Printing and Women
The Gendered Impact of Printing Technology in China

Nina Liu*

August 24, 2025

Abstract

Economists have long argued that technologies can exhibit gender-biased impacts, often focusing on gender-specific innovations such as plowing and weaving. This study shifts the focus to woodblock printing—a seemingly gender-neutral technology—and examines its long-term gendered impact on educational outcomes in China. Combining historical data on poet presence with cohort-level literacy data from the 1982 Census, I find that women's educational outcomes derived greater marginal benefits than men's from increased availability of woodblock-printed books, a pattern that persisted until the onset of Mao's Mass Education Reform in 1949. To address potential endogeneity, I employ an instrument variable strategy that exploits river distance to locations suitable for bamboo cultivation (a crucial raw material for printing paper) as an exogenous source of variation in woodblock printing availability. I attribute the historical root of female-biased impact to the substitutive role of woodblock printing for formal schooling, from which women were largely excluded during the Imperial Period. Furthermore, triple-difference analysis suggests that the persistence of this effect can be partly explained by woodblock printing's role in altering cultural norms regarding investment in daughters' education. Overall, the findings underscore the importance of institutions in shaping the effects of technology, and vice versa.

Keywords: Technology, gender, education, institutions

*Nina, Ningxi Liu, PhD candidate at King's College London. Email: nina.liu@kcl.ac.uk. I am deeply grateful to my supervisors, Liang Bai and Seyhun Sakalli, for their unwavering guidance and devotion to this project. For their invaluable feedback and support, I also thank James Banks, Guillaume Blanc, Mathias Bühler, Alice Calder, Jeremiah Dittmar, Chicheng Ma, Agustina Paglayan, Nuno Palma, Melanie Xue, and the seminar participants at King's College London, Lewis Lab, Manchester, London School of Economics, Oxford. Sincere apologies to any colleagues whose names I may have inadvertently omitted. I alone am responsible for any remaining errors.

1 Introduction

Technology can exhibit gender-biased impacts. Certain innovations, such as the plough and cotton weaving, can generate persistent gendered effects due to biological differences in comparative advantage: the physical demands of plough agriculture favored men (Alesina, Giuliano and Nunn, 2013), while the fine motor skills required for cotton weaving benefited women (Xue, 2024). However, most technologies developed throughout history fall into a seemingly gender-neutral category—those offering no inherent biological advantage to either sex—such as printing press, internet or mobile phones. Yet a critical question remains underexplored: Can gender-neutral technologies produce asymmetric outcomes through their interaction with prevailing gender-biased institutions?

This paper examines one such technology: woodblock printing, the dominant printing technology in China from the 14th to the 19th century. Although woodblock printing is ostensibly a gender-neutral innovation, its educational impact likely interacted with deeply gender-biased institutions. For centuries, entrenched Confucian culture excluded women from formal schooling[1], making household access to printed materials their primary access for literacy acquisition. In contrast, men could leverage both formal schooling and supplementary resources such as libraries and scholarly networks. This institutional asymmetry generates two opposing hypotheses: (1) if woodblock printing functions as a substitute for formal schooling, women should experience greater marginal returns, while (2) if it complements formal education, the primary beneficiaries would be men.

Which gender ultimately gained more from the increased availability of woodblock printing? I assemble two datasets to investigate this question. The first is a two-period panel dataset spanning the Ming (1368–1644) and Qing (1644–1911) dynasties, which records the number of woodblock-printed books and the number of poets across 268 prefectures. This panel structure enables control for time-invariant prefecture characteristics. In Imperial China, poetry writing was the most common literary activity for educated men and, increasingly in the later period, for women (Cai, 2008, p.367), offering a unique window to assess educational attainment—particularly female literacy in a period lacking systematic social statistics.

The second dataset extends the analysis into the modern period by integrating cohort-level literacy estimates from the 1982 Chinese Population Census with historical data on the presence of woodblock-printed books (1368–1911). This dataset tracks educational attainment

[1] Confucius (551–479 BCE) said,"Only women and petty men are hard to educate." (*Analects* 17.25); while Ming scholar Zhang Dai (1597–1684) famously asserted "it is a woman's virtue to be ignorant"—both exemplifying traditional opposition to female education and the orthodox beliefs underpinning state policy.

across fifteen five-year birth cohorts, comprising: (1) eight pre-reform cohorts (born 1897–37) who received their primary education during the Republic of China period (born 1912–49), the era when girls first gained equal formal access to primary schooling alongside boys, and (2) seven reform-era cohorts (born 1938–72) educated during the Maoist era (1949–76), when the Mass Education Movement was implemented as part of Mao's reforms to "eliminate illiteracy." This division reflects fundamental institutional differences between the two regimes, particularly in their policy stances on women's formal education.

To identify the causal effect of woodblock printing on educational outcomes, I employ an instrumental variables (IV) strategy that exploits river distance to locations suitable for bamboo cultivation. Bamboo was a critical raw material for printing paper production during the Ming-Qing era (1368–1911), and book printing was often constrained by the availability of such resources. Using ecological factors linked to bamboo growth (higher precipitation, greater soil organic carbon, suitable pH levels, and lower elevation), I predict suitable cultivation sites based on records of bamboo presence from local gazetteers. The first-stage result shows that prefectures closer to these predicted bamboo sites by river are correlated with higher woodblock-printed book density, enabling variation in book that is plausibly unrelated to unobserved demand-side factors.[2]

I implement a two-stage least squares (2SLS) estimation using a two-way fixed effects (TWFE) framework, exploiting panel variation across the Ming and Qing periods. The results reveal a significant gender asymmetry: a 10% increase in the density of woodblock-printed books leads to a 9.4% increase in female poet density, whereas the corresponding estimate for male poet density is 2.3% and statistically insignificant. These estimates are robust to a rich set of historical controls, including measures of initial economic prosperity (population density), human capital (degree-holder and school density), geographic conditions (caloric suitability index, terrain ruggedness, proximity to coastlines and navigable rivers, and prefecture area), and political influence (distance to national capital and dummy indicating provincial capital). The finding suggests that female poets, who lacked access to formal educational institutions, benefited more from woodblock printing due to the increased availability of printed books for home-based learning.

The cohort-level analysis of modern literacy estimates using the same IV reveals that the female-biased effect of woodblock printing persisted exclusively for pre-reform cohorts (1898–1937), with no significant impact on reform-era cohorts (1938–72). A 10% increase in book

[2]Prior research by Chen, Kung and Ma (2020) used average river distance to bamboo and pine sites (documented in Qing gazetteers) as a proxy for the density of degree holders, arguing that printing inputs aided civil examination success. However, such measures may suffer from endogeneity if bamboo cultivation responded to rising book demand. See Section 4 for details.

density corresponds to an average 0.2 percentage point increase in gender equality in literacy (measured by the ratio of female-to-male literacy rate, where 1 indicates perfect equality) for pre-reform cohorts. These results suggest that even when formal schooling rights were nominally equal for boys and girls in the pre-reform era, prefectures with greater historical availability of woodblock printing developed more gender-equitable educational outcomes. This pattern persisted until the onset of Mass Education Movement under Mao, which substantially reduced barriers to women's primary education and narrowed gender gaps in literacy.

To interpret the persistence of this female-biased educational effect across generations despite formal institutional changes, I employ a triple-difference framework to test whether historical woodblock printing fostered a progressive cultural shift in investment in daughters' education. Using microdata from the 1982 Census, I construct an individual-level literacy dataset stratified by sex, prefecture of origin (bamboo-suitable or not), and parental literacy status, for pre-reform cohorts. This empirical strategy isolates the cultural effect by comparing the female-male educational gap between literate and illiterate households across treatment (bamboo) and control regions. The results indicate that residing in a treatment prefecture increased the likelihood of literacy for daughters (relative to sons) by an additional 17.5 percentage points in households with at least one literate parent (compared to illiterate households). These findings, while suggestive, provide evidence that greater historical availability of woodblock printing was associated with a more progressive culture regarding investment in daughters' education.

Printing has been recognized as one of three greatest innovations–alongside gunpowder and the compass–since Francis Bacon first highlighted their transformative impacts throughout the world[3]. While all three originated in China, scholars have noted a striking paradox: these inventions appear to have generated more significant socioeconomic impacts in Europe than in their country of origin.

This study re-examines the economic effects of printing technologies by shifting focus from Europe's movable-type—invented by Gutenberg in the 15th century Mainz—to China's woodblock—dating back to the 9th century.[4] While extant economic research has extensively documented how movable-type printing contributed to European economic growth (Dittmar, 2011), facilitated the Protestant Reformation (Rubin, 2014; Boerner, Rubin and Severgnini, 2021), and promoted civic engagement (Cagé and Rueda, 2016), scholars on woodblock

[3]Francis Bacon, *Novum Organum*, book I, CXXIX

[4]Printed in 868 CE using woodblock technology, *The Diamond Sutra* represents both the world's oldest dated printed book and a masterpiece of Tang Dynasty Buddhist art. This complete scroll, discovered in the Dunhuang caves, is now housed in the British Library.

printing remain largely passive—ranging from early dismissals of the technology as "outlived their usefulness" (Steinberg, 2017, p.113) to contemporary analyses of its limitations relative to movable-type (Angeles, 2017). This study advances current research on printing by showing the positive, but previously neglected, role of woodblock printing in promoting gender equality in education in China, where it served as a substitute for formal schooling in women's education. This finding invites further comparative evidence on how these two printing technologies functioned within different institutional and cultural contexts.

The female-favoring impact of woodblock printing is somewhat surprising, given the brute fact that the technology was not originally intended to promote women's education but rather to serve men's use. While we have a good understanding of gender-biased impact can arise—from gender-specific technologies such as plough (Alesina, Giuliano and Nunn, 2013) and cotton weaving (Xue, 2024)—or from gender-specific shock, such as rising tea prices benefiting women who were more efficient in tea plucking (Qian, 2008) (see Duflo (2012) for a detailed review), much less is known about how ostensibly gender-neutral technologies or shocks can generate gender-biased outcomes. My findings underscore the role of gender-biased institutions in determining the effects of technologies. This carries important policy implications: future technology implementations require ex ante institutional analysis to prevent unintended distributional consequences and ensure equitable benefit allocation.

This study also contributes to the broader literature on formal and informal institutions (Alesina and Giuliano, 2015), with particular focus on the ones affect human capital. It engages with research examining how formal institutions—such as welfare law adoption (Dittmar and Meisenzahl, 2020), religiosity (Squicciarini, 2020), and inclusive local autonomy (Serafinelli and Tabellini, 2022)—shape human capital formation. Additionally, the persistent female-biased effects of woodblock printing observed during the pre-reform era provide new insights into the literature on the cultural origin of parental investment decision in children's human capital (Ashraf et al., 2020; Bau, 2021).

Last but not least, this study adds to a thriving body of literature on measuring historical female representation (Goldin, 2006; Fernández, 2013; Bertocchi and Bozzano, 2016; Nekoei and Sinn, 2021; Bühler, Vollmer and Wimmer, 2024). This study extends the existing literature in two ways, first, I focus on poet groups, which offers an alternative comparable proxy for educational attainment by gender; second, I show that the rise of early female human capital originates from the increasing availability of woodblock printing, a pattern distinct from male human capital development in China. Without such alternative measures of female literacy, assessing technology's historical impact would remain profoundly challenging.

The rest of the paper is organized as follows. Section 2 provides historical background on Imperial China, focusing on the development of woodblock printing and the evolution of educational institutions from the 14th to the 20th century. Section 3 describes the data sources and construction. Section 4 details the instrumental variable strategy used to identify causal effects. Section 5 presents the main results on the contemporaneous impact of printing, while Section 6 examines the persistence of these effects into the modern period, exploring associated shifts in cultural norms regarding education. Section 7 concludes.

2 Background

2.1 Printing Technology in Imperial China

When printing was declared one of the greatest innovations in human history, little acknowledgment was paid to its Chinese origins—woodblock. Woodblock printing emerged in the Tang dynasty (618–907), expanded for religious and governmental purposes during the Song dynasty (960–1279), and ultimately achieved commercialization in the Ming-Qing period (1368–1911) (Tsien, 1985). This technique involved carving entire pages of text onto wooden blocks, which were then inked and pressed onto paper. In contrast, Western movable-type printing, developed by Gutenberg in the 15th century, relied on individual metal letters that could be rearranged to form different pages (Eisenstein, 1979). This fundamental technological divergence ultimately determined their differential historical recognition and impact.

Although movable-type was first invented in China by Bi Sheng (972–1051) during the Song dynasty—five centuries before Gutenberg—it never supplanted woodblock printing due to the complexity of the Chinese writing systems, which required at least 8,000 distinct characters to print a single book (Tsien, 1985; Twitchett et al., 1983). As a result, woodblock printing remained the dominant method of text reproduction in China until the introduction of lithography in the late 19th century (Wilkinson, 2012, p.909). Despite its ostensibly lagged techniques, Englishman John Barrow (1764–1848), a member of the Macartney mission to Qing China, observed that Chinese printing was "as free as in England, and the profession of [woodblock] printing open to everyone" (Barrett, 2008, p.11).

A defining feature of woodblock printing was the immobility of its production infrastructure. Unlike movable-type printing, which relied on mechanical presses and allowed for flexible relocation, woodblock printing required a dedicated set of carved wooden blocks for each book. These blocks were bulky, difficult to transport, and expensive to reproduce, making

Table 1: Total Number of Woodblock-printed Books in Each Province, 1368–1911

Province	Ming (1368–1644)	Qing (1644–1911)
Jiangsu	4,690	48,060
Zhejiang	1,052	24,692
Hunan	93	12,574
Hubei	103	10,171
Guangdong	135	10,028
Zhili	918	9,424
Sichuan	142	5,809
Anhui	607	4,971
Jiangxi	386	4,738
Shandong	235	2,738
Fujian	371	2,662
Henan	175	2,321
Shanxi	266	1,706
Shaanxi	104	1,528
Guangxi	12	1,250
Gansu	11	524
Guizhou	1	479
Yunnan	6	475
Total	9,307	144,150

Source: National Census of Ancient Book Database. See section 3 for detailed data construction.

them inherently immobile. As a result, printing workshops were typically stationary and served local markets (Brokaw, 2007*a*). Furthermore, maintaining and storing large volumes of blocks demanded substantial space and capital investment, reinforcing the localized and capital-intensive nature of the technology. Online Appendix Figure A1 shows a sample woodblock used to print a single page of a religious text, highlighting the physical scale and storage requirements that constrained workshop mobility.

This geographic immobility contributed to a decentralized and fragmented printing industry. Most workshops produced only a narrow range of titles—typically four to five—and served primarily local markets. As a result, the diffusion of printed materials was highly localized, with limited spillovers beyond nearby regions. Unlike the European movable-type printing revolution, which supported the development of integrated commercial book markets, China's woodblock printing remained tied to local networks of kinship, officials, and elite scholarship (McDermott, 2006; Chia, 2020). This constrained dissemination had important implications for educational access, particularly for women, whose learning was confined to the domestic sphere and relied heavily on the availability of books within the household.

The spatial distribution of book production further illustrates these constraints. Table 1 reports the number of woodblock-printed books published in 18 provinces during the Ming and Qing periods using The National Census of Ancient Book database (National Library of China, 2024). Major centers such as Jiangsu and Zhejiang accounted for over 50% of all printed books nationwide, becoming hubs for scholarly and literary output. In contrast, peripheral regions like Yunnan and Guizhou had substantially lower book production. Furthermore, the data indicate that total output during the Ming dynasty remained relatively modest compared to the output during the Qing dynasty. This pronounced regional and temporal variation in the prevalence of woodblock-printed books provides a valuable measure of differential availability of woodblock printing across time and space.

2.2 Gender-biased Educational Institutions in China

2.2.1 Education before 1911

Prior to the Chinese Revolution of 1911, formal education and employment opportunities were largely inaccessible to women due to prevailing Confucian cultural norms, which excluded them from public life and further constrained their educational prospects (Rawski, 1979, p.7). The imperial examination system—the principal pathway to official positions and social mobility—was entirely closed to women (Elman, 2000), further reducing the incentive and institutional infrastructure for female schooling. Even as late as 1904, *The Regulation for Schools* promulgated by the Qing government explicitly stated that only men were permitted to attend formal schools. As a result, for centuries women's education relied almost exclusively on domestic instruction—a mode of learning that depends heavily on the availability of printed materials within the household.

Further compounding institutional barriers was the widespread practice of foot-binding—a custom that involved breaking and tightly binding young girls' feet to achieve a culturally idealized appearance. This practice resulted in lifelong disability and severely limited mobility, reinforcing the ideal of female domesticity (Ko, 2005). Ironically, for elite women, this physical confinement sometimes encouraged intellectual pursuits such as reading and writing. As Ko (1995) notes, access to printed materials enabled some gentry women to pursue self-education, and the spread of woodblock printing significantly enhanced this possibility by increasing the availability of texts within the home.

The gendered structure of education was further institutionalized through family instructions (*jiaxun*), quasi-legal documents that regulated behavior across generations and enforced patriarchal norms.[5] For example, *The Family Instructions for the Yu Clan*, compiled in 1604, stated: "When boys and girls reach the age of five or six, boys should be sent to school to learn the basic subjects, while girls should remain within the confines of their homes and be taught basic texts, such as the biographies of virtuous women and rules of conduct for women." This explicit differentiation in educational expectations reinforced the divide between male and female learning environments and further gendered the use of printed materials.

2.2.2 Progress Toward Equality in the 20th Century

The early twentieth century brought a gradual shift in attitudes toward women's education. Following the fall of the Qing dynasty, the Republican government (1912–49) adopted a more progressive stance, increasingly recognizing women's right to education. Reformist elites promoted gender equality in principle, and female enrollment rose, especially in normal schools and missionary institutions (Bailey, 2007). However, educational expansion during this period was limited in scope, and progress in female education lagged significantly behind that of males. Although state rhetoric supported education for both sexes, implementation was often vertical rather than horizontal—targeting elite inclusion rather than mass accessibility. Consequently, a large portion of women remained excluded from formal schooling, and significant gender disparities in primary education persisted. (Further statistical evidence will be presented in detail in Section 3.)

A more radical transformation took place after the establishment of the People's Republic of China in 1949. The new Communist regime prioritized mass literacy and framed education as a tool of national development and ideological transformation. Under Mao's leadership, the state launched the Mass Education Movement, aiming to eradicate illiteracy among both men and women, particularly in rural areas (Gao, 2015). Unlike the Republican period, the Communist approach was universal, compulsory, and secular. Public primary schools rapidly expanded, and female enrollment rates surged as the state actively promoted gender equality in access to education. This inclusive educational institutions helped close the gender gaps in primary education, marking a fundamental departure from previous regimes. The contrast across institutional contexts is consistent with global evidence suggesting that non-democratic regimes can, at times, expand literacy effectively through large-scale, state-led education campaigns (Paglayan, 2021; Palma and Reis, 2021).

