Information Research - Vol. 30 No. iConf (2025)

The promoting effect of internet use on the physical health of Chinese older adults: an empirical study based on the difference-in-differences model

Haoyuan Sun, Zhenkang Fu, Zhengtong Pu, and Qinghua Zhu
DOI: https://doi.org/10.47989/ir30iConf46894

Abstract

Introduction. The widespread adoption of information technology (IT) has had a profound impact on economic and social development. With China confronting an increasingly serious aging population issue, it becomes critical to examine the effects of Internet use on the physical health of older adults.

Method. This study seeks to investigate the impact of Internet use on older adults' physical health by employing a pooled cross-sectional difference-in-differences (DID) model, drawing on data from the China General Social Survey (CGSS) covering the years 2012 to 2021.

Conclusion. The results demonstrate that Internet use significantly enhances the physical health of older adults, with these findings proving robust. Mechanism tests further reveal that Internet use contributes to improved physical health by increasing the frequency of physical exercise. Additionally, heterogeneity analysis indicates that the positive effects of Internet use are more pronounced among older adults in economically developed regions and those with poor self-perceived health.

Introduction

With the swift progression of China's economy and ongoing social changes, the issue of population aging has become increasingly significant. Data from the National Bureau of Statistics (2024) shows that the number of individuals aged 60 and over in China has reached 297 million, making it one of the fastest-aging nations in the world. As people grow older, the overall health of the older adults tends to deteriorate. It is estimated that more than 190 million older adults in China are affected by chronic illnesses, with over 40 million experiencing full or partial disability (Ma, 2022). Addressing the health concerns of this large older adult population has therefore become a pressing social challenge. However, in the short term, current healthcare resources are inadequate to meet the substantial medical needs of this demographic. The ***Healthy China Plan 2030***, a key national health strategy, has consistently highlighted the importance of information technology in providing support and guidance. As a result, the Internet has emerged as a promising approach to help address the physical health challenges faced by older adults.

The Internet has not only profoundly transformed the lifestyles of older adults but has also had a far-reaching and significant impact on their physical health. While some people worry that using the Internet may have a negative impact on the health of the older adults by making their living habits worse (Gilleard et al., 2007), numerous studies have refuted this notion. Research demonstrates that Internet use does not reduce the amount of time older adults dedicate to exercise; on the contrary, it significantly increases their likelihood and frequency of engaging in physical activities(Guo et al., 2022; Sasaki et al., 2022). Moreover, older adults can access health-related information online, which enhances their health awareness and self-management abilities (Ang et al., 2021; Han & Zhao, 2021). From a methodological perspective, most current studies adopt an internal view of information behaviour, utilizing models such as ordinary least squares (OLS) regression to quantify the promoting effects of Internet use on the physical health of older adults. However, this approach has certain limitations: first, it tends to overlook external factors such as the implementation of national macro policies, the widespread availability of Internet infrastructure, and changes in the channels through which the Chinese population accesses information. Second, issues related to endogeneity between variables may compromise the accuracy and reliability of the results.

In addition to examining the positive effects of Internet use on the health of older adults and the underlying mechanisms involved, this paper seeks to address the limitations of previous research. To account for external influencing factors and resolve the issue of endogeneity between variables, this paper employs a difference-in-differences (DID) model. Furthermore, to address the challenge of small sample sizes, this study integrates six years of statistical data into pooled cross-sectional data. This approach not only increases the sample size but also allows for a more nuanced examination of the similarities and differences in the impact of Internet use on the physical health of older adults over different periods. Specifically, this study modifies the classic DID model to construct a pooled cross-sectional DID model, supplemented by a series of robustness tests to ensure the scientific validity of the conclusions.

Theoretical hypotheses

Social support theory emphasizes how interactions within social networks promote physical health by providing informational and instrumental support (Cohen & Wills, 1985). Older adults can access knowledge about disease prevention, healthy eating, and exercise through the Internet, and they can use health-monitoring devices or online medical services to enhance their ability to manage their health. At the same time, the core of self-determination theory lies in the influence of autonomy, competence, and relatedness on an individual's behavioural choices (Deci & Ryan, 2000). The Internet enables older adults to autonomously choose their health management methods, such as personalized exercise plans and dietary programs, thereby enhancing their sense of autonomy. The use of smart devices to monitor health data further strengthens their sense of competence. By joining health-related interest communities, they can receive encouragement, which promotes the persistence of long-term healthy behaviour, ultimately contributing to improved physical health.

