePORT)*. National Institutes of Health. <https://report.nih.gov/funding/nih-budget-and-spending-data-past-fiscal-years/success-rates>
- Nygaard, L. P., Piro, F. N., & Aksnes, D. W. (2022). Gendering excellence through research productivity indicators. *Gender and Education*, 34(6), 690–704. <https://doi.org/10.1080/09540253.2022.2041186>
- Ou, G., Zhao, K., Zuo, R., & Wu, J. (2024). Effects of research funding on the academic impact and societal visibility of scientific research. *Journal of Informetrics*, 18(4), 101592. <https://doi.org/10.1016/j.joi.2024.101592>
- Payne, A. A., & Siow, A. (2003). Does Federal Research Funding Increase University Research Output? *The B.E. Journal of Economic Analysis & Policy*, 3(1), 1–24. <https://doi.org/10.2202/1538-0637.1018>

- Rosenbaum, P. R., & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70(1), 41–55. JSTOR.
<https://doi.org/10.2307/2335942>
- Schreiber, M. (2008). A modification of the h-index: The hm-index accounts for multi-authored manuscripts. *Journal of Informetrics*, 2(3), 211–216.
<https://doi.org/10.1016/j.joi.2008.05.001>
- Stephan, P. (2012). *The economics of science*. Harvard University Press.
- Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society Interface*, 4(15), 707–719.
<https://doi.org/10.1098/rsif.2007.0213>
- Sum, N.-L., & Jessop, B. (2013). Competitiveness, the Knowledge-Based Economy and Higher Education. *Journal of the Knowledge Economy*, 4(1), 24–44.
<https://doi.org/10.1007/s13132-012-0121-8>
- Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2), 175–199.
<https://doi.org/10.1016/j.jeconom.2020.09.006>
- Tammi, T. (2009). The competitive funding of university research: The case of Finnish science universities. *Higher Education*, 57(5), 657–679. <https://doi.org/10.1007/s10734-008-9169-6>
- Times Higher Education. (2025). *World University Rankings 2025*. Times Higher Education.
<https://www.timeshighereducation.com/world-university-rankings/2025/world-ranking>

UNESCO Institute for Statistics. (2026). *SDG 9.5.1: Research and Development Expenditure as a Proportion of GDP* [Dataset]. UNESCO Institute for Statistics (UIS).

<https://databrowser.uis.unesco.org/>

University Grants Committee. (2018). *Research Assessment Exercise 2020: General Panel Guidelines*. University Grants Committee.

https://www.ugc.edu.hk/doc/eng/ugc/rae/2020/gpg_sep18.pdf

University Grants Committee. (2024). *Research Grants Council announces funding results of General Research Fund and Early Career Scheme 2024/25*. University Grants Committee.

https://www.ugc.edu.hk/eng/ugc/about/press_speech_other/press/2024/pr20240627.html

University Grants Committee. (2025a). *Early Career Scheme 2026/27: Scheme Overview and Guidance Notes (ECS2)*. University Grants Committee.

<https://www.ugc.edu.hk/doc/eng/rgc/form/ECS2.pdf>

University Grants Committee. (2025b). *UGC announces new initiatives to support higher education in Hong Kong*. University Grants Committee.

https://www.ugc.edu.hk/eng/ugc/about/press_speech_other/press/2025/pr20250627a.html

Waltman, L., & van Eck, N. J. (2015). Field-normalized citation impact indicators and the choice of an appropriate counting method. *Journal of Informetrics*, 9(4), 872–894.

<https://doi.org/10.1016/j.joi.2015.08.001>

Wang, J. J., Wei, S., & Ye, F. Y. (2025). How research funding shapes academic outputs: Evidence from communication research paper characteristics and thematic trends in China. *Journal of Data and Information Science*, 10(4), 1–24.

<https://doi.org/10.2478/jdis-2025-0040>

- Wang, Y., Jones, B. F., & Wang, D. (2019). Early-career setback and future career impact. *Nature Communications*, 10(1), 4331. <https://doi.org/10.1038/s41467-019-12189-3>
- Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press. JSTOR. <http://www.jstor.org/stable/j.ctt5hhcfr>
- Zellner, A. (1962). An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of the American Statistical Association*, 57(298), 348–368. JSTOR. <https://doi.org/10.2307/2281644>