[5] Family instructions were widely regarded as moral codes and binding rules established by clan ancestors to govern future conduct (see Knapp, 2005; Deng and Lamouroux, 2005).

3 Data

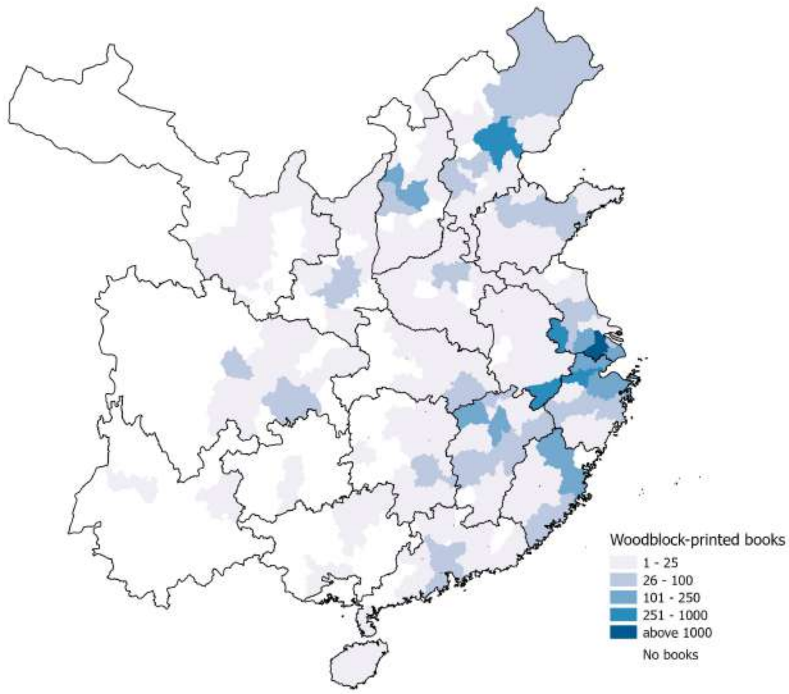
3.1 Woodblock-printed Books

The data on woodblock-printed books are sourced from The National Census of Ancient Book database (hereafter, "Book Census"), which includes all extant ancient books held in 318 national, provincial, and county-level libraries and museums (National Library of China, 2024). The database provides detailed information on book titles, number of volumes, publication places, and publication dates. As of May 2024, the Book Census contains records for 825,362 ancient books, of which 600,008 (73%) are identified as woodblock-printed books (*keben*)—the predominant medium of text reproduction in historical book publishing.

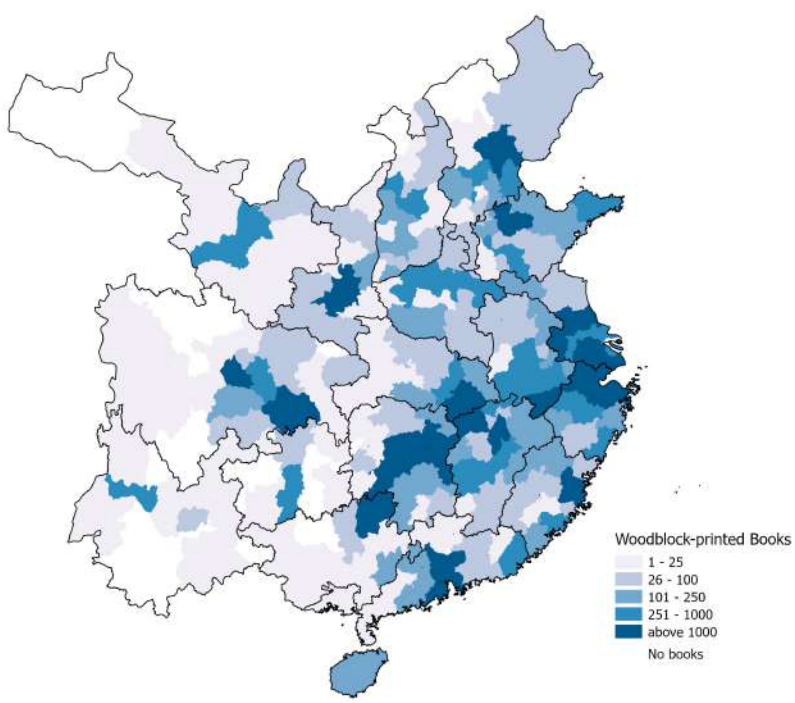
I systematically scrape all woodblock-printed entries from the database and geocode their printing locations based on explicitly listed publication sites or identifiable publisher names[6], leaving a final analytical sample of 9,307 books from the Ming dynasty and 144,150 books from the Qing dynasty, representing 25.6% of the woodblock-printed books in the original sample. This relatively low retention rate is primarily due to missing publication place information. Two key historical factors explain this prevalent data absence: first, woodblock printing was a highly localized and decentralized industry (Brokaw, 2007*b*), often producing for smaller, local circulations that did not require explicit origin statements. Second, the absence of legal protections, such as copyright law, attenuated the incentive for publishers to clearly identify the origin of a printed work (Alford, 1995).

The identified publication locations of books from the Book Census are georeferenced to one of 268 second-level administrative divisions (prefectures) in 1820 using historical GIS data from the China Historical GIS project (CHGIS, 2016). I aggregate the total number of woodblock-printed books published in each prefecture during the Ming (1368–1644) and Qing (1644–1911) dynasties. Figure 1 illustrates the spatial distribution of woodblock-printed books during this period, with darker shading indicating greater presence. While printing activity was geographically widespread, it remained heavily concentrated in southeastern China. Because books published during the Ming dynasty (1368–1644) remained in circulation throughout the Qing (1644–1911), the combined total of 153,457 books from both dynasties is used to represent historical book presence during the Qing period.

[6]To ensure accurate spatial identification, I utilize the China Biographical Database Project (CBDB, 2019) to match publisher names to known historical individuals. I further restrict matches to cases where the individual's recorded lifespan overlaps with the book's publication year, thereby ensuring temporal consistency.



(a) Ming (1368–1644)



(b) Qing (1644–1911)

Figure 1: The Distribution of Woodblock-printed Books in Ming and Qing Dynasties

Note: Panel (a) displays the distribution of woodblock-printed books during the Ming Dynasty. Panel (b) shows the distribution for the Qing period, calculated as the cumulative sum of books in both the Ming and Qing dynasties. Solid black lines indicate provincial boundaries.
Source: National Census of Ancient Books.

To assess the representativeness of the surviving book data, I also digitized an authoritative historical source on Ming-period imprints, *Quanming Fensheng Fenxian Keshu Kao* ("A Catalogue of the Books Printed in the Ming", hereafter, "Du Catalogue") by Du and Du (2001), documenting 7,544 imprints with exact printing locations. The prefecture-level counts from the Book Census and the Du Catalogue reveal a high correlation of 77%. Although the Book Census data is limited, it offers two primary advantages over other sources like the Du Catalogue. First, it better captures the number of books in circulation, as standard catalogues typically record distinct titles but lack information on print runs per title. Second, to my knowledge, the Book Census is the most suitable database for constructing a panel dataset of printing activity spanning both the Ming and Qing dynasties.[7]

To adjust for population size differences across prefectures, I construct a normalized measure of book density by dividing the total number of books (plus one) by the average population (in millions) during the corresponding dynasty. This density is then log-transformed to ensure comparability across regions and to correct for right-skewness in the distribution, preserving all 268 prefectural observations in the analysis.

3.2 Outcome Variables

3.2.1 Poets as a Proxy for Educational Outcome

To measure the educational outcomes for both men and women from the 14th century onward, I exploit a unique historical source: the recorded presence of poets during the Ming (1368–1644) and Qing (1644–1911) dynasties. In a historical context where systematic social statistics on women's education are virtually absent, poetry offers a rare and valuable proxy for educational outcome across genders.

Poetry held a central position in Chinese literary and cultural life for the educated population. Unlike other forms of writing and creative work (such as painting and composing), poetry did not require professional or bureaucratic training and was a relatively accessible medium of expression for any literate individual. While rooted in a long tradition for educated men, it was also widely practiced by educated women. As Mann observes in her study of the High Qing era, "No educated woman I have encountered in the documents from this period complains that she cannot write poetry..." (Mann, 1997). This suggests the remarkable

[7]For example, due to the extensive number of books printed during the Qing Dynasty, Du was unable to complete a corresponding complete catalogue for Qing-period imprints, leaving only 4,447 entries before his passing.

ubiquity of poetry writing even among women, who were otherwise systematically excluded from most forms of public intellectual life.

Poetry was typically circulated within kinship or local networks rather than through commercial channels. Many women's collected works were published posthumously by family members, underscoring the domestic and retrospective nature of their literary recognition (Fong, 2017). For example, the preface to *Jiang Zhu Ge Xiu Yu Cao*, a collection by the female poet Wu Xiuzhu (1808–27), notes that her two younger sisters compiled her manuscripts and arranged their publication after her death. The modest appearance of the printed volume and absence of a known publisher suggest that the collection was produced for familial rather than commercial purposes. Such examples reflect that female authorship was not driven by market incentives, strengthening the credibility of poet records as a measure of literacy.

To demonstrate the comparability of literary output across genders, I include two sample poems: one by Yun Zhu (1771–1833), a prominent female poet, and another by Yuan Mei (1716–97), a leading male poet of the Qing dynasty. Each poem adheres to classical Chinese prosody, suggesting equivalent standards of literary and intellectual skill. Since poetic composition presupposes a high degree of literacy, the presence of poets in historical records provides an indirect yet meaningful indicator of educational attainment. The prevalence of recorded female poets, in particular, offers a window into women's literacy under conditions of institutional exclusion.

暮春（惲珠）	Fading Spring (Yun Zhu)
不人春將去，	[Spring departs without a sound,]
天涯人未還。	[You are still nowhere to be found.]
可憐巾上淚，	[Silent tears keep falling on—]
點點盡成斑。	[Every drop, a petal gone.]
所見（袁枚）	What I Spot (Yuan Mei)
牧童騎黃牛，	[He sings atop his ox,]
歌聲振林樾。	[His voice shakes pine and rock.]
意欲捕鳴蟬，	[A cicada suddenly calls—]
忽然閉口立。	[Still as midnight clock.] [8]

Data on female poets are sourced from the open-access database *Ming Qing Women's Writings* (MQWW, University, 2025). As of December 2022, the database contains 426 published and manuscript poetry collections. These collections were mainly published during the Ming

[8] Author's translation.

and Qing dynasties; some were published in the Republican period (1912–49); and a small proportion of them were hand-written poetry. MQWW database also provides accompanying biographical information including approximate birth years and locations for each female poet. After geocoding these locations to one of 268 prefectures in China Proper as of 1820, I include a total of 2,907 female poets: 263 born during the Ming and 2,644 during the Qing.

Data on male poets are drawn from two major sources: *Ming Shizong* ("Ming Poetry Collection", Zhu Yizun, 1705) and *Wanqing Yishihui* ("Qing Poetry Collection", Xu Shichang, 1929). These anthologies provide short biographies and representative poems for each poet.[9] I manually geocode each poet's birthplace using biographical information, mapping them to prefecture-level administrative divisions as defined in 1820. After restricting the sample to individuals born within China Proper and with identifiable birthplaces, the final dataset includes 2,288 male poets from the Ming dynasty and 5,285 from the Qing dynasty, representing 69% and 86% of the total sample from each source, respectively.

To evaluate the contemporaneous impact of woodblock printing on educational outcomes, I construct a two-period (Ming and Qing) panel of poet density measures for each prefecture. Specifically, I calculate poet density by dividing the number of male and female poets (plus one to accommodate zero values) by the average population during the corresponding dynasty. I also employ log-transformation to overcome the skewness.

A key rationale for employing a two-period panel structure stems from data limitations concerning poet birthyears. A significant portion of poets, particularly females, lack precise birthyear information and are only identifiable by their dynasty of birth. To maximize sample retention and mitigate selection bias that would arise from excluding these observations, I aggregate the data into these two broad dynastic periods rather than employing a finer temporal segmentation.

3.2.2 Modern Literacy using 1982 Census

To assess whether the impact of woodblock printing extended beyond elite groups to the broader population, and to extend the analysis into the modern period, I turn to modern

[9] The compilation methodologies of these anthologies inform their representativeness. Zhu Yizun stated in his preface that his selection of poets for the *Ming Shizong* was based solely on literary merit, without regard to political status or official achievement. However, he also acknowledged a practical constraint: the difficulty of locating older manuscripts meant the collection overrepresented more recent poets. For the *Wanqing Yishihui*, a 1921-09-13 report in *Huangbao* (government newspaper) noted that Xu Shichang and his staff amassed thousands of books from across the country, drawing from both published literature and private manuscripts. Faced with this abundance, their editorial standard shifted toward maximal inclusion rather than strict curation.

literacy estimates drawn from the Third National Population Census 1982, obtained via IPUMS International (Ruggles et al., 2025). This dataset provides comprehensive individual-level microdata, comprising a 1% sample of the total population (over 10 million individuals). Its coverage of cohorts who received their primary education during both the Republican (1912–49) and Communist (1949–) eras makes it particularly valuable for tracing educational outcomes across distinct institutional periods. Another advantage of this source is that it is the earliest available national census that details demographic (e.g., gender, age) and educational information at the individual level.

Based on the census questionnaire, respondents were asked to classify their highest educational attainment based on six options: illiterate or semi-literate, primary school, junior middle school, senior middle school, undergraduate, and college graduate. Following official guidelines, I classify an individual as literate if they reported completing primary school or higher. This criterion is robust because, according to the *Instructions for Filling Out the Questionnaire of the Third National Population Census* (Population and Statistics Department, 1982), individuals with informal schooling (e.g., private schools, private tutors) could self-report a comparable formal level. Specifically, anyone knowing roughly 1,500 or more characters and demonstrating practical literacy skills (e.g., reading a newspaper, writing simple notes) could be classified at the primary school level. This definition thus captures literacy acquired outside the formal school system, mitigating underestimation bias for older cohorts and women. Using this measure, I construct literacy rates for fifteen five-year birth cohorts born between 1898 and 1972.

To assess the effect of institutional change, I divide fifteen cohorts into eight pre-reform cohorts (born 1898–1937) and seven reform cohorts (born 1938–72). The 1938 cutoff is chosen because individuals born in this year would have been of primary school age (7–11 years old) in 1949—the year the Communist regime assumed power and initiated a revolutionary educational institutions. My analysis focuses specifically on primary education due to its uniquely extensive coverage—by 1946, public primary schools already accounted for approximately 95% of all schools (Zhu, 1948), establishing the foundational infrastructure that enabled the rapid implementation of broad-based educational reforms after 1949—a level of penetration not yet achieved by higher educational institutions at that time.

Figure 2 reveals a steady rise in literacy across successive cohorts. Among males, literacy rates increased from 29% for the earliest cohort to 67% for the final pre-reform cohort. Female literacy also exhibited an upward trajectory, rising from 2% to 24% over the same period, yet a substantial gender gap persisted until 1949. In the three decades following the 1949 reforms, these trends accelerated dramatically: male literacy reached 96%, while female literacy saw an

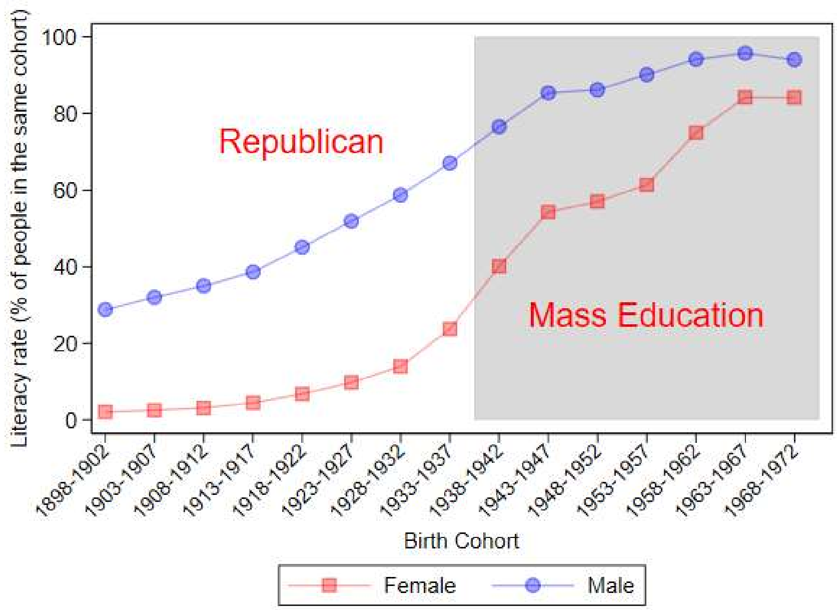


Figure 2: Literacy Rates by Gender Across Birth Cohorts

Note: This figure plots average literacy rates by gender for each five-year birth cohort. The shaded grey region marks reform-era cohorts who received primary education during the Mass Education Movement under Mao's regime since 1949.
Source: The Third National Population Census 1982.

even steeper increase to 84%, suggesting the strong equalizing effects of Mao-era educational reform. The convergence of these trends provides a powerful empirical setting to examine how the historical legacy of woodblock printing interacted with modern institutional changes to shape long-term educational outcomes.

3.3 Control Variables

I incorporate three strands of control variables to account for confounding factors correlated with woodblock-printed books and educational outcomes. Descriptive statistics for all variables across the 268 prefectures are summarized in Online Appendix Table A1. Unless otherwise noted, most spatial data are derived from the China Historical GIS (CHGIS, 2016).

3.3.1 Baseline Controls

Population density. To capture pre-industrial economic development, I include population density in 1393 (early Ming) and 1680 (early Qing), sourced from Cao (2024), a commonly used proxy in historical economic research.

Jinshi density. To account for upper-tail human capital of males, I include the density of *jinshi*—holders of the highest degree under the imperial examination system. The *jinshi* data are sourced from Zhu and Xie (1980), which documents all 46,908 individuals who attained the degree between 1371 and 1904. I classify these degree holders by the period in which they obtained the title: those who earned the *jinshi* degree during the Ming dynasty (1368–1644) and those during the Qing dynasty (1644–1911). Degree holder density is calculated using the same definition applied to book and poet density.

School density. To capture the broader educational environment, I include the density of schools established before 1842, a reference point in the late imperial period. The school data are manually collected from the *Unified Chorograph of the Qing Dynasty* (*Jiaqing Chongxiu Yitong Zhi*, 1842), which records 3,549 educational institutions—both official and private—along with their geographic locations and founding year. I classify these schools by founding period: those established before 1368, during the Ming dynasty (1368–1644), and during the Qing dynasty (1644–1911). School density is calculated using the same definition applied to book and poet density.[10]

3.3.2 Geographical Controls

Caloric suitability index. To account for regional variation in agricultural productivity, I include the caloric suitability index from Galor and Özak (2016). This measure captures the potential for crop-based caloric output across regions, which may influence patterns of population settlement and economic prosperity, and in turn, affect investment in education.