Hypothesis 1: internet use behaviour positively impacts the physical health of the older adults.

Hypothesis 2: internet use behaviour promotes the physical health of the older adults through the path of increasing exercise frequency.

Research design

Selection of External Impact Nodes

From a policy perspective, the ***Healthy China Plan 2030***, released by the Chinese government in 2016, serves as a top-level design for the health sector in China. It clearly states that the health issues of older adults should receive significant attention in the process of promoting national health. The policy particularly emphasizes the role of information technology as a vital tool in addressing the health challenges faced by older adults. From the perspective of Internet accessibility, official statistics indicate that in the second half of 2015, China's Internet penetration rate exceeded 50% for the first time (China Internet Network Information Center, 2016), marking the widespread adoption of digital technology across various sectors of society. Meanwhile, the ways in which older adults access information underwent significant changes during this period (as shown in Figure 1). Compared to traditional channels such as newspapers, magazines, and radio, the proportion of older adults accessing information via the Internet steadily increased, particularly between 2015 and 2017, when the proportion of older adults using the Internet as their primary information source saw a significant rise. Against this background, this paper identifies 2016 as a critical point when information technology began to exert an external influence on the information acquisition behaviour of older adults. It further explores the potential impact on older adults' health management and the implementation of related policies.

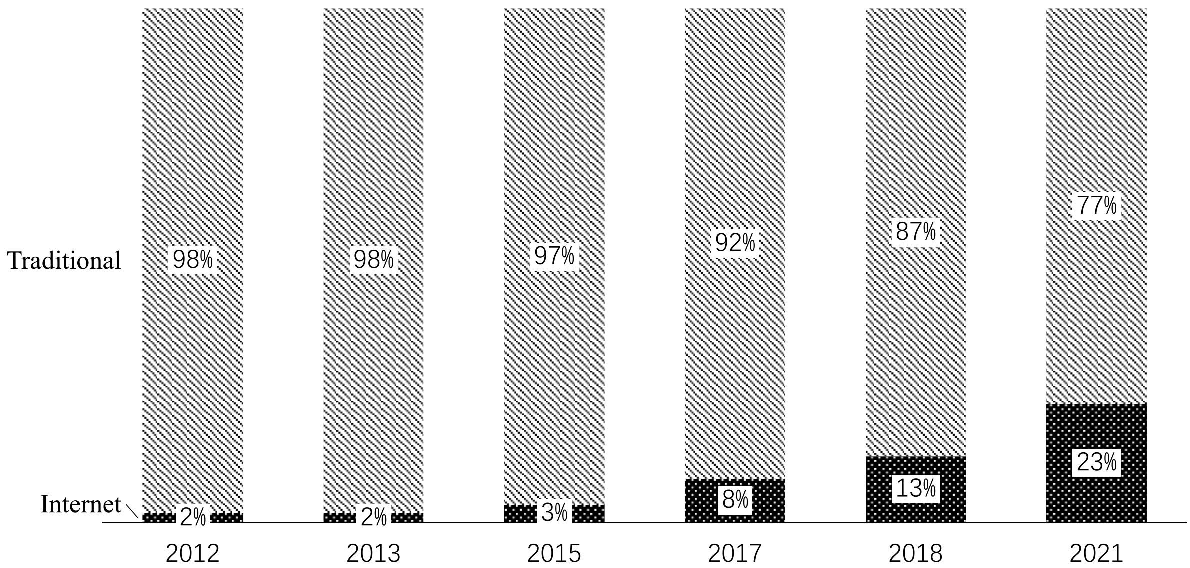

Figure 1. Information access channels for Chinese older adults (Source: The data are collected from CGSS database)

Construction of model

CGSS started in 2003 is a nationwide, comprehensive and continuous academic survey project by Renmin University of China. As one of the most important social surveys in China, CGSS has highly representative and widely applicable data. The CGSS data used in this paper are pooled cross-sectional data, observing different research subjects at different points in time. Therefore, this paper modify the classical DID model to construct a pooled cross-sectional DID model (Kiel & Mcclain, 1995; Yang et al., 2024). Formula (1) considers the time fixed effect based on the pooled cross-sectional DID model. Here, Yrepresents the dependent variable, the physical health of the older adults, with subscripts i and t representing the individual and the year, respectively. $treat_i \times post_t$ is the core explanatory variable, the interaction term between the post-policy dummy variable and the treatment group dummy variable. $treat_i$is the treatment group dummy variable, assigned a value of 1 for individuals with Internet use behavior, and 0 otherwise. Meanwhile, $post_t$is the time dummy variable. The external impact node selected in this paper is the year the ***Healthy China Plan*** (HCP) was issued, with post 2016 assigned a value of 1, and 0 otherwise. X_{it} represents other control variables, including individual and family-level variables. γ_t represents the time fixed effect. Since the data selected in this paper are pooled cross-sectional data, observing different individuals at different points in time, individual fixed effects are not considered. ε_{it} is the random disturbance term of the model. The estimated value of the coefficient. β_1 is the focus of this paper. If hypothesis 1 holds, the coefficient should be significantly positive.

$$Y_{it} = \beta_0 + \beta_1 treat_i \times post_t + \alpha X_{it} + \gamma_t + \varepsilon_{it} \qquad (1)$$

This study draws on data from the China General Social Survey (CGSS) for six selected years: 2012, 2013, 2015, 2017, 2018, and 2021, to construct the research model. Individuals aged 60 and above from each survey year are chosen as the study subjects, with samples containing outliers or missing values in key variables excluded from the analysis. The dependent variable is determined by the question: '*In the past week, has your health affected your ability to work*?' This variable is rated on a five-point scale: always, often, sometimes, rarely, and never, with corresponding values from 1 to 5. '*Always*' is assigned a value of 1, while '*Never*' is assigned a value of 5. The core explanatory variable, DID, represents the interaction between the post-policy dummy variable and the treatment group dummy variable. Control variables encompass individual and family-level characteristics, as outlined in Table 1.

Variables	N	Mean	Median	S.D.	Min	Max
Physical	21244	3.480	4	1.212	1	5
DID	21244	0.181	0	0.385	0	1
Gender	21244	1.503	2	0.500	1	2
Education	21244	3.537	3	2.432	1	13
Registration	21244	1.852	1	1.181	1	8
Matrimony	21244	3.999	3	1.759	1	7
Children	21244	2.549	2	1.484	0	24
Status	21244	4.176	4	1.769	1	10
Familyeco	21244	2.551	3	0.770	1	5

Table 1. Descriptive statistical table.

Empirical tests

Baseline regression

Column (1) of Table 2 demonstrates that Internet use exerts a significant positive effect on the physical health of older adults when control variables are not included. Once control variables are introduced, the coefficient of the core explanatory variable decreases slightly but remains positive, highlighting the robustness of the positive influence of Internet use. In column (3), after accounting for time-fixed effects, the coefficient of the core explanatory variable continues to be significantly positive, with statistical significance at the 1% level. Overall, these results provide support for Hypothesis 1.

Variables	(1) Physical	(2) Physical	(3) Physical
DID	0.536***	0.260***	0.289***
	(26.24)	(12.16)	(11.70)
Gender		-0.122***	-0.121***
		(-7.49)	(-7.41)
Education		0.042***	0.042***
		(11.23)	(11.09)
Registration		0.068***	0.068***
		(9.20)	(9.22)
Matrimony		-0.010**	-0.010**
		(-2.04)	(-2.02)
Children		-0.080***	-0.080***
		(-13.64)	(-13.67)
Status		0.065***	0.065***
		(12.20)	(12.29)
Familyeco		0.228***	0.225***
		(18.37)	(18.02)
Constant	3.383***	2.728***	2.739***
	(369.39)	(59.85)	(55.45)
N	21,244	21,244	21,244
R-squared	0.029	0.119	0.121
Year FE	NO	NO	YES

Notes: (1) Robust t-statistics in parentheses; (2) *** p<0.01, ** p<0.05, * p<0.1

Table 2. Baseline regression results.

Robustness tests

Parallel trend test

This paper follows the methodology of Liu and Qiu (2016) to examine the trend of changes between the treatment and control groups. The results shown in Figure 2 indicate that, within the 99% confidence interval, all regression outcomes prior to the impact point are insignificant, suggesting that the treatment and control groups experienced similar trends before 2016. After the impact point, the physical health of older adults in the treatment group showed a significant improvement. However, five years later, in 2021, the reduced activity range of older adults due to the pandemic negatively impacted their health, causing the parallel trends test to fail for that year. Despite this, the sample used for model construction overall passes the parallel trend test

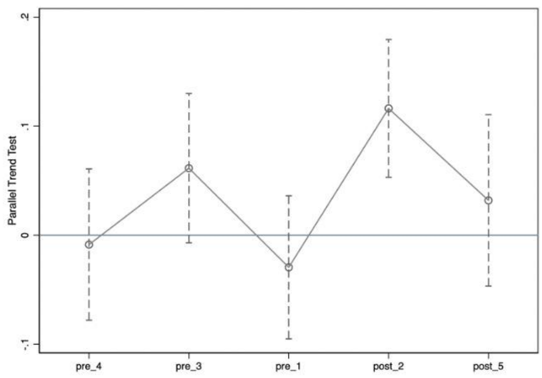

Figure 2. Parallel trend test results.