Terrain ruggedness index. To account for geographic barriers to agriculture and transportation, I control for the terrain ruggedness index, constructed using data from the Digital Elevation Model (USGS, 2015) and following the methodology of Nunn and Puga (2012).

Distance to coast and navigable river. To capture access to trade and transport networks, I compute each prefecture's shortest distance to the coastline and to navigable rivers, using river data from Matsuura (2009) and spatial data from CHGIS (2016).

Prefecture area. To account for variation in spatial scale, I include the logarithm of each prefecture's land area. Larger administrative units may encompass more diverse economic activities, which could influence educational and cultural development.

[10] There was no significant change in the establishment of schools after 1842, and the 1842 source represents the most complete surviving documentation of imperial-era schools.

3.3.3 Political Controls

Administrative proximity. Although woodblock printing became increasingly commercialized during the Ming and Qing dynasties, it was also shaped by state demand and administrative geography. To account for this, I include the distance from each prefecture to Beijing (measured in 100 kilometers) and a dummy variable indicating whether a prefecture was a provincial capital. These variables proxy the influence of central and regional governments in shaping both book production and educational development.

3.4 Additional Variables for Predicting Bamboo Suitability

Here outlines the four primary data used to construct the instrument variable: river distance to locations suitable for bamboo cultivation. Following agronomic guidelines from Farrelly (1996), I identify four key ecological factors necessary for bamboo growth: precipitation, soil organic carbon, soil pH, and elevation. As Farrelly notes, bamboo thrives in well-watered, rich in organic matter, slightly acidic soils, and generally prefers lower elevations. The full empirical strategy is detailed in Section 4.

Precipitation. Data are drawn from the FAO Climate Research Unit (CRU) CL 2.0 dataset (New et al., 2002), which provides annual rainfall at a 10 arc-minute resolution ($\sim$16.7 km) for 1961–1990. Although high-frequency historical rainfall data for the Ming-Qing period are unavailable, long-run climatic averages are sufficiently stable over the relevant time horizon to serve as a reliable proxy for historical growing conditions.

Soil organic carbon and soil pH. Soil data come from the Harmonized World Soil Database (HWSD v2.0, Nachtergaele et al., 2023), offering 1 km spatial resolution. I use the average organic carbon level as a continuous measure and generate a binary indicator for optimal soil pH (5.5–6).

Elevation. Data are obtained from the Harvard China Historical GIS Digital Elevation Model (CHGIS, 2016, based on GTOPO-30 from the USGS), spaced at 30-arc-second pixels ($\sim$1km sq / pixel) across the entire surface of the earth on a geographically projected map. I compute the mean elevation for each of the 268 prefectures.

Recorded bamboo presence. To identify historically suitable locations for bamboo cultivation, I utilize data on recorded bamboo presence during the Qing period, originally compiled by Chen, Kung and Ma (2020). These data were extracted from systematic reviews of provincial gazetteers from the Qing dynasty. (See Chen, Kung and Ma (2020) for a complete listing of original sources). Of the 268 prefectures, 55 reported bamboo resources. Online Appendix Table A3 presents a balance table comparing four ecological factors between prefectures with and without recorded bamboo sites. The results reveal significant differences in these factors, suggesting that the presence of recorded bamboo sites is systematically associated with these ecological characteristics.

4 Proximity to Bamboo as Instrument Variable

4.1 Rationale of IV

To overcome potential endogeneity concern that the availability of woodblock printing may be associated with preexisting demand for education (e.g., regions with a stronger educational culture may exhibit greater demand for woodblock printing, leading to upward bias of estimate), I leverage an instrumental variable—river distance to locations suitable for bamboo cultivation, which affects the supply of woodblock printing presses but should be orthogonal to the demand for education.

During the Ming and Qing dynasties, bamboo became the dominant raw material for printing paper due to its abundance, durability, and affordability (Tsien, 1985; Brokaw, 2007*a*). Prior to the Ming dynasty, paper production relied on bark and other plant fibers, which lacked the resilience necessary for large-scale printing. This improvement in papermaking enabled the widespread use of bamboo paper, reducing production costs and expanding access to printed materials (Zhang and Han, 2006). Online Appendix Figure A2 illustrates the stages of bamboo-based paper-making in detailed, which was first documented in *Tiangong Kaiwu* (*The Exploitation of the Works of Nature*) published in 1637. This fact suggests that the use of bamboo for mass paper production was a relatively new and innovative development of the Ming era.

As a result, in the woodblock printing industry with relatively fixed operating costs (e.g., workshop rent, block carving, labor), regions with closer access to bamboo—facilitated by river transport—faced lower marginal costs for printing paper. This cost advantage stimulated the establishment and operation of local woodblock printing presses, thereby increasing the regional supply of books. Proximity to bamboo suitable areas via navigable rivers thus served as a critical determinant of regional variation in woodblock printing output.[11]

[11] River transport was the principal mode of bulk freight in Imperial China. For instance, in western Fujian saw bamboo paper was shipped downstream in loads of approximately 30,000 sheets (1.8 metric tons) to major commercial hubs like Guangzhou and Chaozhou (Brokaw, 2007*a*, p.116).

This approach builds on the intuition of Chen, Kung and Ma (2020), who use river distance to bamboo and pine sites—identified from Qing-era provincial gazetteers—as an instrument for degree-holder density. They reasoned that bamboo and pine, as the raw materials for paper and ink during the Ming-Qing period, contributed to the publication of reference books essential for success in civil exams. Therefore, proximity to these raw materials by river provided civil examinees with a marginal advantage in attaining degrees through the imperial examination system.

However, a potential concern of using historically documented sites is that these recorded sites may have been endogenous to local demand for printed materials; that is, regions with higher demand for woodblock printing may have been more likely to cultivate and record such resources. Therefore instead of using these recorded sites of bamboo cultivation, I predict suitable sites for bamboo cultivation using exogenous agroecological factors—specifically precipitation, soil organic carbon and suitable pH level, and elevation (Farrelly, 1996)—which are unrelated to historical demand for printing.

Moreover, I focus exclusively on bamboo rather than combining it with pine, as printing paper was the only key input that experienced significant technological innovation during the Ming period; no comparable innovation occurred for ink production using pine (Needham, 1974, p.5). This exclusive focus is empirically validated in Online Appendix Table A2, which shows that proximity to historically recorded bamboo sites is strongly correlated with woodblock-printed book density, whereas proximity to pine sites exhibits no significant relationship—further justifying the choice of instrument.

4.2 Construction of IV

The construction of the IV involves two steps: first, predicting locations suitable for bamboo cultivation based on agroecological factors; and second, calculating the river distance from each prefecture to the closest predicted bamboo location.

To predict, I first employ a probit regression of recorded bamboo presence on the four ecological variables associated with bamboo growth, see Section 3 for a review of data source:

$$\begin{aligned} &Pr(\text{Whether had bamboo on records} = 1|X_i) = \\ &\quad \beta_1 \text{Precipitation}_i + \beta_2 \text{Organic carbon}_i + \beta_3 \text{Suitable soil pH}_i + \beta_4 \text{Elevation}_i + \epsilon_i \end{aligned} \tag{1}$$

Online Appendix Table A4 reports the results. Precipitation, organic carbon, and suitable pH are positively associated with recorded bamboo presence, while elevation is negatively associated. Using the estimated coefficients, I compute the predicted probability of bamboo suitability for each prefecture and classify a prefecture as suitable if its predicted probability exceeds 0.5; this criterion identifies 36 out of 268 prefectures as naturally suitable for bamboo cultivation.

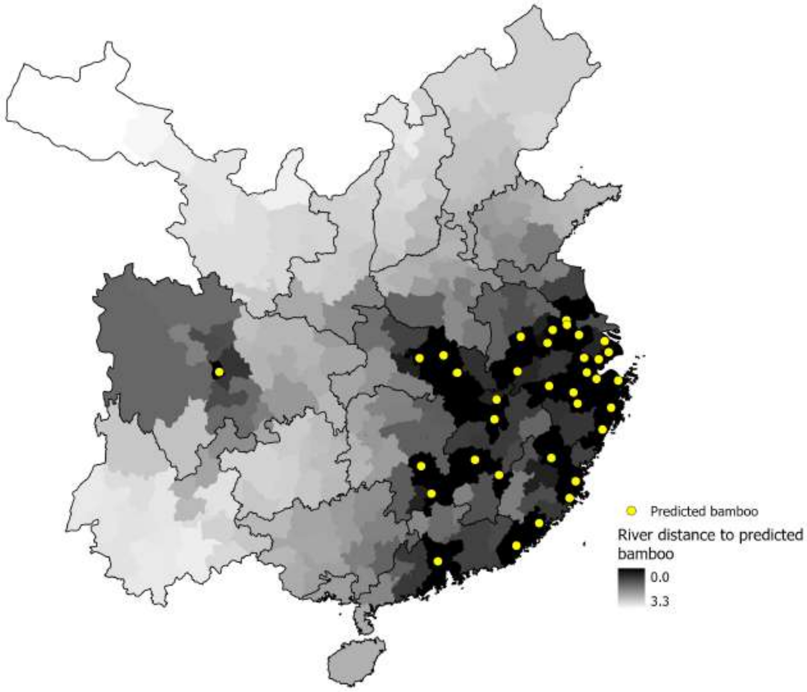
This approach distinguishes ecological suitability from historically recorded presence, as the latter may reflect factors like commercial demand for books, state-led documentation efforts, or cultural symbolism rather than natural growing conditions. Online Appendix Figure A3 maps the distributions of 36 predicted bamboo locations, 33 historically recorded bamboo locations, and the number of woodblock-printed books. For instance, Xi'an—a major administrative center—appears in historical records as a bamboo-producing site despite having a low predicted ecological suitability (0.2), suggesting its documented presence likely reflects its institutional prominence rather than a natural comparative advantage. By contrast, prefectures in the southeastern region exhibit a strong alignment between predicted locations and historical records, which corresponds closely with their high concentration of woodblock printing activity.

To capture access to bamboo resources via river transport, I construct the instrumental variable as the shortest river distance (in kilometers) from each prefecture seat to the nearest predicted bamboo site.[12] To visualize the first-stage relationship, Figure 3a maps the predicted bamboo sites against the log river distance to these sites, while Figure 3b maps these sites against the log density of woodblock-printed books during the Qing period. Both figures reveal a closely aligned spatial pattern, with concentrations predominantly in southeastern regions.

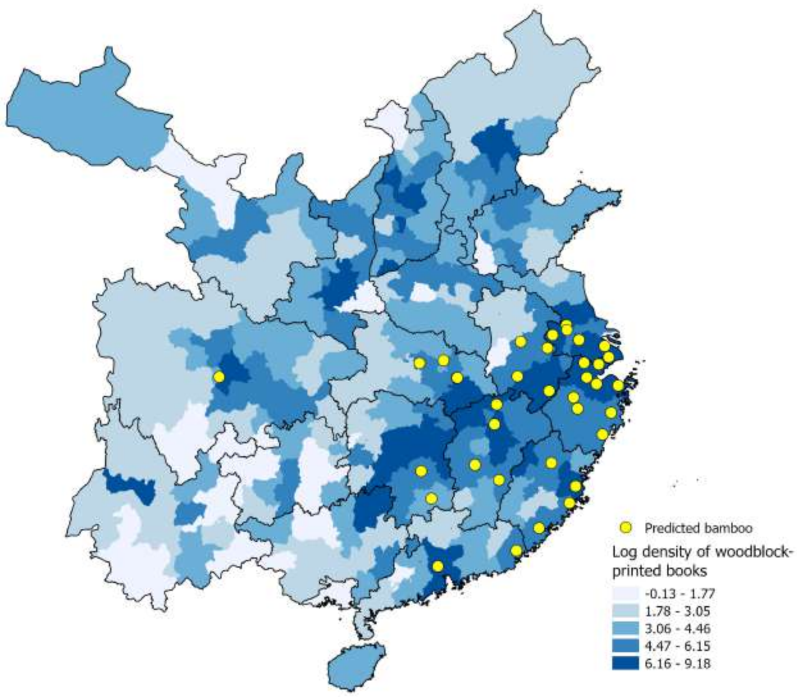
4.3 Validity of IV

To assess the validity of the instrumental variable, I begin by estimating a two-period panel specification that exploits variation across the Ming and Qing dynasties. Specifically, I regress the log density of woodblock-printed books in prefecture i and dynasty t on the interaction between log river distance to locations suitable for bamboo cultivation and an indicator for the Qing dynasty. The model is specified as follows:

[12]River distances are computed along the navigable river network defined by Matsuura (2009), which served as the primary infrastructure for bulk commodity transport during the Ming-Qing period. Online Appendix Figure A4 depicts the river system on map. This network remained largely unchanged throughout the era, ensuring temporal consistency for the instrument.



(a) IV—Log River Distance to Predicted Bamboo



(b) Log Density of Woodblock-printed Books in Qing

Figure 3: Predicted Bamboo, Book Density, and River Distance to Predicted Bamboo

Note: Panel (a) displays the distribution of 36 predicted bamboo prefectures (predicted probability > 0.5) alongside the log density of woodblock-printed books during the Qing period. Panel (b) shows the distribution of predicted bamboo and the instrumental variable—river distance to predicted bamboo sites. Darker shading indicates closer proximity to predicted bamboo cultivation location. Navigable river data are sourced from Matsuura (2009).

Table 2: First-stage: River Distance to Predicted Bamboo and Book Density

	(1)	(2)	(3)
		Log density of books	
Log river distance to predicted bamboo × Qing	-0.585***	-0.379***	-0.452***
	(0.099)	(0.092)	(0.100)
Observations	536	536	536
R^2	0.838	0.867	0.879
Mean	3.337	3.337	3.337
First-stage F-stat	42.761	18.685	20.637
Baseline Controls	Y	Y	Y
Geographical Controls		Y	Y
Political Controls			Y

Note: This table reports the first-stage estimates from a regression of the log density of books on the interaction between log river distance to predicted bamboo locations and a Qing dynasty indicator. Baseline controls include the log density of population, log density of schools, and log density of degree holders. Geographical controls include the caloric suitability index, terrain ruggedness index, log distance to the coast and nearest navigable river, and log prefecture area. Political controls include log distance to the national capital and a dummy indicating provincial capitals. All regressions include prefecture and dynasty fixed effects. Standard errors, reported in parentheses, are adjusted for spatial autocorrelation as described in the main specification. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

$$\ln(\text{book density})_{it} = \theta_i + \delta_t + \beta \ln(\text{river distance to predicted bamboo})_i \times \text{Qing} + \Gamma X_{it} + \Lambda X_i \times \text{Qing} + \epsilon_{it} \tag{2}$$

Where prefecture $i = \{1, 2, \ldots, 268\}$ and dynasty $t = \{Ming, Qing\}$. θ_i is prefecture fixed effects, δ_t is time fixed effects. Column 1 of Table 2 includes only time-varying baseline controls X_{it}, specifically the log density of population, degree holders, and schools (Table 2 Column 1). Column 2 adds a set of time-invariant geographic controls (caloric suitability index, terrain ruggedness index, log distance to coast and river, log prefecture area) interacted with a Qing dynasty indicator. Column 3 further incorporates political controls (log distance to national capital and dummy indicating provincial capital), also interacted with the Qing dummy.

The results in Table 2 suggest that prefectures located closer to predicted bamboo locations by river experienced a greater increase in woodblock-printed book density during the Qing relative to the Ming dynasty. This finding supports the relevance of the instrument, suggesting that bamboo proximity by river became more closely tied to printing activity as the sector commercialized. Across all specifications, the first-stage F-statistic exceed the conventional threshold of 10 recommended by Stock and Yogo (2002), thus mitigating concerns about weak instrument.

While predicted bamboo is determined by exogenous ecological factors, it is crucial to rule out alternative channels through which predicted bamboo proximity might influence educational outcomes, such as through early economic prosperity or broader economic development during the Ming-Qing period. For instance, one concern that bamboo was used for other commercial goods, thereby increasing economic development and demand for education. A potential concern is that bamboo may have been used to produce other commercial goods, thereby stimulating local economic development and, in turn, demand for education.

To assess the validity of the instrument and support the exclusion restriction, I therefore test whether proximity to predicted bamboo is correlated with pre-existing economic conditions or subsequent population growth. Given that the commercialization of printing began during the Ming dynasty (1368–1644), I test whether river distance to predicted bamboo is associated with initial economic conditions at the onset of the Ming dynasty. Table 3 reports regressions of river distance to predicted bamboo on three key indicators: (i) log density of population in 1393, (ii) the caloric suitability index, and (iii) the log density of schools established before 1368. In addition, I test whether river distance to predicted bamboo locations is associated with population growth during the Ming (column 4) and Qing (column 5) periods. Across all specifications, the estimates of river distance to predicted bamboo are statistically insignificant, suggesting that the instrument is orthogonal to early economic conditions that could directly affect educational outcomes, as well as to broader economic development during the analysis period. These findings provide empirical support for the exclusion restriction and the validity of the identification strategy.

Historical evidence further corroborates these results. During the Ming and Qing eras, bamboo was used locally for construction and transportation—such as in scaffolding, simple dwellings, fences, ladders, baskets, and river rafts. However, these applications were largely subsistence-oriented and non-commercial, with bamboo functioning as a minor forest product rather than a commodity in large-scale trade or industry (Liu et al., 2018). Consequently, bamboo proximity is unlikely to have influenced economic development through channels other than its role in reducing the cost of paper production for printing.

Table 3: River Distance to Predicted Bamboo and Early Economic Conditions

	(1) Pop den 1393	(2) Cal suit index	(3) Sch den 1368	(4) Pop den 1680	(5) Pop den 1820
Log river distance to predicted bamboo	-0.257	-0.043	-0.041	-0.042	-0.047
	(0.173)	(0.025)	(0.089)	(0.070)	(0.122)
Log density of population in 1393				0.643***	
				(0.083)	
Log density of population in 1680					0.527***
					(0.107)
Observations	268	268	268	268	268
Adjusted R^2	0.660	0.460	0.260	0.894	0.777
Mean	4.802	0.720	3.268	5.739	6.679
Geographical Controls	Y	Y	Y	Y	Y
Political Controls	Y	Y	Y	Y	Y

Note: This table reports estimates of the association between river distance to predicted bamboo-suitable locations and early economic conditions at the beginning of the Ming dynasty (1368–1644). The dependent variable in column 1 is the log density of population in 1393; column 2 uses the caloric suitability index; and column 3 uses the log density of schools established before 1368. Geographic controls include terrain ruggedness index, log distance to the coast, log distance to the nearest navigable river, and log prefecture area. Political controls include log distance to the national capital and an indicator for provincial capital status. All regressions include province fixed effects. Standard errors, clustered at the province level (18 clusters), are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5 Contemporaneous Impact of Printing

5.1 Two-way Fixed Effects as Benchmark Regression

To assess the contemporaneous effect of woodblock printing on educational outcomes, I utilize the two-period (Ming and Qing) panel setting and estimate a two-way fixed effects (TWFE) model. This approach controls for time-invariant differences across prefectures and common temporal shocks across dynasties. The empirical specification is:

$$\ln(\text{poet density})_{it} = \theta_t + \delta_i + \beta \ln(\text{book density})_{it} + \Gamma X_{it} + \Lambda X_i \times \text{Qing} + \epsilon_{it} \tag{3}$$

Where outcome is either log density of male or female poets for prefecture $i = \{1, 2, \ldots, 268\}$ in dynasty $t = \{Ming, Qing\}$. θ_i is prefecture fixed effects, δ_t is time fixed effects. X_{it} is a group of time-varying controls including log density of population, log density of degree holders, and log density of schools. $X_i \times$ Qing is a vector of time-invariant variables including geographical controls (caloric suitability index, terrain ruggedness index, distance to coast and river, log prefecture area) and political controls (log distance to national capital and dummy indicating provincial capital) interacted with the dummy indicating the Qing period. Standard errors are corrected for spatial auto correlation across prefectures (Conley, 1999; Bester, Conley and Hansen, 2011; Colella et al., 2019).

5.2 TWFE and IV Results

Table 4 reports results from TWFE and 2SLS specifications. Columns 1–4 focus on female poet density; Columns 5–8 on male poet density. Columns 1–3 and 5–7 present TWFE estimates with progressively richer controls. Columns 4 and 8 display the 2SLS results, instrumenting book density using river distance to locations suitable for bamboo cultivation.

The 2SLS estimate for female poet density in Column 4 rises to 0.94, a substantial increase from the TWFE estimate of 0.33 in Column 3. In contrast, the 2SLS estimate for male poet density shifts to 0.23 and becomes statistically insignificant (Column 8) from TWFE estimate of 0.24 but significant. These results imply that a 10% increase in book density is associated with a 9.4% increase in female poet density but only a 2.3% (and insignificant) change in male poet density. Online Appendix Figure A5 presents the corresponding binscatter plots for Columns (4) and (8), illustrating the quasi-linear relationship between predicted book density and poet density by gender.

Table 4: Woodblock-printed Books Density and Poet Density: TWFE and 2SLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log density of female poets				**Log density of male poets**			
		TWFE		2SLS		TWFE		2SLS
Log density of books	0.454***	0.361***	0.331***	0.943***	0.310***	0.249***	0.240***	0.226
	(0.027)	(0.031)	(0.031)	(0.216)	(0.026)	(0.034)	(0.037)	(0.187)
Log density of *jinshi*	-0.096*	0.011	-0.002	-0.132*	0.425***	0.488***	0.486***	0.489***
	(0.050)	(0.043)	(0.044)	(0.079)	(0.054)	(0.060)	(0.062)	(0.072)
Log density of schools	0.681***	0.653***	0.672***	0.272	-0.025	-0.011	-0.011	-0.001
	(0.107)	(0.076)	(0.074)	(0.169)	(0.084)	(0.089)	(0.085)	(0.158)
Log density of population	-0.005	-0.065	-0.073	0.066	-0.049	-0.076	-0.086	-0.089
	(0.064)	(0.053)	(0.049)	(0.070)	(0.053)	(0.055)	(0.060)	(0.082)
Observations	536	536	536	536	536	536	536	536
R^2	0.858	0.879	0.882	0.770	0.833	0.840	0.845	0.845
Mean	1.431	1.431	1.431	1.431	2.221	2.221	2.221	2.221
First-stage F-stat				20.637				20.637
Geographical controls		Y	Y	Y		Y	Y	Y
Political controls			Y	Y			Y	Y

Note: Each column incrementally adds a set of control variables as described in the text. Geographic controls include the caloric suitability index, terrain ruggedness index, log distance to the coast, log distance to the nearest navigable river, and log prefecture area. Political controls consist of log distance to the national capital and a dummy indicating provincial capitals. All regressions include prefecture and dynasty fixed effects. Standard errors, adjusted for spatial autocorrelation at the prefecture level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Among the findings, two key observations merit further discussion. First, the 2SLS estimates for the effect of book density are significantly larger and more robust for women (Column 4) than for men (Column 8). This gender-differentiated impact suggests that woodblock printing, while ostensibly a neutral technology, disproportionately benefited female educational attainment. I attribute this female-favoring effect to the substitutive role of woodblock printing for formal schooling. Consequently, women, who were excluded from traditional educational institutions, derived greater marginal benefits from increased access to printed materials compared to men, who could leverage a wider array of external educational resources.

The second important finding from the results is the downward bias in the 2SLS estimates of book density compared to the TWFE estimates, which is observed for female poets but not for males. This suggests that some unobserved factors positively correlated with woodblock printing but negatively associated with the presence of female poets—while this unobserved factor had little correlation with presence of male poets. A plausible explanation lies in the differential impact of Confucian cultural constraints on the presence of poets by gender.

Confucian ideology, which emphasized hierarchical gender roles and the subordination of women to male authority, played an non-negligible role in shaping norms around authorship and publication in Imperial China. While male literary achievement was publicly celebrated and often linked to official careers and social prestige, women's writing was expected to remain private, modest, and confined within the domestic sphere. In this cultural context, female authorship was often deemed inappropriate or even shameful, particularly if it risked attracting public attention. Consequently, the preservation and publication of women's literary work were subject to strict familial and societal scrutiny, resulting in a gender asymmetry in literary representation.

Anecdotal evidence supports this interpretation. Shen Shanbao (1808–62), a respected female poet and anthologist from Zhejiang, lamented that many talented women were never published because their families feared that literary fame might bring shame rather than honor. In another example, Tang Qingyun (1813–32) from Jiangsu wrote over 800 poems, but her mother burned the manuscripts after her death, as she believed women's public authorship conflicted with Confucian propriety. These cases suggest that gender norms, reinforced by Confucian culture, shaped not only women's access to education but also the visibility and preservation of their intellectual contributions.

5.3 Robustness Check

Due to unavoidable potential selection issues inherent in historical data, I conduct a series of robustness checks to assess the reliability of the benchmark results presented in Table 4. These tests include modifications to variable definitions, alternative model specifications, and different estimation strategies. Full results are reported in the Online Appendix.

5.3.1 Non-Poetry Books as an Alternative Measure

A potential concern is that the inclusion of poetry books in the construction of book density may bias the estimated relationship between printing and poet density, especially if poetry titles directly stimulate literary output. To isolate the impact of printing on general literacy rather than literary skill, I re-estimate the main specification using only non-poetry woodblock-printed books. Poetry titles are defined as any title containing the Chinese characters 詩 (shi) or 詞 (ci); all other titles are classified as non-poetry.

Online Appendix Table A5 presents the results. The estimates closely mirror the benchmark results in Table 4: non-poetry book density is significantly associated with increased female poet density but has no significant effect on male poets. This pattern suggests that the observed rise in female literary participation was driven not by greater exposure to the poetry genre per se, but by broader improvements in literacy and educational access facilitated by the increasing availability of woodblock printing.

5.3.2 The Impact of Western Printing

Another concern is that the observed increase in poet density may be confounded by the influence of Western (movable-type) printing, introduced to China in the early 19th century via missionary activity. To address this, I compile data on 118 Western printing presses active between the 19th and early 20th centuries, based on historical sources compiled by Zhang, Pang and Zheng (1999). These presses—owned by Protestant and Catholic missions as well as Western commercial publishers—were located in 29 prefectures.

For each prefecture, I compute the river distance to the nearest Western printing press and re-estimate the benchmark specification using this measure as the explanatory variable. Results are reported in Online Appendix Table A6, Columns 1–3 (female poets) and 5–7 (male poets). In Columns 4 and 8, I include both Western and woodblock printing measures to jointly assess their effects.

The coefficient on river distance to Western printing is statistically insignificant when putting all control variables (Columns 3 and 6). By contrast, the coefficient on woodblock-printed book density remains positive and significant for poets (Columns 4 and 8). These findings confirm that the educational effect documented here is driven by traditional woodblock printing rather than the later Western printing technology.

5.3.3 Alternative Models—Pooled OLS, Probit and Poisson

To assess the robustness of the results to model choice, I estimate three alternative specifications: Pooled OLS, Probit, and Poisson regressions. Due to the scale-dependent nature of log transformations, interpreting "log-like" outcome variables as percentage changes has been contested, as unit changes can distort estimates and hinder interpretability (Chen and Roth, 2024). Following Chen and Roth (2024), I adopt Probit (specification 5) and Poisson (specification 6) models to estimate scale-invariant, normalized parameters. These models also allow for a distinction between the extensive margin (the probability of having any poets) and the intensive margin (the number of poets, conditional on presence). Implementing Probit and Poisson models in a fixed-effects panel setting would lead to a substantial loss of observations, as many prefectures exhibited no change in poet presence or count across the two dynasties. To preserve the full sample and maintain comparability across models, I adopt a pooled estimation framework for all three specifications.

- A Pooled OLS model, replacing prefecture fixed effects with province fixed effects, to capture broader regional variation;

$$\ln(\text{poet density})_{it} = \delta_t + \theta_p + \beta \ln(\text{book density})_{it} + \Gamma X_{it} + \epsilon_{it} \tag{4}$$

- A Probit model, where the dependent variable is a binary indicator for whether a prefecture has any recorded poets;

$$Pr(\text{Whether had a poet} = 1 \mid X_{it}) = \delta_t + \theta_p + \beta \ln(\text{book density})_{it} + \Gamma X_{it} + \epsilon_{it} \tag{5}$$

- A Poisson model, appropriate for count data, estimating the number of poets per prefecture.

$$\text{Number of poets}_{it} = \exp\left(\delta_t + \theta_p + \beta \ln(\text{book density})_{it} + \Gamma X_{it} + \epsilon_{it}\right) \tag{6}$$

Where prefecture $i = \{1, 2, \cdots, 268\}$, dynasty $t = \{Ming, Qing\}$, and province $p = \{1, 2, \cdots, 18\}$. Control variables X_{it} remain as in the benchmark analysis as indicated in Specification 3, including both time-varying and time-invariant characteristics.

Results are summarized in Online Appendix Tables A7 (Pooled OLS), A8 (Probit), and A9 (Poisson). Across all specifications, book density is significantly associated with greater female poet representation—whether measured by density, probability, or count. The results for male poets remain smaller and often statistically insignificant, with some specifications even producing negative coefficients.

Taken together, these robustness checks confirm the central finding: the greater availability of woodblock printing had a disproportionately positive effect on women's educational outcomes, and this pattern holds across different model specifications, measures of book content, and historical controls.

6 Modern Impact of Woodblock Printing

6.1 Cohort-analysis Using 1982 Census

To investigate the long-run impact of historical woodblock printing on modern educational outcomes, I construct a cohort-prefecture-level dataset covering 252 modern prefectures—defined by 1982 administrative boundaries—across 15 five-year birth cohorts born from 1898 to 1972. These units are matched to historical prefectures from 1820 by assigning each modern prefecture the identifier of the 1820 unit whose administrative seat lies within its boundaries.[13]

I employ the same IV strategy—river distance to locations suitable for bamboo cultivation—in the cohort-level analysis. The 2SLS estimation specifications are as follows:

$$\ln(\text{book density})_i = \ln(\text{river distance to predicted bamboo})_i + X_i + \epsilon_i \tag{7}$$

$$Y_{ic} = \sum_{c=1898-1902}^{1968-1972} \beta_c \ln(\widehat{\text{book density}})_i + \sum_{c=1898-1902}^{1968-1972} \Lambda_c X_i + \epsilon_{ic} \tag{8}$$

Let $i = \{1, 2, \ldots, 252\}$ index modern prefectures and $c = \{1, 2, \ldots, 15\}$ index five-year birth cohorts. The key explanatory variable is the log density of woodblock-printed books in the Ming–Qing period. Book density is calculated as the total number of woodblock-printed books (plus one) published in prefecture i, normalized by its average historical population during Ming-Qing period.

The baseline controls X_i include: (i) log population density and school density during the Ming–Qing period (normalized using the same method as book density;) (ii) geographic characteristics—caloric suitability index, terrain ruggedness index, distances to navigable rivers and the coast, and log prefecture land area; and (iii) political variables—log distance to the national capital and dummy indicating historical provincial capital. All regressions are weighted by cohort size to account for regional and temporal variation in population structure. Standard errors are corrected for spatial auto correlation across prefectures.

[13] If no historical prefectural seat falls within a modern unit, I assign the 1820 prefecture covering the largest share of land area. Historical variables are merged one-to-one. For modern prefectures containing multiple historical seats, I aggregate count variables (e.g., books, population, schools) and average continuous variables (e.g., terrain ruggedness, caloric index, distance). This ensures a consistent spatial alignment across datasets.

The first-stage results using specification 7 (Online Appendix Table A10) confirm a strong and significant correlation between river distance to location suitable for bamboo cultivation and historical book presence at the cross-sectional level. A first-stage F-stat exceeding 10 meets the conventional threshold (Stock and Yogo, 2002), indicating that the instrument is not weak.

I focus on two main outcome variables Y_{ic}. The first is relative literacy by gender, defined as the literacy rate in prefecture i and cohort c divided by the national cohort average. This normalization enables meaningful comparisons across cohorts by adjusting for secular trends in literacy. The coefficient β thus captures percentage deviations from the national cohort-level mean, allowing for a more interpretable measure of literacy advancement—particularly in early cohorts with low absolute levels.

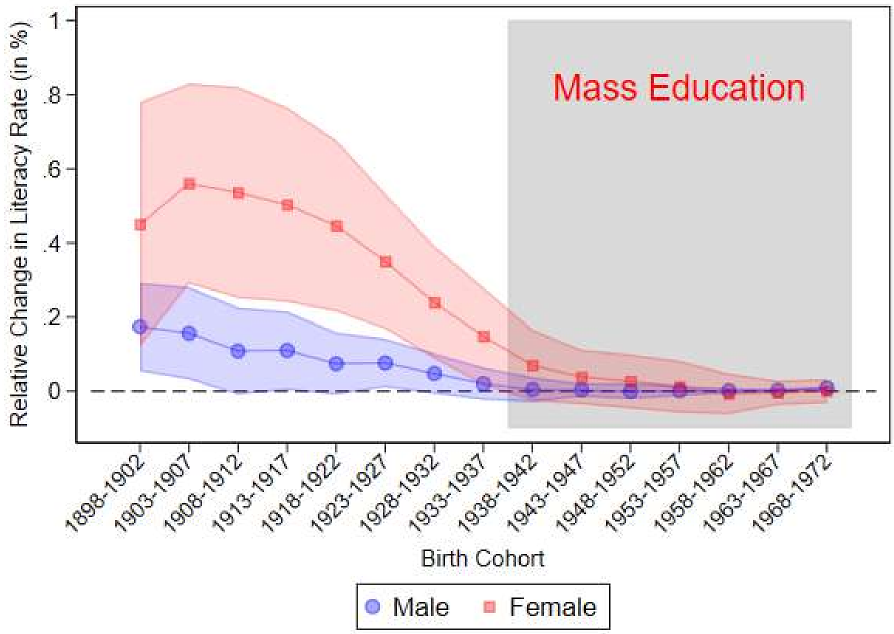
The second outcome is the ratio of female-to-male literacy rate within each prefecture-cohort cell. This provides a direct measure of gender equality in educational attainment at the primary level, with values closer to one indicating greater parity. Given women's historically limited access to formal education, this variable serves as a proxy for the inclusiveness of local educational institutions and cultural attitudes toward female literacy.

6.2 Interpretation of Results

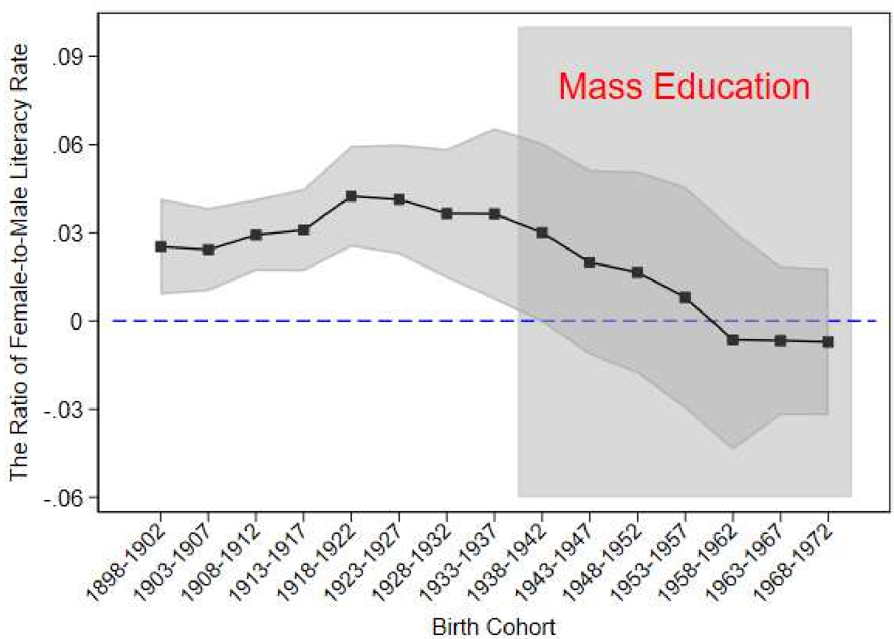
Figure 4 presents the 2SLS estimates of the impact of historical woodblock printing (measured by the log density of woodblock-printed books) on modern educational outcomes. Panel (a) demonstrates the effect of printing on relative literacy gains by gender, and panel (b) illustrates the effect of printing on the ratio of female-to-male literacy rate across the fifteen birth cohorts. For comparation, OLS estimates are provided in Online Appendix Figure A6; the subsequent discussion focuses on the interpretation of the 2SLS estimates.

This analysis yields two key findings. First, Panel (a) shows that among pre-reform cohorts, women in regions with greater historical availability of woodblock printing experienced larger relative gains in literacy, with this correlation substantially stronger than that for their male counterparts. These results suggest that, prior to the Mass Education Movement, women still derived greater marginal benefits from the historical availability of woodblock printing for literacy acquisition compared to men. Panel (b) further demonstrates that this female-biased effect translated into improved gender equality in literacy (as measured by the female-to-male literacy ratio), again only for pre-reform cohorts.