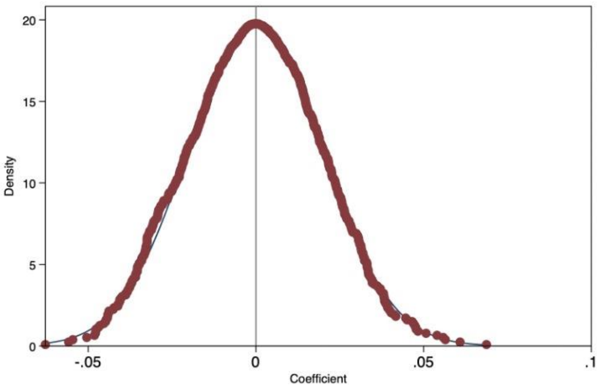

Figure 3. Placebo test results.

Placebo Test

To mitigate the potential influence of omitted variables on the estimation results, this paper adopts the approach of Cai (2016) , randomly selecting variables from the entire sample to form a pseudo treatment group, while assigning the remaining samples to a pseudo control group. Figure 3 presents the estimated coefficient kernel density, and the corresponding p-value distribution based on 1,000 random selections. The findings indicate that the average regression coefficient is close to zero, with the distribution of the treatment effect centered near zero, forming a symmetrical bell-shaped curve. These results suggest that omitted variables do not significantly bias the estimation, further confirming the robustness of the baseline regression outcomes.

PSM-DID

To address the issue of endogeneity in the sample, this paper employs propensity score matching (PSM) before conducting the difference-in-differences estimation. Following the approaches of Böckerman (2009) and Yi (2024), periodic propensity score matching is applied to the relevant data, using the same covariates for matching as the control variables in the baseline regression. The regression results are presented in Table 3. The first column of the table represents 1:1 nearest neighbor matching without replacement, while columns (2) through (4) show 1:1 to 1:3 nearest neighbor matching with replacement. Columns (5) to (7) present 1:1 to 1:3 caliper matching, and column (8) reports kernel matching. After performing propensity score matching to mitigate endogeneity, the core explanatory variables remain significantly positive, confirming the robustness of Hypothesis 1.

	(1) Nearest	(2) Nearest	(3) Nearest	(4) Nearest	(5) Caliper	(6) Caliper	(7) Caliper	(8) Kernel	(9) Repeated
DID	0.338***	0.325***	0.312***	0.302***	0.299***	0.307***	0.296***	0.288***	0.289***
	(6.76)	(7.86)	(9.48)	(10.11)	(9.49)	(10.42)	(10.52)	(11.63)	(12.00)
Constant	3.278***	2.985***	3.006***	3.033***	2.927***	2.986***	3.006***	2.753***	2.739***
	(31.57)	(27.73)	(35.23)	(39.26)	(34.09)	(38.30)	(40.98)	(55.14)	(58.00)
Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	6,554	5,146	8,100	9,992	8,833	10,260	11,443	20,993	21,244
R-squared	0.117	0.080	0.087	0.089	0.087	0.090	0.091	0.119	0.121

Notes: (1) Robust t-statistics in parentheses; (2) *** p<0.01, ** p<0.05, * p<0.1

Table 3. PSM-DID and repeated random sampling.

Repeated random sampling

To address the issue of sample selection bias, this paper applies the Bootstrap method for repeated random sampling, following the approach of Li (2021), with a sample size of 10,000 and 500 random extractions. The empirical results in column (4) of Table 3 show that the outcomes from repeated random sampling are largely consistent with the baseline regression results, further reinforcing the robustness of the conclusions.

Mechanism tests

Increase the frequency of exercise

The Internet can enhance the frequency of exercise among older adults, which in turn impacts their physical health. To examine this pathway, a mediation effect model is employed. Columns (1)-(3) of Table 4 present the results: (1) shows the baseline regression, (2) indicates that the variable representing the frequency of physical exercise in older adults is significantly positive, and (3) demonstrates that both the core explanatory variable and the mediation variable are significantly positive, with a reduction in the coefficient of the core explanatory variable, providing preliminary evidence for the existence of this pathway. Additionally, the Bootstrap test and Sobel test are conducted to validate the mediation effect. According to Table 5, the confidence interval for the indirect effect does not include zero, confirming the presence of the mediation effect. Moreover, the Sobel test returns a p-value of 0.001, significant at the 1% level, further establishing the mediation effect and supporting Hypothesis 2.