Second, woodblock printing had a significant impact exclusively on pre-reform cohorts, with



(a) Relative Change in Literacy Rate by Gender



(b) The Ratio of Female-to-Male Literacy Rate

Figure 4: The 2SLS Estimates of Printing on Modern Educational Outcomes

Note: (a) 2SLS estimates of the effect of woodblock-printed book density on the relative change in literacy rate (in percentage terms). (b) 2SLS estimates of the effect of woodblock-printed book density on the female-to-male literacy rate ratio. Shaded areas around the point estimates represent 95% confidence intervals. The grey region indicates cohorts exposed to the Mass Education Movement. Regressions control for all historical covariates listed in Specification 8. Observations are weighted by cohort population size. Standard errors are adjusted for spatial autocorrelation at the prefecture level.

no observable effect for reform-era cohorts. This indicates that the influence of historical woodblock printing diminished as universal, state-led educational reforms expanded access and reduced pre-existing gender disparities in literacy.

The first observation highlights the persistent female-biased impact of historical woodblock printing on educational outcomes, even amid formal institutional transformations. While the Republican regime formally endorsed women's right to education, women in regions with greater historical access to woodblock-printed books remained significantly more likely to attain literacy. This suggests that some informal "rules of the game" continued to shape educational outcomes after changes in formal institution (i.e., codified law). These findings are consistent with broader research on the discussion of culture and institution, which shows that deeply rooted gender biases can continue to influence outcomes even as economies evolve (e.g., Alesina and Giuliano, 2015).

The second key observation concerns the diminished role of woodblock printing following the Maoist revolution. Unlike the Republican-era reforms, which were limited in scope and largely concentrated in urban centers, Maoist education policies were comprehensive, compulsory, and penetrated deeply into rural areas (Karl, 2020, p.25, 165). The efficacy of these educational policies is evident in the rapidly narrowing gender gap in literacy for reform-era cohorts, as shown in Figure 2. Consequently, while woodblock printing served to mitigate gender gaps in education prior to the Mao era, the subsequent revolutionary reforms diminished its relative importance by fundamentally altering the institutional landscape and directly guaranteeing universal access to primary education.

Online Appendix Figure A7 presents 2SLS estimates for different levels of educational attainment (middle and high school), and Online Appendix Figure A8 examines heterogeneous effects by urban versus rural populations. Overall, the female-biased impact was more pronounced at lower levels of education and among urban populations. This pattern aligns with the historical context: primary education was more representative of women's educational opportunities during this period, and Republican-era educational institutions primarily targeted urban elites.

6.3 Did Woodblock Printing Change Culture?

This section investigates whether historical woodblock printing engendered a cultural shift in parental investment in children's education. Motivated by the persistent female-biased effect of printing despite formal institutional change from Imperial to Republican China, I test the hypothesis that greater historical availability of woodblock printing facilitated the development of more progressive culture norms surrounding female education. This cultural shift is posited to have increased parental investment in daughters' human capital, thereby mitigating the gender gap in education during the pre-reform period.

I employ a triple-difference (DDD) framework to test this hypothesis. The empirical strategy leverages variation across three dimensions:

1. Region: "Treatment" prefectures (those with predicted bamboo, $T = 1$) vs. "Control" prefectures (those without, $T = 0$).

2. Child's gender: Daughters ($F = 1$) vs. Sons ($F = 0$).

3. Parental literacy: Households with at least one literate parent ($L = 1$) vs. those with illiterate parents ($L = 0$).

The triple-difference estimator isolates this effect by comparing the female-male educational gap between literate and illiterate households across treatment and control regions. Simplified estimating equation 9 is specified as follows. All variables in this basic setup are dummy variables, and no additional control variables added.

$$\begin{aligned} Y_{ihp} = \beta_0 + \beta_1 T_{ihp} + \beta_2 F_{ihp} + \beta_3 L_{ihp} + \beta_4 T_{ihp} \times F_{ihp} + \beta_5 T_{ihp} \times L_{ihp} + \beta_6 F_{ihp} \times L_{ihp} \\ + \beta_7 T_{ihp} \times F_{ihp} \times L_{ihp} + \epsilon_{isp} \end{aligned} \tag{9}$$

With Y_{ihp} being the outcome variable of interest corresponding to child i in household h and prefecture p, and T_{ihp} being an indicator for being in a "treated" cohort (p is bamboo prefecture), and F_{ihp} being an indicator for being female, and L_{ihp} is an indicator for household with literate parent. β_7 is the DDD estimator.

Table A4 provides the full results of the triple-difference framework; and to facilitate interpretation, I also plot the two DiD estimates in Figure 5. It is important to note that these results should be interpreted as highly suggestive, given the low sample retention rate inherent in the stringent criteria required for this intergenerational analysis. The sample is restricted to pre-reform cohorts (respondents aged 45 and above in the 1982 census) to ensure no exposure to the Mass Education Movement. Furthermore, parental literacy is defined stringently: households are classified as having "illiterate parents" only if both the father and mother were illiterate, and as having "literate parent(s)" if at least one parent was literate. To mitigate some sample size concerns, I incorporate data from the Fourth National

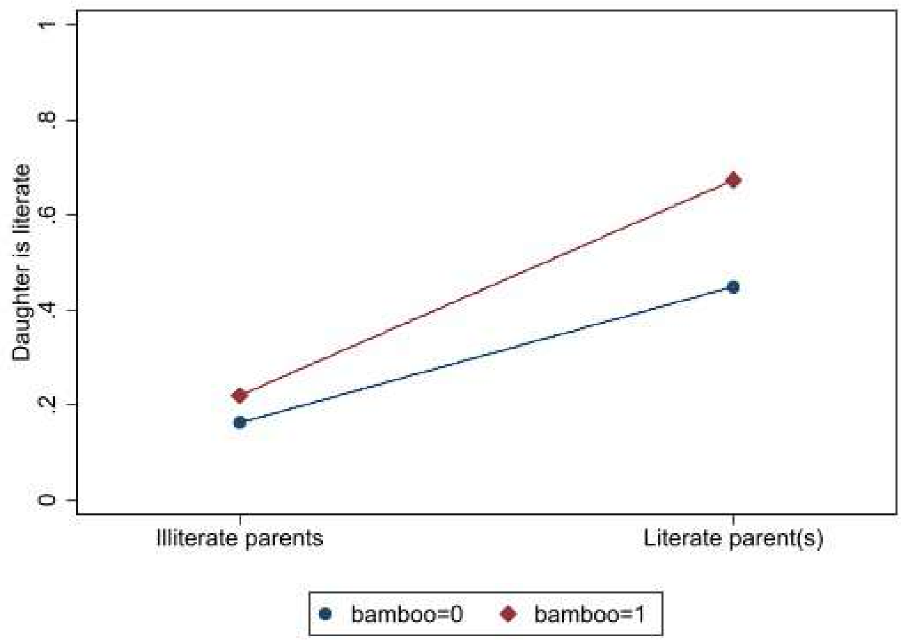
Table 5: DDD Estimates of the Impact of Bamboo on Intergenerational Transmission of Literacy

Location/Parent literacy	Illiterate	Literate	Family difference for location
A. Treatment Individuals: Women			
Bamboo prefectures	0.220	0.674	0.454***
	(0.059)	(0.024)	(0.070)
	[50]	[371]	
Non-bamboo prefectures	0.163	0.450	0.287***
	(0.018)	(0.014)	(0.026)
	[443]	[1288]	
Location difference for a given family	0.057	0.224***	
	(0.056)	(0.029)	
Difference-in-difference	0.167**		
	(0.075)		
B. Control Individuals: Men			
Bamboo prefectures	0.603	0.893	0.289***
	(0.025)	(0.008)	(0.021)
	[388]	[1275]	
Non-bamboo prefectures	0.520	0.817	0.297***
	(0.009)	(0.005)	(0.010)
	[2914]	[5179]	
Location difference for a given family	0.083***	0.075***	
	(0.027)	(0.012)	
Difference-in-difference	-0.008		
	(0.026)		
DDD	**0.175****		
	(0.074)		

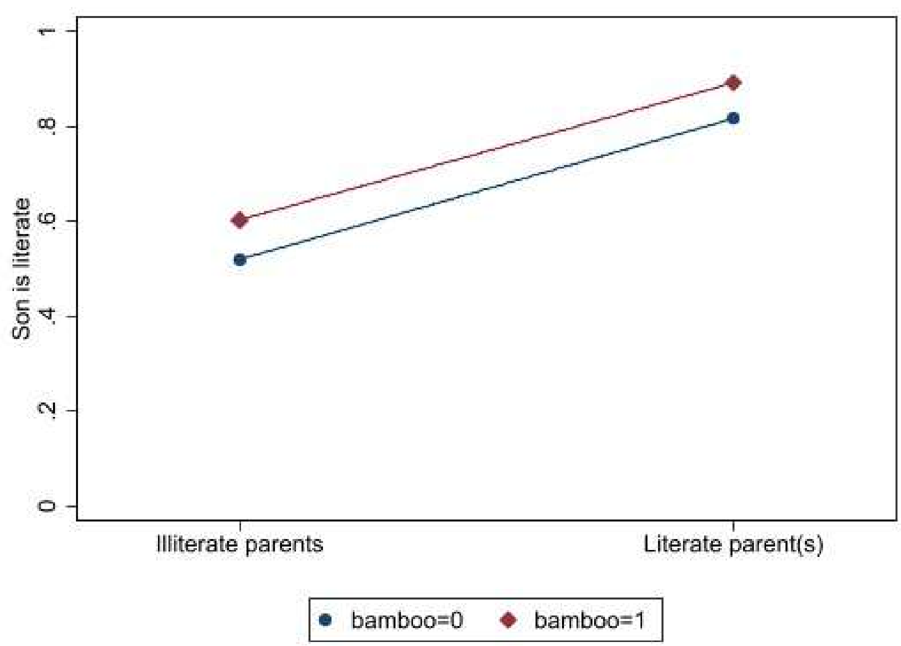
Note: This table reports all derived difference-in-differences (DiD) estimates and the derived triple-difference (DDD) estimate. Standard errors are shown in parentheses; the number of observations is indicated in brackets.
Source: The Third National Population Census 1982.

Population Census 1990. While this expansion did not substantially increase the sample size, the estimated effects remained consistent (Table A11), bolstering confidence in the results.

I focus the interpretation on three key estimates from Table A4: two difference-in-differences (DiD) estimates and the triple-difference (DDD) coefficient. The first DiD estimate indicates that a woman born in a bamboo-suitable prefecture was 16.7 percentage points more likely to be literate if she had at least one literate parent, compared to a similar woman born in a non-bamboo prefecture. This suggests a stronger intergenerational transmission of literacy



(a) DiD Estimates for Daughters



(b) DiD Estimates for Sons

Figure 5: DiD Estimates of the Impact of Bamboo on Intergenerational Transmission of Literacy

Note: Panel (a) demonstrates the DiD estimates of the impact of treatment (bamboo to proxy woodblock printing) on intergenerational transmission of literacy for daughters. Panel (b) demonstrates the DiD estimates for sons.

for daughters in regions with a historical availability of woodblock printing. In contrast, the corresponding effect for sons is -0.8 percentage point and statistically insignificant. This suggests that there is no significant cultural difference in transmission of literacy for sons in or not in woodblock printing regions, for whom primary educational investment was already a stable norm.

The core DDD estimator confirms that the difference in the parental literacy effect between women and men is itself significantly larger in treatment regions than in control regions. The results indicate that residing in a treatment prefecture increased the likelihood of literacy for daughters (relative to sons) by an additional 17.5 percentage points in households with at least one literate parent (compared to illiterate households). This pattern reflects a distinct shift in cultural attitudes toward educational investment in daughters relative to sons. Consequently, women in historically print-rich regions faced fewer disadvantages relative to men in terms of intra-household educational allocation. This mechanism—whereby woodblock printing fostered progressive cultural norms around gender and education—helps explain the persistence of its female-biased impact throughout the pre-reform era.

This represents a uniquely different cultural environment regarding daughter-specific educational investment in historical printing centers, resulting in that women in historically print-rich regions faced fewer disadvantages relative to men in terms of intra-household educational culture. This mechanism helps explain the persistence of the female-biased impact of woodblock printing in the pre-reform era. Figure A9 provides further evidence of this cultural shift, displaying literacy rates for females and males in treatment (bamboo-suitable) and control prefectures. The figure shows a growing disparity in female literacy rates between bamboo and non-bamboo regions over time for pre-reform cohorts, while the gap in male literacy rates remained relatively stable across these same cohorts.

These findings speak to the literature on the role of culture in shaping educational outcomes (e.g., Bau, 2021; Ashraf et al., 2020), while offering a novel extension: it demonstrates that cultural norms are themselves endogenously shaped by the broader technological and institutional context. Overall, these results underscore the complex, recursive interplay between technology, culture, and formal institutions. They contribute new evidence to a vibrant field of study examining how institutional environments mediate the social effects of technological change and, conversely, how technological advancements can reshape institutional and cultural landscapes.

7 Conclusion

This study examines the a seemingly gender-neutral technology–woodblock printing–and its long-run gender-biased impacts on educational outcomes in China. By interacting with gender-asymmetric institutions, particularly the exclusion of women from formal schooling, woodblock printing generated persistent female-biased impacts on educational outcome due to its substitutive role for formal schooling.

Using historical data on poet presence and cohort-level literacy rates from the 1982 Census, I trace both the origins and persistence of this gendered effect across institutional regimes. To address endogeneity concerns related to education demand, I employ an instrumental variable strategy that exploits river distance to ecologically suitable bamboo locations—key inputs in book production—as an exogenous source of variation in the woodblock printing supply.

The results reveal that higher density of woodblock-printed books is strongly associated with increases in female poet density, while the effects on male poet density are smaller and less robust. This asymmetry reflects institutional constraints that restricted women's access to formal schooling, making home-based learning via books a relatively more important channel for female literacy accumulation. In contrast, men could rely on broader educational infrastructure, diluting the marginal effect of availability of woodblock printing.

The female-biased impact of printing persisted into the twentieth century. 2SLS estimates using modern literacy data show that historical woodblock printing significantly raised relative female literacy rates and improved gender parity in literacy—especially for cohorts who received primary education before the Mass Education Movement. This persistence can be partly explained by an evolving culture of investment in daughters' education, as suggested by a DDD estimator that leverages variation across parental literacy, historically bamboo-suitable, and child sex. These results underscore the potential cultural shift facilitated by the historical proliferation of woodblock printing.

However, this gendered effect declined sharply with the expansion of universal public schooling under Maoist reforms. The state-led Mass Education Movement—compulsory and universal in scope—effectively neutralizing the cultural and historical asymmetries that had previously amplified the impact of woodblock printing for women.

Taken together, these findings underscore how technologies—even those without an inherent gender bias—can generate long-lasting, gender-biased impacts when mediated by unequal institutional contexts. This study contributes to the broader literature on the interplay between technology and institutions, both formal and informal.

References

Alesina, Alberto, and Paola Giuliano. 2015. "Culture and institutions." *Journal of economic literature*, 53(4): 898–944.

Alesina, Alberto, Paola Giuliano, and Nathan Nunn. 2013. "On the origins of gender roles: Women and the plough." *The quarterly journal of economics*, 128(2): 469–530.

Alford, William P. 1995. *To steal a book is an elegant offense: Intellectual property law in Chinese civilization.* Stanford University Press.

Angeles, Luis. 2017. "The great divergence and the economics of printing." *The Economic History Review*, 70(1): 30–51.

Ashraf, Nava, Natalie Bau, Nathan Nunn, and Alessandra Voena. 2020. "Bride price and female education." *Journal of Political Economy*, 128(2): 591–641.

Bailey, Paul J. 2007. *Gender and education in China: Gender discourses and women's schooling in the early twentieth century.* Routledge.

Barrett, Timothy Hugh. 2008. *The woman who discovered printing.* Yale University Press.

Bau, Natalie. 2021. "Can policy change culture? Government pension plans and traditional kinship practices." *American Economic Review*, 111(6): 1880–1917.

Bertocchi, Graziella, and Monica Bozzano. 2016. "Women, medieval commerce, and the education gender gap." *Journal of Comparative Economics*, 44(3): 496–521.

Bester, C Alan, Timothy G Conley, and Christian B Hansen. 2011. "Inference with dependent data using cluster covariance estimators." *Journal of Econometrics*, 165(2): 137–151.

Boerner, Lars, Jared Rubin, and Battista Severgnini. 2021. "A time to print, a time to reform." *European Economic Review*, 138: 103826.

Brokaw, Cynthia J. 2007*a*. *Commerce in Culture: The Sibao Book Trade in the Qing and Republican Periods.* Harvard University Asia Center Publications Program.

Brokaw, Cynthia Joanne. 2007*b*. "Book history in premodern China: The state of the discipline I." *Book History*, 10(1): 253–290.

Bühler, Mathias, Leonhard Vollmer, and Johannes Wimmer. 2024. "Female education and social change." *Journal of Economic Growth*, 29(1): 79–119.

Cagé, Julia, and Valeria Rueda. 2016. "The long-term effects of the printing press in sub-Saharan Africa." *American Economic Journal: Applied Economics*, 8(3): 69–99.

Cai, Zong-qi. 2008. *How to Read Chinese Poetry: A Guided Anthology.* Columbia University Press.

Cao, Shuji. 2024. "The Population History of China (1368–1953)." *The Population History of China (1368–1953).*

CBDB. 2019. "China Biographical Database (CBDB)."

Chen, Jiafeng, and Jonathan Roth. 2024. "Logs with zeros? Some problems and solutions." *The Quarterly Journal of Economics*, 139(2): 891–936.

Chen, Ting, James Kai-sing Kung, and Chicheng Ma. 2020. "Long live Keju! The persistent effects of China's civil examination system." *The economic journal*, 130(631): 2030–2064.

CHGIS. 2016. "Version: 6."

Chia, Lucille. 2020. *Printing for profit: the commercial publishers of Jianyang, Fujian (11th-17th centuries).* Vol. 56, BRILL.

Colella, Fabrizio, Rafael Lalive, Seyhun Orcan Sakalli, and Mathias Thoenig. 2019. "Inference with arbitrary clustering."

Conley, Timothy G. 1999. "GMM estimation with cross sectional dependence." *Journal of econometrics*, 92(1): 1–45.

Deng, Xiaonan, and Christian Lamouroux. 2005. "The" Ancestors' Family Instructions": Authority and Sovereignty in Song China." *Journal of Song-Yuan Studies*, , (35): 79–97.

Dittmar, Jeremiah. 2011. "Information technology and economic change: the impact of the printing press." *The Quarterly Journal of Economics*, 126(3): 1133–1172.

Dittmar, Jeremiah, and Ralf Meisenzahl. 2020. "Public goods institutions, human capital, and growth: Evidence from German history." *The Review of Economic Studies*, 87(2): 959–996.

Duflo, Esther. 2012. "Women empowerment and economic development." *Journal of Economic literature*, 50(4): 1051–79.

Du, Xinfu, and Tongshu Du. 2001. "Comprehensive Study of Ming Dynasty Printed Books by Province and County: Jiangsu Volume."

Eisenstein, Elizabeth. 1979. *The printing press as an agent of change.* Cambridge, Cambridge University Press.