	(1) Physical	(2) Exercise	(3) Physical
DID	0.289***	0.410***	0.244***
	(11.70)	(11.66)	(9.97)
Exercise			0.112***
			(21.67)
Constant	2.739***	0.927***	2.632***
	(55.45)	(14.51)	(53.82)
Control	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
N	21,244	21,201	21,201
R-squared	0.121	0.181	0.141
Sobel Test (P Value)		0.000***	

Notes: (1) Robust t-statistics in parentheses; (2) *** p<0.01, ** p<0.05, * p<0.1

Table 4. Mediation effect results.

Effect	Coef.	Std. Err.	z	P>|z|	Normal Based [95% Conf. Interval]	
Indirect	0.055	0.004	12.990	0.000***	0.047	0.063
Direct	0.205	0.021	9.890	0.000***	0.165	0.246
Total	0.260	0.021	12.550	0.000***	0.220	0.301

Notes: (1) *** p<0.01, ** p<0.05, * p<0.1

Table 5. Bootstrap test result.

Heterogeneity analysis

Economic development heterogeneity

Based on the GDP levels of various regions in 2016, this paper divides the sample into economically developed areas and economically less developed areas. The results, presented in columns (1) and (2) of Table 6, indicate that the positive effect of Internet use on the physical health of older adults is more pronounced in economically developed regions compared to less developed ones. This difference may be attributed to the limited availability of exercise facilities and equipment for older adults in less developed areas.

	(1) Developed	(2) Less Developed	(3) Good	(4) Poor
DID	0.310***	0.227***	0.091***	0.318***
	(10.77)	(4.70)	(3.42)	(9.52)
Constant	2.874***	2.491***	4.037***	2.375***
	(48.76)	(27.74)	(70.24)	(39.93)
Control	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	14,627	6,581	8,215	13,019
R-squared	0.120	0.114	0.050	0.115
Inter-group Difference Test	0.071*		0.000***	

Notes: (1) Robust t-statistics in parentheses; (2) *** p<0.01, ** p<0.05, * p<0.1

Table 6. Heterogeneity analysis results.

Self- perceived health heterogeneity

Older adults exhibit varying perceptions of their own health. This section therefore examines how internet use affects physical health across different self-perceived health groups. Accordingly, the sample is divided into two groups: those with good self-perceived health and those with poor self-perceived health. The results, as shown in columns (3) and (4) of Table 6, indicate that internet use has a significantly stronger positive effect on the group with poor self-perceived health. These findings provide important insights for future policymaking: for older adults with chronic illnesses or disabilities, internet use proves to be an effective means of improving physical well-being.

Conclusion

With the rapid development of China's economy and the growing challenges posed by an aging population, physical health issues among older adults have become increasingly prominent. In this context, addressing these health concerns requires effective intervention measures. The expansion of Internet access offers a promising solution to this problem. This paper aims to examine the impact of Internet use on the physical health of older adults, providing a scientific foundation for the formulation of relevant policies and interventions. Grounded in social support theory and self-determination theory, the study employs a pooled cross-sectional difference-in-differences (DID) model, using data from the CGSS database spanning 2012 to 2021.

The findings indicate that Internet use has a significantly positive impact on the physical health of older adults, and the results remain robust. Furthermore, the mechanism analysis shows that Internet use enhances their physical health by increasing the frequency of exercise. The heterogeneity analysis reveals that Internet use has a stronger positive effect on older adults in economically developed regions and those with poor self-perceived health.

Future research should further investigate how different types of Internet usage behaviour affect the physical health of older adults to provide more comprehensive empirical evidence for policy development. Additionally, attention should be given to improving access to exercise facilities and equipment in economically less developed areas. Through continued research and policy refinement, this paper aims to further enhance the physical health of older adults and contribute to the realization of the Healthy China initiative. Such efforts will not only improve the quality of life for older adults but also significantly contribute to social harmony and stability.

Acknowledgements

This work is supported by the National Social Science Foundation of China (No.22&ZD327).

About the authors

Haoyuan Sun is a master student at the School of Information Management, Nanjing University, Nanjing, China. His research interests are in health informatics and digital humanities. He can be contacted at haoyuan.sun@smail.nju.edu.cn.