Elman, Benjamin A. 2000. *A cultural history of civil examinations in late imperial China.* Univ of California Press.

Farrelly, David. 1996. *The book of bamboo: a comprehensive guide to this remarkable plant, its uses, and its history.* University of California Press.

Fernández, Raquel. 2013. "Cultural change as learning: The evolution of female labor force participation over a century." *American Economic Review*, 103(1): 472–500.

Fong, Grace S. 2017. *Herself an Author: Gender, Agency, and Writing in Late Imperial China.* University of Hawaii Press.

Galor, Oded, and Ömer Özak. 2016. "The agricultural origins of time preference." *American economic review*, 106(10): 3064–3103.

Gao, Pei. 2015. "Risen from chaos: the development of modern education in China, 1905-1948." PhD diss. London School of Economics and Political Science.

Goldin, Claudia. 2006. "The quiet revolution that transformed women's employment, education, and family." *American economic review*, 96(2): 1–21.

Karl, Rebecca E. 2020. *Mao Zedong and China in the twentieth-century world: A concise history.* Duke University Press.

Knapp, Keith N. 2005. *Selfless offspring: filial children and social order in medieval China.* University of Hawaii Press.

Ko, Dorothy. 1995. *Teachers of the Inner Chambers: Women and Culture in Seventeenth-Century China.* Stanford University Press.

Ko, Dorothy. 2005. *Cinderella's sisters: A revisionist history of footbinding.* Univ of California Press.

Liu, Weiyi, Chaomao Hui, Fang Wang, Meng Wang, and Guanglu Liu. 2018. "Review of the Resources and Utilization of Bamboo in China." *Bamboo-current and future prospects*, 16(4): 133–142.

Mann, Susan. 1997. *Precious records: Women in China's long eighteenth century.* Stanford University Press.

Matsuura, Akira. 2009. *Research on the History of Inland Water Transport in the Qing Dynasty:* 清ロ内河水㷺の研究. Kansai University Press: Osaka.

McDermott, Joseph P. 2006. *A social history of the Chinese book: books and literati culture in late imperial China.* Vol. 1, Hong Kong University Press.

Nachtergaele, Freddy, Harrij van Velthuizen, Luc Verelst, Dave Wiberg, Matieu Henry, Frederica Chiozza, Yusuf Yigini, Ece Aksoy, Niels Batjes, Enoch Boateng, et al. 2023. *Harmonized world soil database version 2.0.* Food and Agriculture Organization of the United Nations.

National Library of China. 2024. "National Census of Ancient Book." *http://202.96.31.78/xlsworkbench/publish*.

Needham, Joseph. 1974. *Science and civilisation in China.* Vol. 5, Cambridge University Press.

Nekoei, Arash, and Fabian Sinn. 2021. "HERSTORY: The rise of self-made women."

New, Mark, David Lister, Mike Hulme, and Ian Makin. 2002. "A high-resolution data set of surface climate over global land areas." *Climate research*, 21(1): 1–25.

Nunn, Nathan, and Diego Puga. 2012. "Ruggedness: The blessing of bad geography in Africa." *Review of Economics and Statistics*, 94(1): 20–36.

Paglayan, Agustina S. 2021. "The non-democratic roots of mass education: evidence from 200 years." *American Political Science Review*, 115(1): 179–198.

Palma, Nuno, and Jaime Reis. 2021. "Can autocracy promote literacy? Evidence from a cultural alignment success story." *Journal of Economic Behavior & Organization*, 186: 412–436.

Population, and State Statistical Bureau Statistics Department. 1982. "Instructions for Filling Out the Questionnaire of the Third National Population Census." *https://microdata.worldbank.org/index.php/catalog/461/download/13839*.

Qian, Nancy. 2008. "Missing women and the price of tea in China: The effect of sex-specific earnings on sex imbalance." *The Quarterly Journal of Economics*, 123(3): 1251–1285.

Rawski, Evelyn Sakakida. 1979. "Education and popular literacy in Ch'ing China." *(No Title)*.

Rubin, Jared. 2014. "Printing and Protestants: an empirical test of the role of printing in the Reformation." *Review of Economics and Statistics*, 96(2): 270–286.

Ruggles, Steven, Lara Cleveland, Rodrigo Lovaton, Sula Sarkar, Matthew Sobek, Derek Burk, Dan Ehrlich, Quinn Heimann, Jane Lee, and Nate Merrill. 2025. "IPUMS, International: Version 7.6 [dataset]." *https://doi.org/10.18128/D020.V7.6*.

Serafinelli, Michel, and Guido Tabellini. 2022. "Creativity over time and space: A historical analysis of European cities." *Journal of Economic Growth*, 27(1): 1–43.

Squicciarini, Mara P. 2020. "Devotion and development: Religiosity, education, and economic progress in nineteenth-century France." *American Economic Review*, 110(11): 3454–3491.

Steinberg, Sigfrid Henry. 2017. *Five hundred years of printing.* Courier Dover Publications.

Stock, James H, and Motohiro Yogo. 2002. "Testing for weak instruments in linear IV regression."

Tsien, Tsuen-Hsuin. 1985. "Paper and Printing, vol. 5, part 1 of Science and Civilisation in China, ed. Joseph Needham." *Cambridge: Cambridge University Press*, 278: 203–60.

Twitchett, Denis Crispin, et al. 1983. "Printing and publishing in medieval China." *(No Title).*

University, McGill. 2025. "Ming Qing Women's Writings." *https://digital.library.mcgill.ca/mingqing/english/index.php*.

USGS. 2015. "Digital Elevation Models." *http://www.usa.gov/publicdomain/label/1.0/*.

Wilkinson, Endymion Porter. 2012. "Chinese history: A new manual." *(No Title).*

Xue, Melanie Meng. 2024. "High-value work and the rise of women: The cotton revolution and gender equality in China." *Available at SSRN 2389218.*

Zhang, Shudong, Duoyi Pang, and Rusi Zheng. 1999. *The General History of Printing in China.* Vol. 1, Beijing Shi : Yin Shua Gong Ye Chu Ban She.

Zhang, Xiumin, and Qi Han. 2006. "The General History of Chinese Printing." *(No Title).*

Zhu, Baojiong, and Peilin Xie. 1980. "Ming-Qing Jinshi Timing Beilu Suoyin."

Zhu, Jiahua. 1948. "Dierci Jiaoyu Nianjian [The Second Education Report](Vol. 15)."

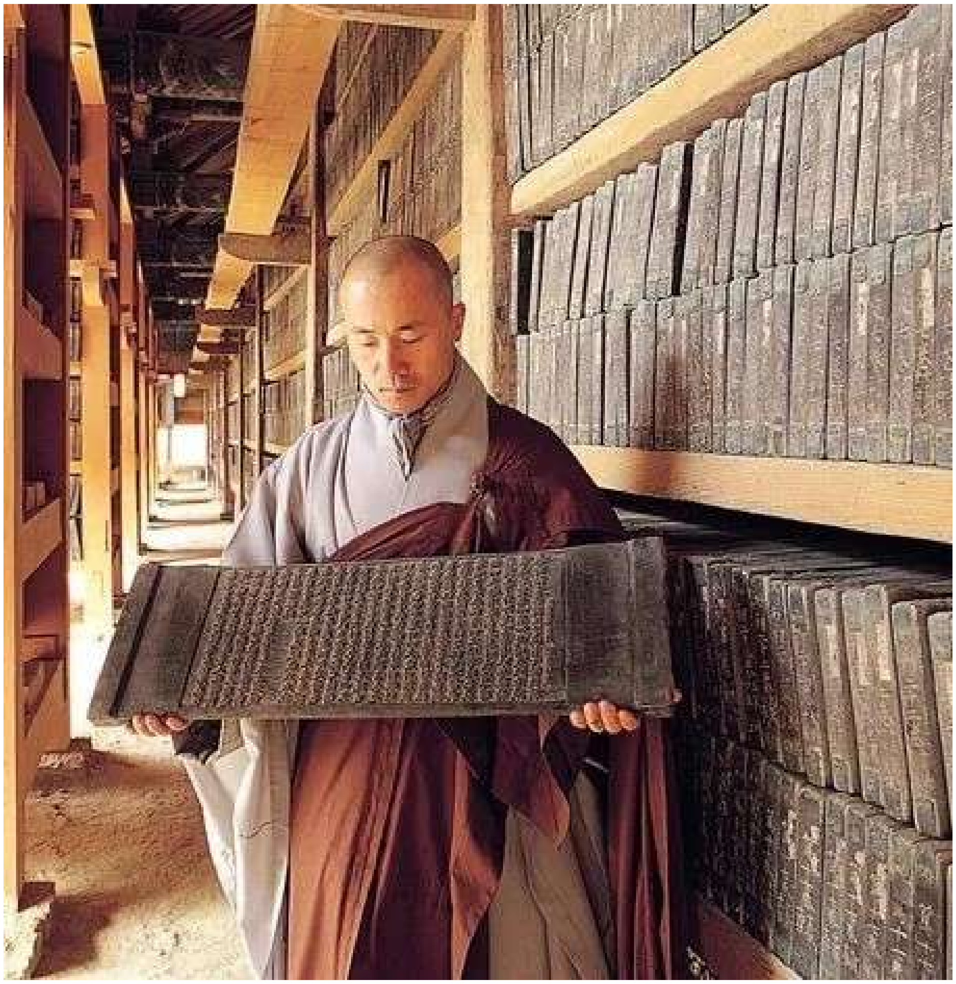


Figure A1: The Illustration of Chinese Woodblock

Note: The image depicts a monk in traditional robes holding a large rectangular woodblock engraved with text, standing amid shelves filled with similar blocks in what appears to be a historical archive or library in a temple. These woodblocks, once used for printing, are remarkably durable but occupy substantial physical space, reflecting both their long-term preservation and the spatial demands of the technology. The extensive storage requirements also highlight the immobility of woodblock printing workshops, which were often anchored in fixed locations due to the logistical challenges of maintaining and transporting such materials.

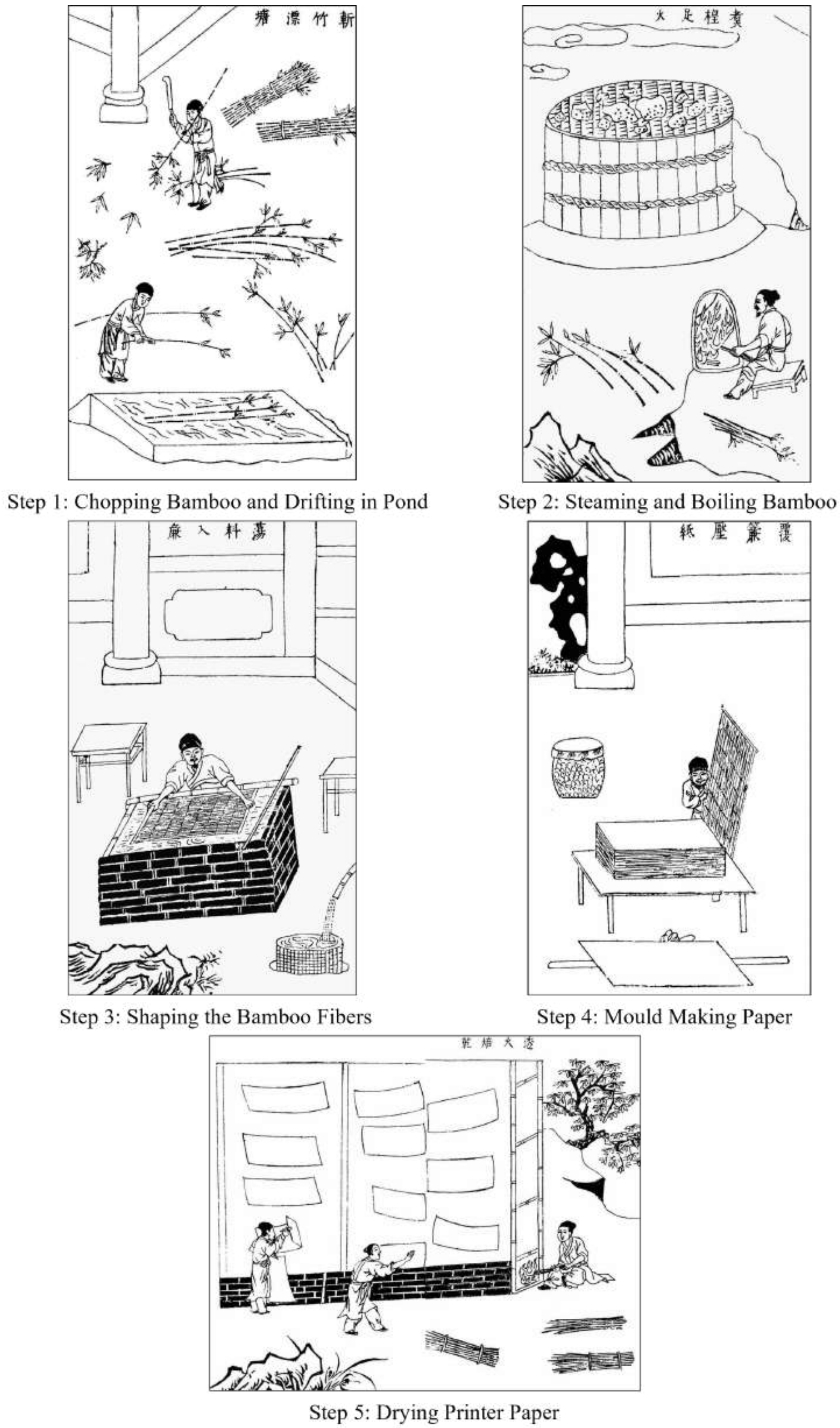


Figure A2: The Illustration of Papermaking using Bamboo

Note: This illustration is taken from the Tiangong Kaiwu (The Exploitation of the Works of Nature), published in 1637. It is the first known manual to document the use of bamboo as the primary raw material for making printer paper. Its publication near the end of the Ming Dynasty (1368–1644) suggests that the adoption of bamboo paper began during the Ming period and was subsequently expanded and widely used during the Qing Dynasty.

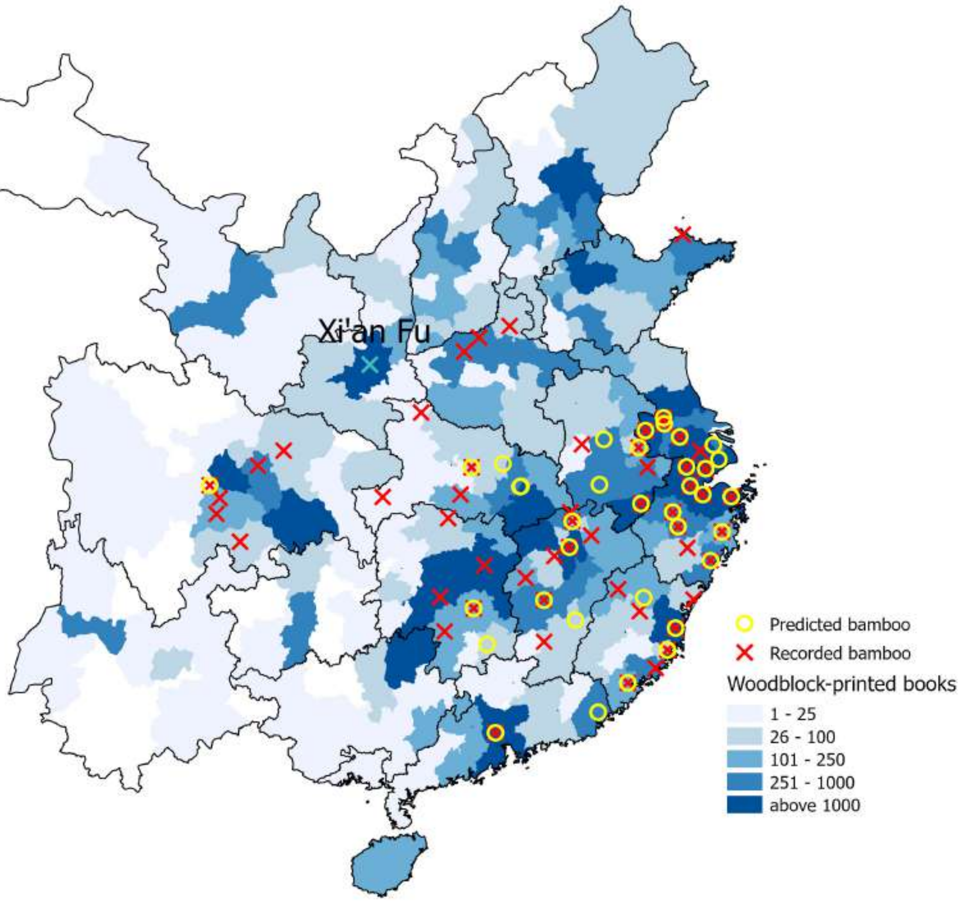


Figure A3: The Distribution of Recorded and Predicted Bamboo, and Woodblock-printed Books

Note: This figure displays the distribution of 36 predicted bamboo prefectures (predicted probability > 0.5), 55 recorded bamboo prefectures, alongside number of woodblock-printed books during the Ming-Qing period.

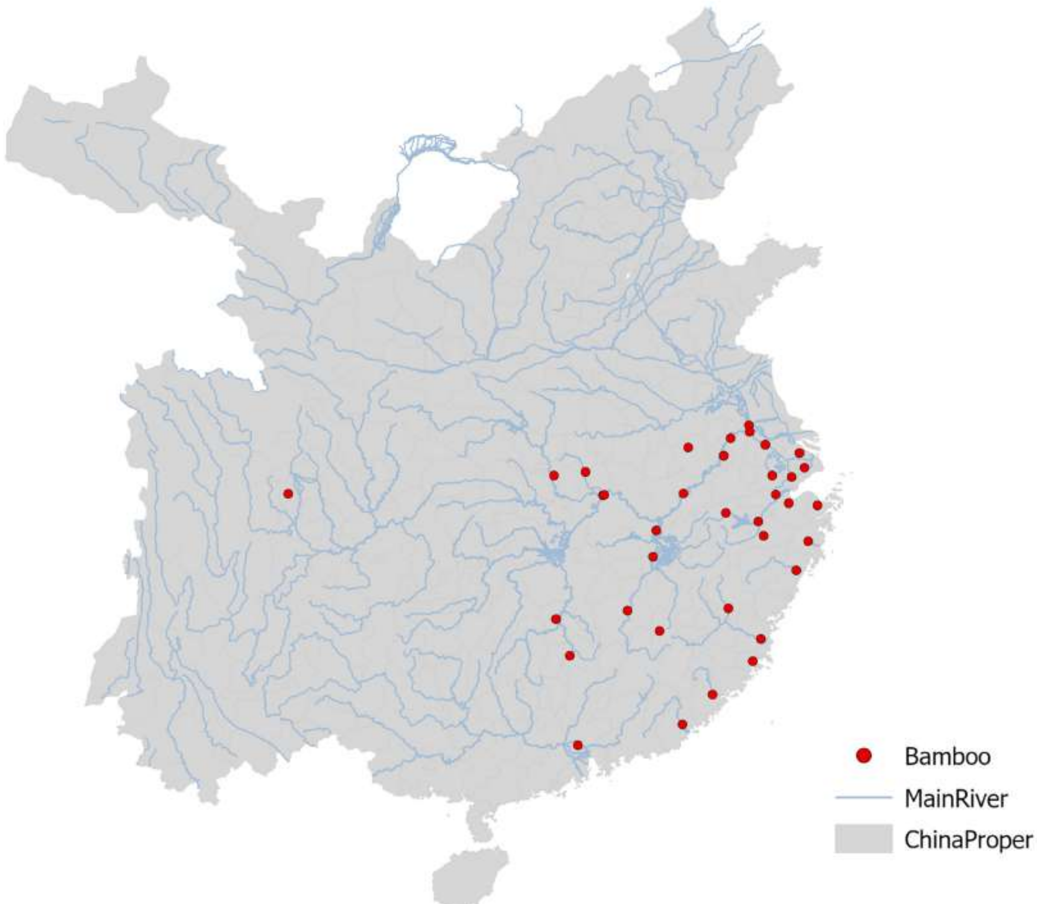


Figure A4: The Distribution of Recorded and Predicted Bamboo, and Navigable River

Note: This figure displays the distribution of 36 predicted bamboo prefectures (predicted probability > 0.5) alongside the navigable river data are sourced from Matsuura (2009).

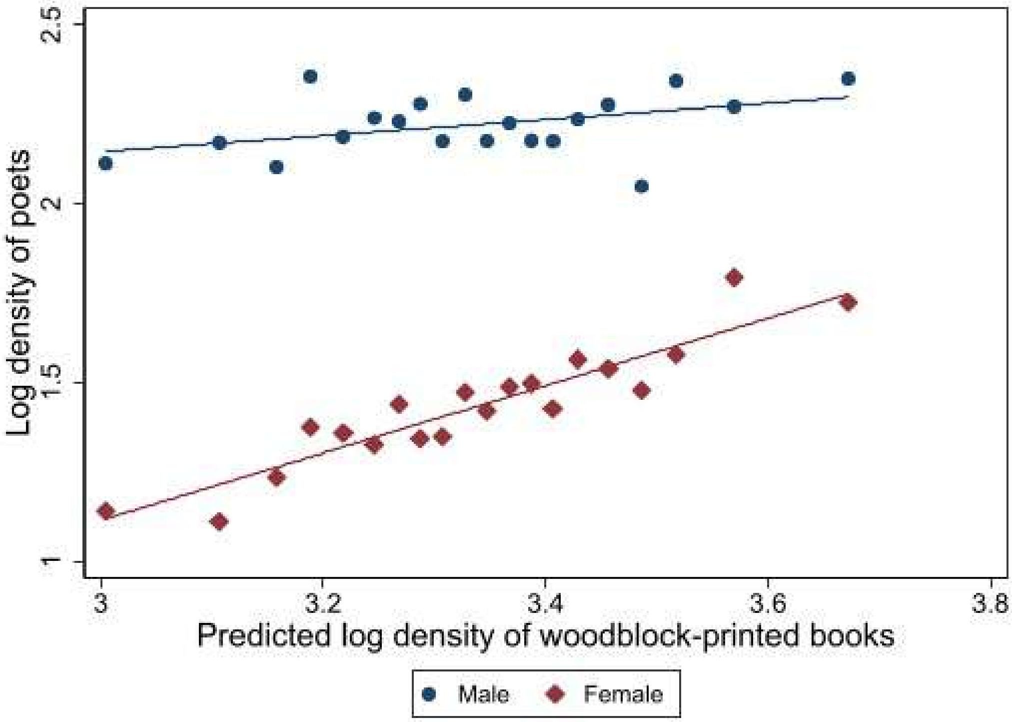
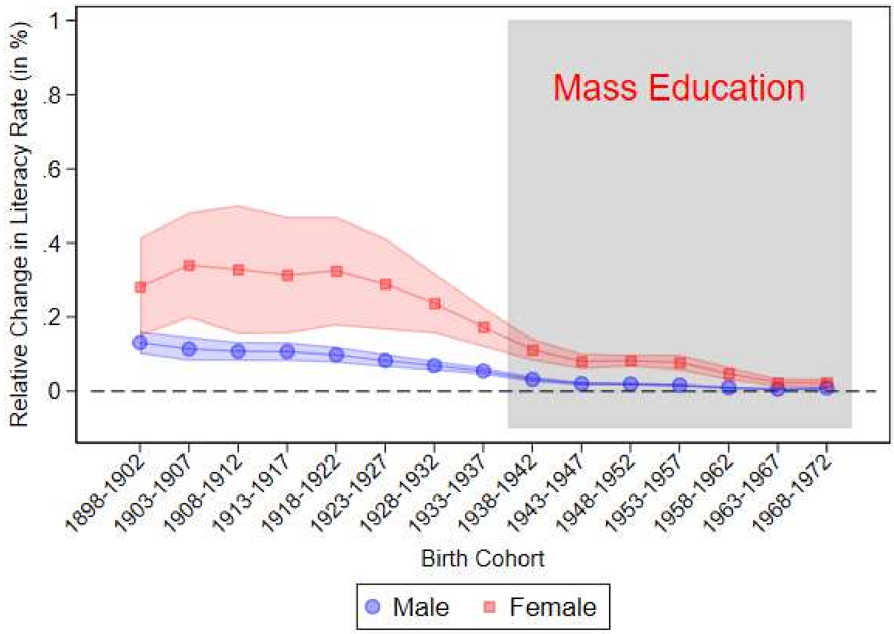
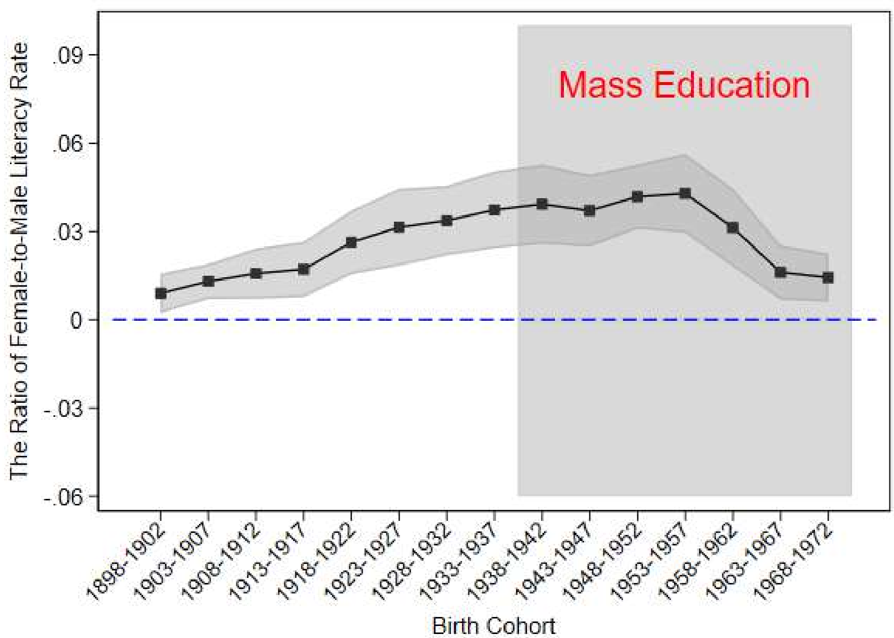


Figure A5: Binscatter Plot of Poet Density and Predicted Book Density

Note: This figure plots the log density of male and female poets against the predicted log density of books using 2SLS regression. The fitted lines illustrate the relationship separately for male (denoted as blue dot) and female (denoted as red diamond) poets. The steeper slope for female poets suggests a stronger association between predicted woodblock-printed books and female poet presence.



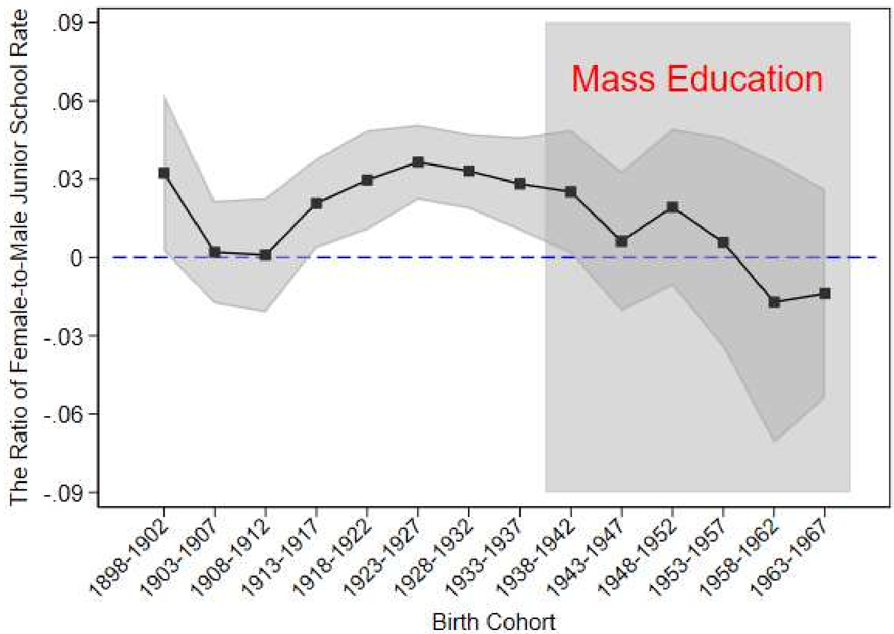
(a) Relative Change in Literacy Rate by Gender



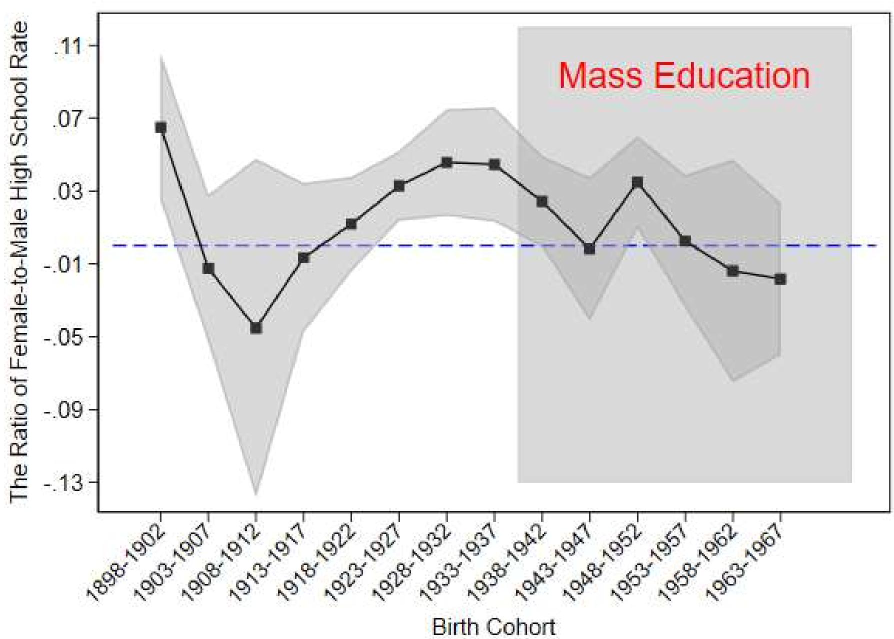
(b) The Ratio of Female-to-Male Literacy Rate

Figure A6: The OLS Estimates of Printing on Modern Educational Outcomes

Notes: (a) OLS estimates of the effect of woodblock-printed book density on the relative change in literacy rate (in percentage terms). (b) OLS estimates of the effect of woodblock-printed book density on the female-to-male literacy rate ratio. Shaded areas around the point estimates represent 95% confidence intervals. The grey region indicates cohorts exposed to the Mass Education Movement. Regressions control for all historical covariates listed in Specification 8. Observations are weighted by cohort population size. Standard errors are adjusted for spatial autocorrelation at the prefecture level.



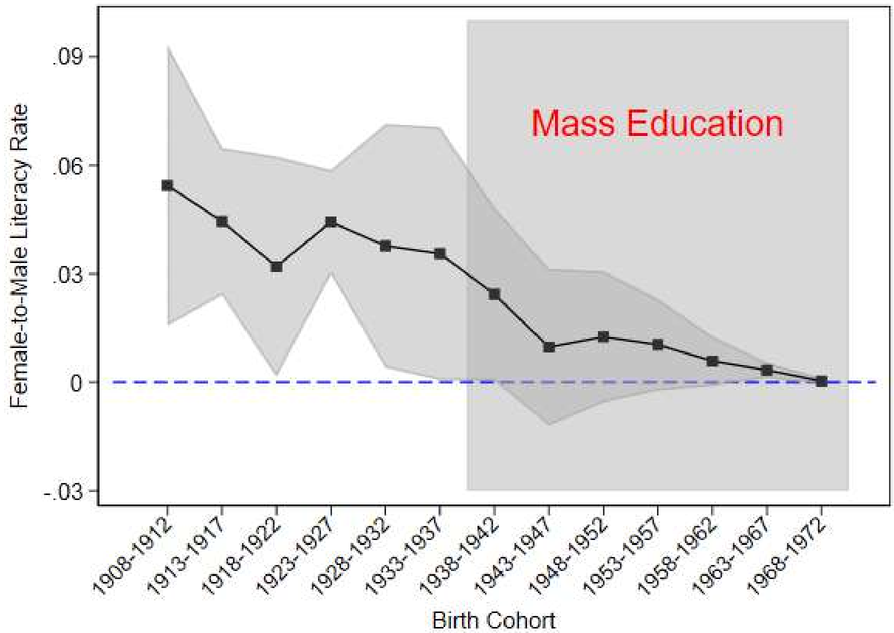
(a) Middle school



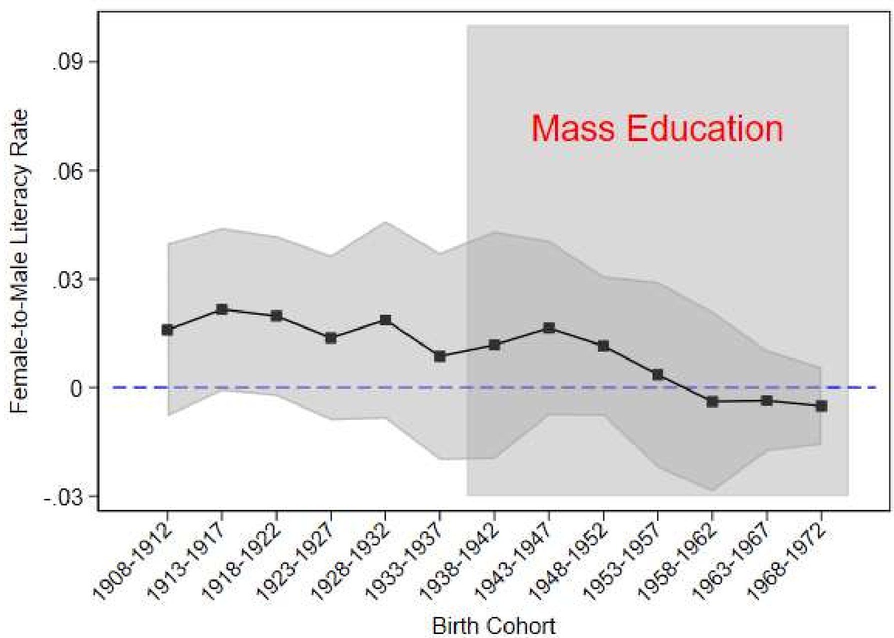
(b) High School

Figure A7: The 2SLS Estimates of Woodblock Printing on The Ratio of Female-to-male Educational Attainment

Notes: (a) The 2SLS estimates of the effect of woodblock-printed book density on the female-to-male middle school attainment ratio. (b) The 2SLS estimates of the effect of woodblock-printed book density on the female-to-male high school attainment ratio. The shaded area around the point estimates indicates 95% confidence intervals. The grey region highlights cohorts exposed to the Mass Education Movement. Regressions control for all historical covariates listed in Specification 8. Standard errors are adjusted for spatial autocorrelation across prefectures.



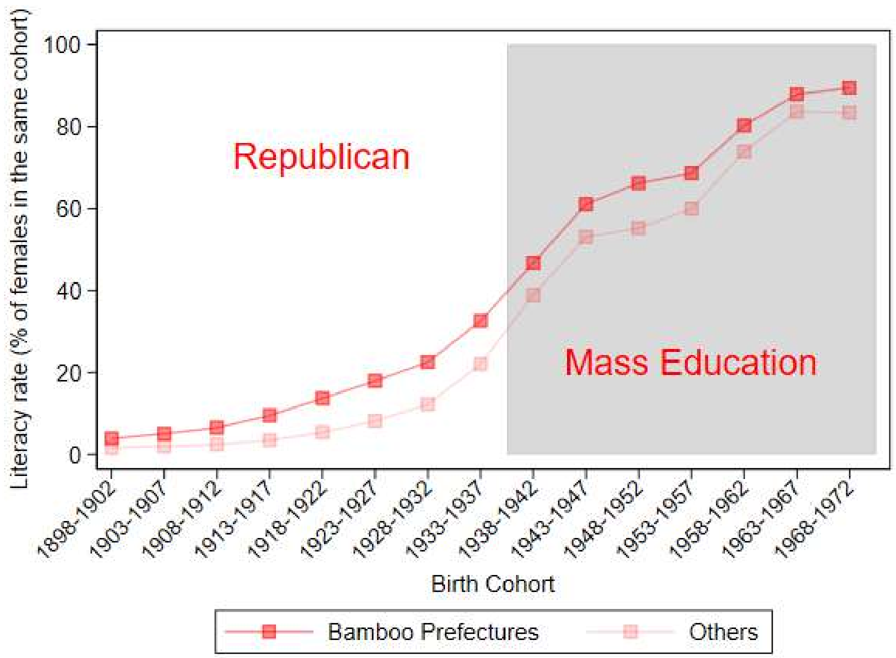
(a) Urban Population



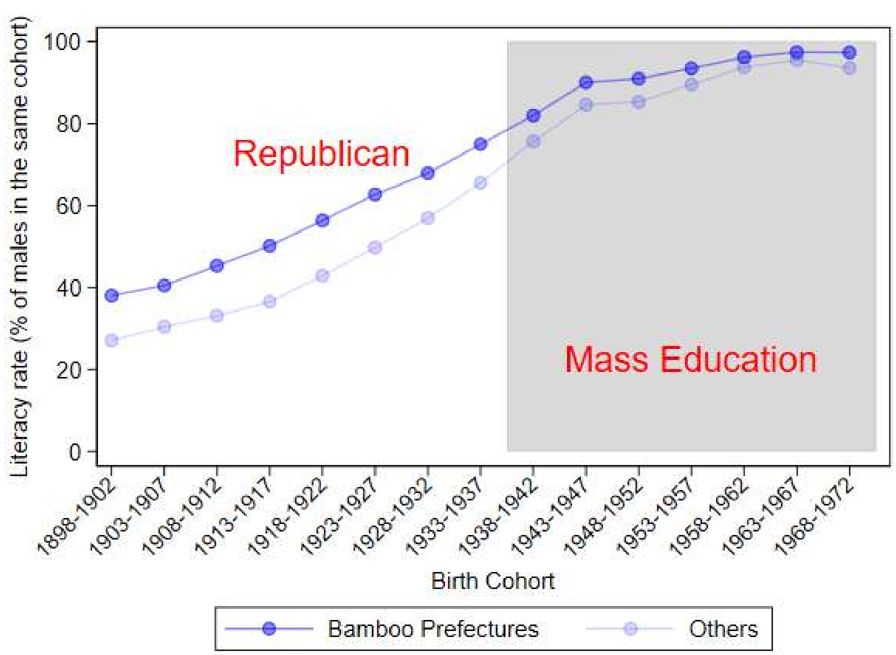
(b) Rural Population

Figure A8: The 2SLS Estimates of Woodblock Printing on The Ratio of Female-to-male Literacy Rate by Household Type

Notes: (a) The 2SLS estimates of the effect of woodblock-printed book density on the female-to-male literacy ratio among the urban population. (b) The 2SLS estimates of the effect of woodblock-printed book density on the female-to-male literacy ratio among the rural population. The analysis uses data from the 1990 Census, the census providing urban-rural classification. Shaded areas around the point estimates represent 95% confidence intervals. The grey region highlights the cohorts affected by the Mass Education Movement. Regressions include all historical controls specified in Specification 8. Standard errors are adjusted for spatial autocorrelation across prefectures.