Zhenkang Fu is a Doctoral candidate at the School of Information Management, Nanjing University, Nanjing, China. His research interests are in technology and innovation management. He can be contacted at fuzhenkang@163.com.

Zhengtong Pu is a master student at the School of Information Management, Nanjing University, Nanjing, China. His research interests are in health informatics and digital humanities. He can be contacted at ztpu@smail.nju.edu.cn.

Qinghua Zhu is a Professor and Doctoral supervisor at the School of Information Management, Nanjing University, Nanjing, China. His research interests are in user information behavior, smart elderly care, and health informatics. He can be contacted at qhzhu@nju.edu.cn.

References

Ang, S., Lim, E., & Malhotra, R. (2021). Health-Related Difficulty in Internet Use Among Older Adults: Correlates and Mediation of Its Association With Quality of Life Through Social Support Networks. GERONTOLOGIST, 61(5), 693–702. https://doi.org/10.1093/geront/gnaa096

Böckerman, P., & Ilmakunnas, P. (2009). Unemployment and self-assessed health: Evidence from panel data. Health Economics, 18(2), 161–179. Scopus. https://doi.org/10.1002/hec.1361

Cai, X., Lu, Y., Wu, M., & Yu, L. (2016). Does environmental regulation drive away inbound foreign direct investment? Evidence from a quasi-natural experiment in China. Journal of Development Economics, 123, 73–85. https://doi.org/10.1016/j.jdeveco.2016.08.003

Cohen, S., & Wills, T. (1985). Stress, Social Support, and the Buffering Hypothesis. PSYCHOLOGICAL BULLETIN, 98(2), 310–357. https://doi.org/10.1037//0033-2909.98.2.310

Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. PSYCHOLOGICAL INQUIRY, 11(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Gilleard, C., Hyde, M., & Higgs, P. (2007). Community and communication in the third age: The impact of Internet and cell phone use on attachment to place in later life in England. JOURNALS OF GERONTOLOGY SERIES B-PSYCHOLOGICAL SCIENCES AND SOCIAL SCIENCES, 62(4), S276–S283. https://doi.org/10.1093/geronb/62.4.S276

Guo, B., Zhang, X., Zhang, R., & Chen, G. (2022). The Association between Internet Use and Physical Exercise among Middle-Aged and Older Adults-Evidence from China. INTERNATIONAL JOURNAL OF ENVIRONMENTAL RESEARCH AND PUBLIC HEALTH, 19(24), 16401. https://doi.org/10.3390/ijerph192416401

Han, J., & Zhao, X. (2021). Impact of Internet Use on Multi-dimensional Health: An Empirical Study Based on CGSS 2017 Data. Frontiers in Public Health, 9. https://doi.org/10.3389/fpubh.2021.749816

Kiel, K., & Mcclain, K. (1995). House Prices During Siting Decision Stages—The Case of an Incinerator from Rumor Through Operation. JOURNAL OF ENVIRONMENTAL ECONOMICS AND MANAGEMENT, 28(2), 241–255. https://doi.org/10.1006/jeem.1995.1016

Li, Q. & Zhang, Y. (2021). Financial openness and resource allocation efficiency: Evidence from foreign banks entering China. Chinese Industrial Economy, 5, 95–113. https://doi.org/10.19581/j.cnki.ciejournal.2021.05.016

Liu, Q., & Qiu, L. D. (2016). Intermediate input imports and innovations: Evidence from Chinese firms' patent filings. Journal of International Economics, 103, 166–183. https://doi.org/10.1016/j.jinteco.2016.09.009

Sasaki, S., Sato, A., Tanabe, Y., Matsuoka, S., Adachi, A., Kayano, T., Yamazaki, H., Matsuno, Y., Nakano, A., & Watanabe, T. (2022). Internet use and physical activity of older adults during the COVID-19 pandemic: A cross-sectional study in a northern Japanese City. BMC GERIATRICS, 22(1), 688. https://doi.org/10.1186/s12877-022-03360-5

Yang, H., Zhang, Y., & Liu, L. (2024). Patent priority examination under mission orientation: Can speed and quality coexist? Technology in Society, 76, 102452. https://doi.org/10.1016/j.techsoc.2023.102452

Yi, Y., Zhang, L., Du, L., & Sun, H. (2024). Cross-regional integration of renewable energy and corporate carbon emissions: Evidence from China's cross-regional surplus renewable energy spot trading pilot. Energy Economics, 135, 107649. https://doi.org/10.1016/j.eneco.2024.107649