(a) Female Literacy Rate



(b) Male Literacy Rate

Figure A9: The Trend of Literacy Rate: Bamboo Prefectures and Others

Note: (a) Average female literacy rates in bamboo and non-bamboo prefectures (b) Average male literacy rates in bamboo and non-bamboo prefectures. This figure plots cohort-level literacy rates by gender using data from the 1982 China Census. Prefectures are classified based on had predicted bamboo or not. Literacy rates are shown for five-year birth cohorts from 1898 to 1972. The shaded gray region denotes cohorts exposed to the Mass Education Movement launched under Mao's regime in 1949.

Table A1: Descriptive Statistics

	mean	sd	min	max	count
Time Variant Variables					
Ming Period (1368-1644)					
Log density of books in Ming	2.77	1.19	-0.19	7.01	268
Log density of female poets in Ming	1.52	1.07	-0.75	5.33	268
Log density of male poets in Ming	2.26	0.98	-0.05	5.33	268
Log density of population in 1393	4.80	1.39	-1.36	8.51	268
Log density of *jinshi* (degree holders) in Ming	4.32	1.29	0.46	6.80	268
Log density of schools in Ming	3.16	0.78	1.19	5.57	268
Qing Period (1644-1911)					
Log density of books in Qing	3.91	1.82	-0.13	9.18	268
Log density of female poets in Qing	1.34	1.08	-1.05	5.49	268
Log density of male poets in Qing	2.20	1.03	-0.27	5.64	268
Log density of population in 1680	5.74	1.49	0.42	8.48	268
Log density of *jinshi* (degree holders) in Qing	3.74	1.11	0.48	6.32	268
Log density of schools in Qing	2.73	0.68	0.63	5.40	268
Time Invariant Variables					
Geographical Controls					
Caloric suitability index	0.72	0.21	0.00	1.00	268
Terrain Ruggedness index	4.99	1.19	1.49	6.88	268
Log distance to coast in km	5.81	1.17	0.16	7.56	268
Log distance to river in km	2.12	1.41	0.06	5.74	268
Log area in km^2	9.33	0.83	6.83	12.20	268
Political Controls					
Log distance to national capital in 100km	2.45	0.57	0.00	3.21	268
Whether provincial capital	0.07	0.26	0.00	1.00	268

Source: See texts.

Table A2: Log book density and log river distance to bamboo or pine

	(1)
	Log density of books
Log river distance to bamboo	-0.440**
	(0.155)
Log river distance to pine	-0.215
	(0.127)
Observations	267
R^2	0.716
Mean	3.906
Baseline controls	Y
Geographical controls	Y
Political controls	Y

Note: The outcome variable is the log density of books published during the Ming–Qing period. Bamboo and pine data are sourced from Chen, Kung and Ma (2020) and calculated by the author using ArcGIS. Baseline controls include the log density of population in 1393, log density of schools, and log density of degree holders during the Ming–Qing period. Geographical controls comprise the caloric suitability index, terrain ruggedness index, log distance to the coast and navigable river, and log prefecture area. Political controls include log distance to the national capital and a dummy indicating provincial capitals. Regression include province fixed effects, and standard errors are clustered at the province level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Balance Table of Environmental Factors: Recorded v.s. Non-recorded Bamboo Sites

	(1)	(2)	(3)	(4)
Statistics		Means		Difference
Variable	Full sample	Bamboo=1	Bamboo=0	(2) vs (3)
Precipitation (in 1000 mm)	1.090	1.346	1.024	0.322***
	(0.434)	(0.313)	(0.437)	(0.063)
Organic carbon (% weight, g/kg)	0.493	0.564	0.474	0.089***
	(0.109)	(0.093)	(0.105)	(0.016)
Suitable soil pH	0.246	0.564	0.164	0.399***
	(0.432)	(0.501)	(0.371)	(0.061)
Elevation (in 1000 km)	0.736	0.333	0.840	-0.507***
	(0.775)	(0.283)	(0.826)	(0.113)
Observations	268	55	213	268

Note: This table summarizes four ecological factors used in predicting bamboo suitability. Precipitation is measured as average annual precipitation (in 1,000 mm), organic carbon as percentage weight (g/kg), soil pH as the average level within the optimal range (5.5–6.0), and elevation in thousands of kilometers. Column 1 presents the mean values across all 268 prefectures in China proper. Column 2 shows means for the 55 prefectures with bamboo presence recorded in provincial gazetteers, while Column 3 reports means for the remaining 213 prefectures without such records. Standard deviations are reported in parentheses. Column 4 presents the difference in means between Columns 2 and 3, with standard errors in parentheses. See text for data sources. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: The Estimates of Environmental Factors on Whether Had Bamboo in Gazetteers

	(1) Whether had recorded bamboo
Precipitation (in 1000 mm)	0.763**
	(0.327)
Organic carbon (% weight, g/kg)	2.243*
	(1.294)
Suitable soil pH	1.001***
	(0.222)
Elevation (in 1000 km)	-0.544***
	(0.191)
Observations	268
Mean	0.205

Note: The outcome variable is a binary variable equal to one if a prefecture had bamboo recorded in Qing-era provincial gazetteers. Precipitation is measured as average annual precipitation (in 1,000 mm); organic carbon is expressed as percentage weight (g/kg); suitable soil pH denotes the average soil pH level within the optimal range (5.5–6.0); and elevation is measured in thousands of kilometers. Robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Non-poetry Density and Poet Density: TWFE and 2SLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log density of female poets				**Log density of male poets**			
	TWFE			2SLS	TWFE			2SLS
Log density of non-poetry	0.442***	0.349***	0.319***	0.864***	0.292***	0.227***	0.217***	0.207
	(0.031)	(0.029)	(0.029)	(0.193)	(0.026)	(0.031)	(0.033)	(0.157)
Log density of *jinshi*	-0.083	0.025	0.012	-0.087	0.434***	0.501***	0.498***	0.500***
	(0.051)	(0.044)	(0.043)	(0.073)	(0.056)	(0.060)	(0.063)	(0.069)
Log density of schools	0.665***	0.639***	0.660***	0.270	-0.027	-0.010	-0.009	-0.002
	(0.127)	(0.088)	(0.084)	(0.183)	(0.082)	(0.088)	(0.084)	(0.144)
Log density of population	-0.010	-0.070	-0.078*	0.042	-0.054	-0.082	-0.093	-0.095
	(0.045)	(0.049)	(0.045)	(0.064)	(0.051)	(0.055)	(0.060)	(0.078)
Observations	536	536	536	536	536	536	536	536
R^2	0.857	0.878	0.882	0.788	0.829	0.838	0.843	0.843
Mean	1.431	1.431	1.431	1.431	2.221	2.221	2.221	2.221
First-stage F-stat				22.573				22.573
Geographical controls		Y	Y	Y		Y	Y	Y
Political controls			Y	Y			Y	Y

Note: Each column incrementally adds a set of control variables as described in the text. Geographic controls include the caloric suitability index, terrain ruggedness index, log distance to the coast, log distance to the nearest navigable river, and log prefecture area. Political controls consist of log distance to the national capital and a dummy indicating provincial capitals. All regressions include prefecture and dynasty fixed effects. Standard errors, adjusted for spatial autocorrelation at the prefecture level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: Western Printing and Poet Density: TWFE

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log density of female poets				**Log density of male poets**			
Log river distance to western printing	-0.851***	-0.340***	-0.176	-0.099	-0.527***	-0.230***	-0.126	-0.071
	(0.135)	(0.119)	(0.151)	(0.122)	(0.093)	(0.084)	(0.106)	(0.090)
Log density of books				0.329***				0.239***
				(0.030)				(0.034)
Log density of degree holders	0.009	0.108*	0.071	0.000	0.491***	0.554***	0.539***	0.487***
	(0.069)	(0.060)	(0.057)	(0.043)	(0.053)	(0.058)	(0.064)	(0.064)
Log density of schools	0.989***	0.891***	0.890***	0.674***	0.190**	0.154*	0.148*	-0.009
	(0.127)	(0.094)	(0.089)	(0.081)	(0.077)	(0.081)	(0.077)	(0.080)
Log density of population	-0.019	-0.119*	-0.134**	-0.065	-0.063	-0.113*	-0.130**	-0.080
	(0.054)	(0.063)	(0.056)	(0.046)	(0.062)	(0.059)	(0.060)	(0.062)
Observations	536	536	536	536	536	536	536	536
R^2	0.802	0.840	0.850	0.883	0.799	0.819	0.826	0.846
Mean	1.431	1.431	1.431	1.431	2.221	2.221	2.221	2.221
Geographical controls		Y	Y	Y		Y	Y	Y
Political controls			Y	Y			Y	Y

Note: Each column incrementally adds a set of control variables as described in the text. Geographic controls include the caloric suitability index, terrain ruggedness index, log distance to the coast, log distance to the nearest navigable river, and log prefecture area. Political controls consist of log distance to the national capital and a dummy indicating provincial capitals. All regressions include prefecture and dynasty fixed effects. Standard errors, adjusted for spatial autocorrelation at the prefecture level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: Woodblock-printed Books Density and Poet Density: Pooled OLS and 2SLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log density of female poets				**Log density of male poets**			
	Pooled OLS			2SLS	Pooled OLS			2SLS
Log density of books	0.368***	0.379***	0.367***	0.623***	0.241***	0.248***	0.225***	0.183
	(0.045)	(0.046)	(0.045)	(0.197)	(0.034)	(0.033)	(0.036)	(0.168)
Log density of *jinshi*	-0.170***	-0.074*	-0.080**	-0.216*	0.262***	0.338***	0.335***	0.358***
	(0.049)	(0.040)	(0.039)	(0.118)	(0.055)	(0.051)	(0.051)	(0.104)
Log density of schools	0.643***	0.366***	0.386***	0.304***	0.184**	-0.011	-0.005	0.009
	(0.102)	(0.110)	(0.108)	(0.117)	(0.081)	(0.081)	(0.080)	(0.090)
Log density of population	-0.108*	-0.243***	-0.246***	-0.227***	-0.077*	-0.163***	-0.183***	-0.186***
	(0.058)	(0.071)	(0.074)	(0.084)	(0.043)	(0.046)	(0.045)	(0.050)
Observations	536	536	536	536	536	536	536	536
R^2	0.543	0.624	0.627	0.562	0.514	0.566	0.573	0.571
Mean	1.431	1.431	1.431	1.431	2.221	2.221	2.221	2.221
First-stage F-stat				21.152				21.152
Geographical controls		Y	Y	Y		Y	Y	Y
Political controls			Y	Y			Y	Y

Note: Each column incrementally adds a set of control variables as described in the text. Geographic controls include the caloric suitability index, terrain ruggedness index, log distance to the coast, log distance to the nearest navigable river, and log prefecture area. Political controls consist of log distance to the national capital and a dummy indicating provincial capitals. All regressions include province and dynasty fixed effects. Standard errors, adjusted for spatial autocorrelation at the prefecture level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: Woodblock-printed Book Density and Whether Had a Poet: Probit and 2SLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Whether had any female poet				**Whether had any male poet**			
	Probit			2SLS	Probit			2SLS
Log density of books	0.517***	0.502***	0.487***	1.019***	-0.014	-0.027	-0.052	-0.631***
	(0.081)	(0.080)	(0.078)	(0.080)	(0.069)	(0.070)	(0.074)	(0.174)
Log density of *jinshi*	0.236**	0.100	0.092	-0.310**	0.825***	0.631***	0.629***	0.791***
	(0.103)	(0.108)	(0.108)	(0.127)	(0.113)	(0.123)	(0.122)	(0.100)
Log density of schools	-0.711***	-0.393**	-0.372**	-0.482***	-1.011***	-0.726***	-0.701***	-0.340
	(0.136)	(0.154)	(0.155)	(0.132)	(0.154)	(0.169)	(0.168)	(0.219)
Log density of population	0.151*	0.276***	0.273***	0.240***	0.142*	0.262***	0.238**	0.113
	(0.081)	(0.101)	(0.103)	(0.081)	(0.084)	(0.100)	(0.101)	(0.101)
Observations	536	536	536	536	490	490	490	490
Mean	0.412	0.412	0.412	0.412	0.684	0.684	0.684	0.684
Geographical controls		Y	Y	Y		Y	Y	Y
Political controls			Y	Y			Y	Y

Note: The dependent variable here is a binary variable that 1 indicates this prefecture had at least one female poet or male poet, otherwise 0. Each column incrementally adds a set of control variables as described in the text. Geographic controls include the caloric suitability index, terrain ruggedness index, log distance to the coast, log distance to the nearest navigable river, and log prefecture area. Political controls consist of log distance to the national capital and a dummy indicating provincial capitals. All regressions include province and dynasty fixed effects. Standard errors, adjusted for spatial autocorrelation at the prefecture level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: Woodblock-printed Book Density and Number of Poets: Poisson and 2SLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Number of female poets				**Number of male poets**			
	Poisson			2SLS	Poisson			2SLS
Log density of books	0.489***	0.418***	0.382***	1.810***	0.228***	0.170***	0.143***	-0.330
	(0.043)	(0.034)	(0.038)	(0.621)	(0.040)	(0.042)	(0.045)	(0.369)
Log density of *jinshi*	0.679***	0.599***	0.615***	-0.525	0.925***	0.902***	0.918***	1.162***
	(0.098)	(0.067)	(0.072)	(0.377)	(0.087)	(0.079)	(0.082)	(0.221)
Log density of schools	-0.944***	-0.561***	-0.547***	-1.064***	-1.237***	-0.936***	-0.940***	-0.916***
	(0.132)	(0.144)	(0.144)	(0.376)	(0.097)	(0.113)	(0.110)	(0.179)
Log density of population	0.269**	0.478***	0.441***	0.628***	0.298**	0.441***	0.405***	0.317***
	(0.124)	(0.096)	(0.100)	(0.218)	(0.149)	(0.109)	(0.113)	(0.121)
Observations	536	536	536	536	536	536	536	536
Mean	5.424	5.424	5.424	5.424	14.129	14.129	14.129	14.129
Geographical controls		Y	Y	Y		Y	Y	Y
Political controls			Y	Y			Y	Y

Note: The dependent variable here is number of poets in each prefecture. Each column incrementally adds a set of control variables as described in the text. Geographic controls include the caloric suitability index, terrain ruggedness index, log distance to the coast, log distance to the nearest navigable river, and log prefecture area. Political controls consist of log distance to the national capital and a dummy indicating provincial capitals. All regressions include province and dynasty fixed effects. Standard errors, adjusted for spatial autocorrelation at the prefecture level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: First-stage: River Distance to Predicted Bamboo and Historical Book Density

	(1)	(2)	(3)
		Log density of books	
Log river distance to predicted bamboo	-1.238***	-1.141***	-1.033***
	(0.085)	(0.328)	(0.271)
Observations	252	252	252
R^2	0.476	0.579	0.618
Mean	4.449	4.449	4.449
First-stage F-stat	40.133	12.108	10.569
Baseline Controls	Y	Y	Y
Geographical Controls		Y	Y
Political Controls			Y

Note: This table reports the first-stage estimate of interaction between log river distance to predicted bamboo and dummy indicating Qing on log density of books. Each column incrementally adds a set of control variables as described in the text. Baseline controls include log density of average population and schools in Ming-Qing period. Geographic controls include the caloric suitability index, terrain ruggedness index, log distance to the coast, log distance to the nearest navigable river, and log prefecture area. Political controls consist of log distance to the national capital and a dummy indicating provincial capitals. All regressions include province fixed effects. Standard errors, adjusted for spatial autocorrelation at the prefecture level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: DDD Estimates of the Impact of Bamboo on Intergenerational Transmission of Literacy

Location/Parent literacy	Illiterate	Literate	Family difference for location
A. Treatment Individuals: Women			
Bamboo prefectures	0.269	0.733	0.465***
	(0.055)	(0.019)	(0.057)
	[67]	[540]	
Non-bamboo prefectures	0.177	0.487	0.311***
	(0.017)	(0.012)	(0.024)
	[532]	[1765]	
Location difference for a given family	0.092*	0.246***	
	(0.050)	(0.024)	
Difference-in-difference	0.154**		
	(0.065)		
B. Control Individuals: Men			
Bamboo prefectures	0.619	0.909	0.290***
	(0.022)	(0.006)	(0.017)
	[504]	[1972]	
Non-bamboo prefectures	0.530	0.835	0.305***
	(0.008)	(0.004)	(0.008)
	[3663]	[7562]	
Location difference for a given family	0.089***	0.073***	
	(0.024)	(0.009)	
Difference-in-difference	-0.015		
	(0.022)		
DDD	**0.169*** **		
	(0.062)		

Note: This table reports all derived difference-in-differences (DiD) estimates and the derived triple-difference (DDD) estimate. Standard errors are shown in parentheses; the number of observations is indicated in brackets.
Source: The Third National Population Census 1982, and The Fourth National Population Census 1990.